Optical Image Evaluation

U. S. Department of Commerce

National Bureau of Standards

Circular 526
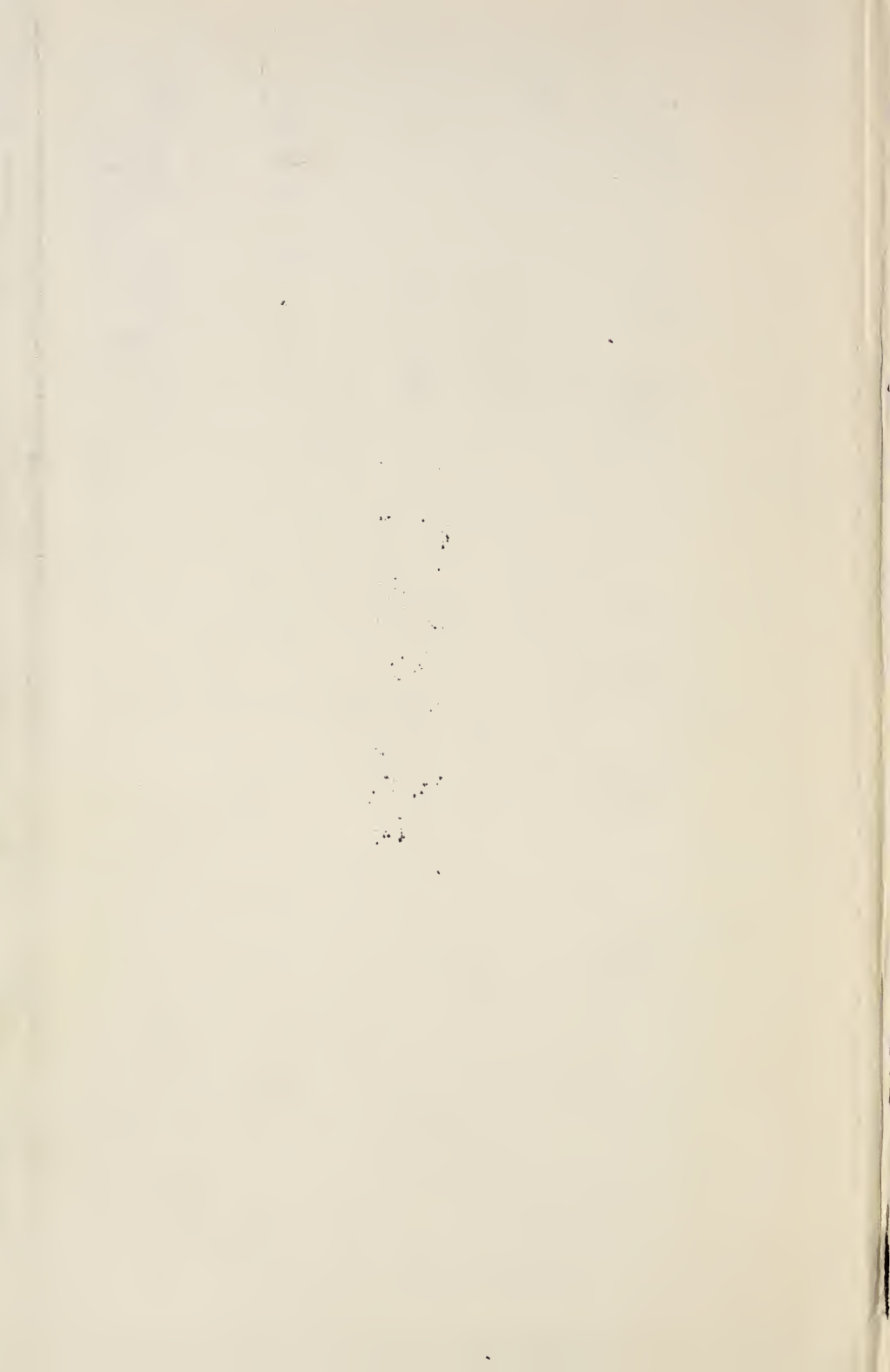
Optical Image Evaluation

Proceedings of the NBS Semicentennial Symposium on Optical Image Evaluation Held at the NBS on October 18, 19, and 20, 1951

Foreword

The Symposium on the Evaluation of Optical Imagery was one of the twelve symposia conducted as part of the scientific program of the National Bureau of Standards in the year 1951 which marked the fiftieth anniversary of its establishment. The subject for this symposium was chosen because of its current interest and because it is one in which the Bureau has been active for many years.

In the field of applied optics it is the generally accepted practice to evaluate optical designs on the basis of geometrical optics, and the performance of optical systems has often been based upon measurements of the geometric aberrations. This practice is justified when the aberrations are so large that diffraction plays but a small part in determining the quality of imagery. Now, however, better optical systems are being produced, automatic computing machines make it feasible to completely test an optical design by computation, the interferometer enables the wave front emergent from an optical system to be completely mapped, and integrating devices enable the diffraction effects to be readily and completely determined. With these contemporary developments it seemed timely to reexamine and compare the different methods of image evaluation with the purpose of placing them on a sound engineering basis and utilizing the principles of physical optics when justified.

The scientific excellence of the symposium derived from the quality of the papers and discussions presented by the participants. The generosity of the speakers in making their material available for publication in this volume is sincerely appreciated.

Acknowledgment is made to Dr. Wallace R. Brode, Dr. H. R. J. Grosch, Dr. A. Maréchal, Dr. Stanley S. Ballard, and Dr. Brian O'Brien who served as chairmen at the different sessions. The generous cooperation of the Office of Naval Research in making this symposium possible is also gratefully acknowledged.

A. V. Astin, *Director,*
National Bureau of Standards.

Contents

Introduction

By Irvine C. Gardner

Geometric optics is one of the older branches of optics, but the parts most directly related to optical design were for a long time not developed into a continuous body of scientific knowledge. This situation arose partly because the earlier designers were expert artisans or craftsmen who did not have the same facilities for communicating with each other as did the contemporary scientists and, also, partly because methods of design were viewed as trade secrets. Optical design as we know it begins with Fraunhofer (1787–1826) because prior to the discovery of the Fraunhofer lines there could be no precise measures of index of refraction. However, it is only in the last few decades that books have attempted to explain in detail how an optical system is designed. "A System of Applied Optics" by H. Dennis Taylor, is one of the earliest notable books that attempted conscientiously to set forth a method of designing an optical system. Even now the method of designing an optical system is not a well-established regular procedure. Successful designers are very often self-taught, each to a large extent has his own methods; and consequently optical design still remains to a considerable extent an art rather than a science.

If the design of an optical system is such that large aberrations play a predominant part in determining the distribution of energy in an image spot, the problem may lie entirely within the domain of geometrical optics. On the other hand for the microscope, the astronomical telescope, or the modern airplane-camera lens with small field angle and long focus, where the distribution of energy in the image spot is largely determined by diffraction, then the final judgment of the result should involve an appeal to physical optics. Reliance upon physical optics has not been general partly because of tradition and partly because in many instances it was not practicable to make the required extensive computations.

Designing an optical system by an established procedure is an engineering task. It has been customary in requesting the design of a new optical system, such as a camera objective, to specify performance characteristics as, for example, the aperture ratio or relative aperture, the focal length, the field of view, and resolving power. However, there is no generally accepted and established method of interpreting the quality of the image, as indicated by the results of ray tracing, or the computed diffraction effects in terms of resolving power. In most instances if good resolving power is required, the designer simply perfects the design as much as possible within the limits of time and cost and hopes that the resulting lens will perform as well as required. This laissez-faire attitude has persisted even up to the present time because usually a system designed solely on the basis of geometric considerations performs better than anticipated. This occurs because of compensations that arise from the physical nature of light. Although

this is the preferable direction in which an error of judgment should lie, it will be understood that too large a factor of safety may be expected to result in optical designs that are unnecessarily complicated or may deter one from designing systems giving a more spectacular performance than is considered possible on the basis of geometric optics. These considerations illustrate the need for a more completely engineered method of image evaluation. The need for this is further indicated when it is noted that recently serious doubt has arisen as to whether or not resolving power is always a correct criterion of lens performance. As is shown by some of the articles in this volume, there are instances in which the image plane of a photographic lens that produces the most desirable photograph is not in the plane of maximum resolving power but in a plane differently defined, which may be designated in a general way as the plane of maximum contrast, or the plane that produces photographs with the steepest density gradient.

Not all of these apparent lacks should be charged against the lens designer. The design of a complicated lens system requires an extremely large program of tedious and time-consuming computations and it has in most cases not been possible to extend this work beyond the stage at which a satisfactory lens is assured. Furthermore, experimental tools, by which the distribution of energy within an image spot could be accurately determined, have not been available. The recent availability of new experimental and computational tools has greatly changed the outlook. The photomultiplier tube, with its spectacular increase of sensitivity beyond that of previous devices used for the same purpose, makes possible measurements of image quality that were previously unattainable. The programmed electronic computing machine not only makes possible much more elaborate algebraic computations or ray tracings in order to arrive at more nearly perfect designs but it also offers possibilities of making extended calculations in connection with the evaluation of an image that were previously impracticable. It is now practicable to trace skew rays, a task that the lens designer has previously avoided, so far as permissible, because of the excessive labor. It is also possible to compute diffraction patterns, thereby evaluating the image quality in terms of physical optics. Thus optical design, one of the older and traditional branches of optics, has again been brought into a nascent state.

In connection with the recognition of the fiftieth anniversary of the establishment of the National Bureau of Standards it was considered appropriate to conduct a series of symposiums dealing with scientific subjects of current interest. In view of the indicated state of flux existing in the theory of optical design it was considered desirable to select some phase of this subject, and accordingly the Symposium on Optical Image Evaluation was planned. The fact that nearly 250 people from all parts of the United States and Europe fully occupied 3 days with the formal presentation of papers and with informal discussion indicates that the choice of subject was a fortunate one and that the renewed vitality of the subject as presented in the preceeding paragraphs is generally realized.

This interest in optical image evaluation is abundantly justified by its importance and complexity. In order to show briefly the different possible courses that image evaluation may take, reference may be made to figure 1 where two main branches are shown. The evalua-

IMAGE EVALUATION

MATHEMATICAL MODEL AVAILABLE

PHYSICAL MODEL AVAILABLE

(a) GEOMETRICAL INTERPRETATION

(b) CONTOUR MAP OF WAVE FRONT

(c) DIFFRACTION PATTERN (RESOLVING POWER)

(a') HARTMAN TEST, KNIFE EDGE TEST, ETC.

(b') LENS INTERFEROMETER TEST

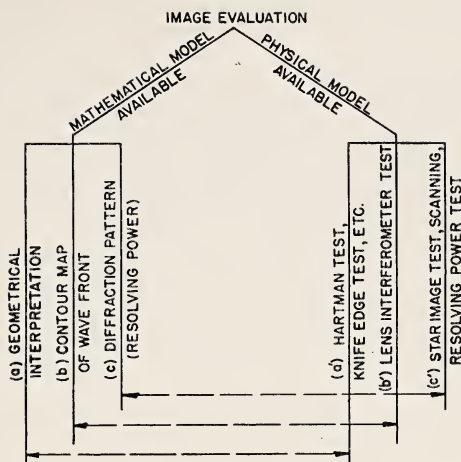(c') STAR IMAGE TEST, SCANNING, RESOLVING POWER TEST

FIGURE 1. *Diagram showing relation between different methods of image evaluation*

tion may be based upon a lens design before a prototype has been actually constructed. In this case one has a constructional specification of the lens embodying such parameters as radii of curvature, thickness, separation, and index of refraction etc. Following Dr. Herzberger, this will be referred to as a mathematical model of a lens. The course of rays through the lens may be determined by computations either by the older or the newer methods. It is now practicable to trace rays sufficient in number for any desired purpose. There are at least three possible evaluations: (a) Purely geometrical interpretation may be based on the distribution of the rays in the neighborhood of an image point; (b) the computation may be made in such a manner that a contour may result showing the distance measured in wavelengths between the emergent wavefront and the desired emergent wavefront, which is usually either plane or spherical; (c) the distribution of intensity in the diffraction pattern may be computed. This gives an interpretation of the image quality in terms of physical optics and usually involves a numerical integration that can be performed by a digital computer or an integration made by a mechanical analog machine. Having determined the diffraction pattern an estimate may be made of the resolving power.

Now follow the other of the two main stems. In this instance one has a prototype or a production sample of the lens and evaluates the quality of the image by an experimental method. The experimental method (a') most closely analogous to the method (a) of the preceding paragraph is the Hartman test and its variations or some variation of the Schlieren test by which the geometric paths of the rays are determined. The lens interferometer (method b') may be used to give information that is directly comparable with that afforded by the method (b). According to method (c') the star image may be photographed on a large scale to show the diffraction pattern, the distribution of energy within the diffraction pattern may be measured by one of the special scanning methods using the photomultiplier tube, or rather scanty information regarding the diffraction pattern may be

obtained by means of the resolving power test. Any of these procedures (method c′) offers results that may be compared with the results of paragraph (c).

The methods outlined permit the characteristics of the imagery to be determined. However, the evaluation, in the full sense of the word, is not completed until these measured characteristics are compared with the desired characteristics. As has been mentioned there are still differences of opinion concerning the characteristics that are desired. And when this subject is more thoroughly studied it will doubtless be found that different performance characteristics of imagery are desired for different purposes. By different purposes one refers not only to such manifest differences as those of visual instruments, photographic instruments, and projection instruments but to divisions of application that exist within these grand divisions. In photography, for example, one has airplane photography in which the object is conspicuously low in contrast, pictorial photography, and process work. In projection one has motion-picture projection where high magnification and good resolving power are required, and television projection for which the detail on the screen is necessarily limited.

This brief résumé gives an idea of the large field of engineering knowledge concerning image evaluation that remains to be filled in. The papers of this symposium touch upon most aspects of the problem in more or less detail but, like most useful scientific work, the papers also suggest the large amount of work that remains to be done.

1. The Diffraction Theory of Aberrations

By F. Zernike [1]

My starting point has been a new series development of the aberration function. Let V represent the optical path from an object point P through any optical system to its Gaussian image P', expressed as a function of the rectangular coordinates y and z of the point of intersection Q with the exit pupil (y in the axial plane). Customarily this is developed into a power series, general term $\alpha_{pq}y^pz^q$, leaving out the dependence on the axial distance of P, with which we are not specially concerned here. Of course y and z may as well be replaced by polar coordinates R and φ, making the general term $a_{kl}R^k\cos^l\varphi$. Opticians are quite accustomed to this development, so much so that specific names have been given to each of the lower terms, each being thought of as a *single aberration*. From a somewhat different problem I came in 1934 to consider a different development, rearranging the powers into polynomials, the general term $b_{nm}S_n^m(y, z)$ being homogeneous of the nth degree and also expressible as $b_{nm}R_n^m(r)\cos m\varphi$, in which $r=R/R_m$, where R_m is the radius of the circular exit pupil. These "circle polynomials" are perfectly determined by one condition, that of *orthogonality*, i. e.,

$$\int_0^1 R_n^m(r)R_{n'}^m(r)r\,dr=0 \qquad \text{for } n\neq n' \tag{1}$$

together with the normalizing condition that the maximum value of each shall be equal to one. Rather than give a table of these polynomials (see literature) I show their properties by figures 1.1 and 1.2, which give their course within the area of the pupil.

With the new development a *single aberration* shall mean an aberration characterized by a single term, i. e., by a circle polynomial. The remarkable advantages of this will appear presently.

Calculating the diffracted amplitude at a point in the receiving plane specified by polar coordinates ρ, ψ measured from P' as origin, one obtains, apart from irrelevant constant factors,

$$\int_0^1 R_n^m(r)r\,dr \int_0^{2\pi} \cos(\rho r\cos\varphi)\cos m(\varphi+\psi)d\varphi=$$

$$\cos m\psi \int_0^1 R_n^m J_m(\rho r)r\,dr = \pm\rho^{-1}J_{n+1}(\rho)\cos m\psi.$$

From this simple result a further development led to a new formula for the amplitude in the neighborhood of an aberrationless focus

$$u_0=\sum_{n=0}^\infty (-i)^n(2n+1)x^{-\frac{1}{2}}J_{n+\frac{1}{2}}(x)\rho^{-1}J_{2n+1}(\rho), \tag{2}$$

[1] Natuurkundig Laboratorium, Groningen, Netherlands.
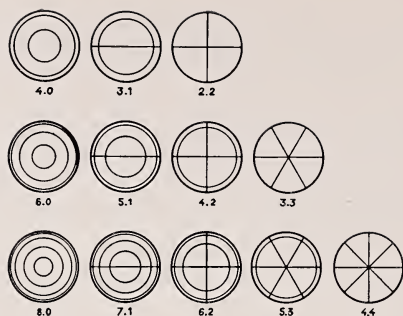
FIGURE 1.1. *Nodal lines for the complete circle polynomials.*

The first row belongs to the third order aberrations, respectively, spherical aberration, coma, astigmatism; the second row to the fifth order aberrations; the third row to the seventh-order aberrations.
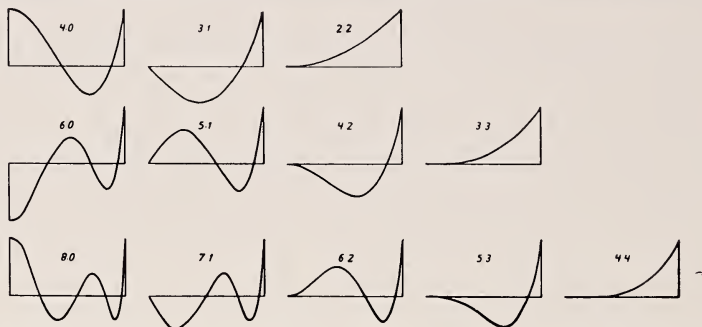
FIGURE 1.2. *Graphical representation of the radial parts, same arrangement as in figure 1.1.*

FIGURE 1.3. *Contour lines of equal intensity near an aberrationless focus.*

The numbers give the percentage of the value at the focus. The straight lines indicate the boundary of the geometric cone.

in which x measures the axial distance from the focal plane. This formula was checked numerically with those of Lommel (1886), which contain power series, different for points inside and outside the geometrical light cone. Figure 1.3 gives the resulting surfaces of equal intensity, the complete three-dimensional figure being generated by revolving the figure around the horizontal axis. A figure of this kind was first given by Berek in a rough but correct sketch from Lommel's results and afterwards repeatedly copied in the German literature with such errors that it finally appeared with marginal rays under an angle four times too small and with rectangular axes interchanged.

The next step is to introduce ordinary spherical aberration by adding a term βR_4^0 to V. Developing the diffraction integral into powers of the coefficient β, the first approximation is found to be

$$u_0 - i\beta\sum(-i)^n(2n+1)x^{-1/2}J_{n+1/2}(x)\rho^{-1}[a_nJ_{2n+5}+b_nJ_{2n+1}+c_nJ_{2n-3}],$$

where a_n, b_n, c_n are simple expressions in n. Representing the bracket expression as the result of an operation O_4 on J_{2n+1}, the same operation must be applied repeatedly to obtain the higher terms and the final series is represented symbolically by

$$u=\sum(-i)^n(2n+1)x^{-1/2}J_{n+1/2}(x)\rho^{-1}e^{-i\beta}O_4J_{2n+1}(\rho), \tag{3}$$

a result that is not only elegant but very useful. For higher order spherical aberrations one has only to substitute a different operator. Figures 1.4 and 1.5 give some results.

In order to judge the quality of an optical image, Strehl as early as 1895 introduced the brightness in the center of the diffraction image as compared to the same without aberrations. This is an efficient way of expressing the deterioration of the image through small aberrations by a single number, which we shall call the "Strehl Definition." For the best performance this should be 90 percent or more. Suppose there are a number of single aberrations present, with coefficients respectively, B_1, B_2, B_3, etc. The Strehl Definition SD will be

$$SD=1-B_1^2-B_2^2-B_3^2- \ \ . \ \ . \ \ .,$$
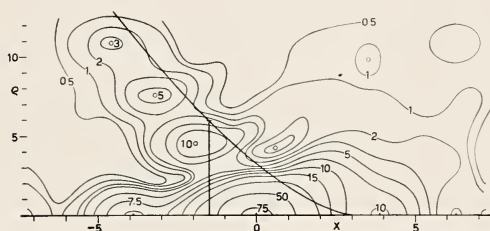
FIGURE 1.4. *Contour lines for third-order spherical aberration,* $\beta=1$, *Strehl Definition 0.81.*

The geometric caustic has been inserted, also the radius of the geometric circle of greatest constriction.
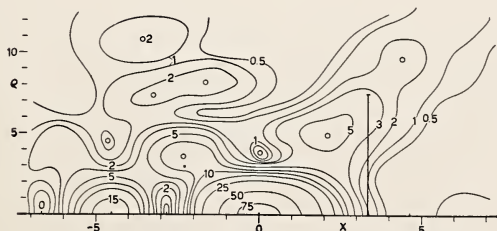

FIGURE 1.5. *Contour lines for fifth order-spherical aberration,* $\beta=1$, *Strehl Definition 0.86.*

The vertical line shows location and size of the circle of greatest constriction in the case of optimum balancing according to geometric optics.

3

where there are no cross products because of the orthogonality. This means that each separate aberration diminishes the *SD* on its own account, uninfluenced by others. In other words, there is no use in leaving a residual of lower aberrations in order to balance a higher one. Or expressed in the old way: Each circle polynomial contains next to the highest power of r a number of lower powers which balance it in the best way. It is worth while to dwell somewhat longer on the question of balancing and on the tolerances that result from these calculations.

The quarter-wave limit was announced in 1878 by Lord Rayleigh, who was then so much ahead of his time that his result was considered for at least 50 years as unsurpassed, and worse, as self-explanatory in the following way.

The waves that come together to a single focus should reinforce each other by being in the same phase. Because of aberrations the phases will differ, but not much harm will be done if the deviations from the mean are less than one quarter, because there will be no opposite components in that case.

However, Rayleigh did much better than this; he calculated the intensity in the paraxial focus for two types of aberration and remarked that it should at least be 80 percent. He omitted the effect of refocusing, which is very marked. As an instance I take the eighth-order spherical aberration, expressed in units of $\lambda/2\pi$ by $V = \beta_8 r^8$. This cannot well be corrected, even in a high-power microscope objective, whereas fourth- and sixth-power terms can be adjusted more or less at will by the designer. If he cancels the fourth and sixth powers the decrease in *SD*, which equals the integral of V^2 over the aperture, will be

$$\int_0^1 \beta_8^2 \left(r^8 - \frac{4}{5} r^2 + \frac{1}{5} \right)^2 2r\,dr = \frac{4}{225}\, \beta_8^2,$$

where the constant term 1/5 appears because of the necessary adjustment of phase, and the term with r^2 because of the adjustment to best focus. The coefficients 1/5 and $-4/5$ have been determined so as to make the integral a minimum. If now a residual of r^4 aberration is introduced for balancing, the three resulting coefficients will have to be calculated anew, with r^6 and r^4 aberrations four new coefficients. The calculation is more straightforward with the circle polynomials. It is easily found that r^8 can be represented by these functions

$$r^8 = \frac{1}{70} R_8 + \frac{1}{10} R_6 + \frac{2}{7} R_4 + \frac{2}{5} R_2 + \frac{1}{5}, \tag{4}$$

and the integrated square of this sum can be written down at once, as the cross products vanish, whereas R_n^2 gives $1/(n+1)$. The numerical result, brought to a common denominator, is

$$\frac{1}{210^2} (1 + 63 + 720 + 2352 + 1764). \tag{5}$$

The sum in parentheses equals 4,900 and the whole expression 1/9. However, as stated above, a constant should in all cases be added to (4) so as to minimize the integral and this clearly should be $-1/5$, thus canceling the term 1,764 in (5). This brings down (5) to 16/225,

4

corresponding to the unchanged (paraxial) focus. By adjusting the focus, the term with R_2 can be cancelled, leaving 784 as the sum in (5), result 4/225, same as stated above. Best balancing with the fourth order will in the same way leave only 64, with fourth and sixth orders only 1. These numbers are the squares of the successive binomial coefficients of the eighth power.

If we allow a tolerance of 10-percent decrease in Strehl Definition, we obtain for the maximum value of β_8, $(16/225)\beta_8^2 = 0.1$ or $\beta_8 = (15/4)0.316$ in the most unfavorable case, no balancing paraxial focus. This is the tolerance expressed in radians, a factor $\lambda/2\pi$ must be added to transform it to the usual linear measure. Calling the new coefficient b_8, we obtain successively

$$b_8 = 0.189\lambda, \ = 0.378\lambda, \ = 1.32\lambda, \ = 10.6\lambda, \tag{6}$$

respectively for no balancing and paraxial focus, no balancing and best focus, balancing with fourth order, balancing with fourth and sixth orders.

In a sense these results may be too favorable for the fully balanced error. We have started from the power series and as an example cut this off after the r^8 term, supposing the higher terms to be negligible. Then the term $b_8 r^8$ was changed to

$$\frac{1}{70} b_8 R_8 = \frac{1}{70} b_8 (70r^8 - 140r^6 + 90r^4 - 20r^2 + 1).$$

It is more consistent to judge the advantage obtained from the circle polynomials by starting from a development $V = \Sigma c_{2n} R_{2n}(r)$ and cutting this off after the term with R_8. I have compared a small number of practical examples (to be treated in a later paper) that show that the ratio of the coefficients $b_8/c_8 = 70$ is only approximated to if the power series converges very fast; if its successive coefficients decrease in a ratio of 10:1, it is 54, if 4:1 then 35. In this last case therefore, the tolerance for b_8 would still be 5.3 λ when fully balanced.

In view of these results, which are specially favorable for spherical aberration, it would seem preferable to state the tolerance for 10-percent decrease in Strehl Definition in the following general way, applicable to all kinds of aberrations:

$$maximum \ rms \ deviation \ of \ path = \lambda/(2\pi\sqrt{10}) = 0.0503 \ \lambda. \tag{7}$$

It is worth while to compare these numbers with the results of balancing according to geometrical optics. Taking once more the path-aberration $\beta_8 r^8$, this corresponds to a transverse aberration of the ray equal to its derivative, $8\beta_8 r^7$, expressed in diffraction units (in which the first dark ring has a radius of 3.83). Now the designer will endeavor to balance this by leaving lower order aberrations, so as to obtain a minimum "circle of greatest constriction" of the beam. For this well-known problem, one should take

$$\beta_8 \left(8r^7 - 14r^5 + 7r^3 - \frac{7}{8} r \right) = \frac{1}{8} \beta_8 T_7(r), \tag{8}$$

in which the polynomial T_n is defined by $T_n = \cos n\Phi$, if $\cos \Phi = r$.

5 .

Evidently, T_n oscillates between $+1$ and -1 in the interval $-1 \rightarrow +1$ of r. In our case, therefore, the greatest constriction has a radius of $1/8$ β_8. Now, it is hardly possible to find a tolerance for this in a consistent way. The conventional view seems to have been that a disk of this kind must be superimposed on the Airy disk. It could then at most be allowed to have a radius of 2, i. e., equal to the mean radius of the diffraction disk. The tolerance would hence be

$$1/8 \ \beta_8 = 2, \quad b_8 = 16 \ \lambda/2\pi = 2.5\lambda \quad \text{(geometric optics)},$$

or only one-fourth of the real diffraction tolerance! The practical result of this geometric balancing is remarkably good. To calculate it, we must obtain the path-aberration by integrating (8) giving

$$\beta_8 \left(r^8 - \frac{7}{3} \ r^6 + \frac{7}{4} \ r^4 - \frac{7}{16} \ r^2 \right) = \beta_8 \left(\frac{1}{70} \ R_8 - \frac{1}{60} \ R_6 - \frac{1}{168} \ R_4 - \ldots \right)$$

and find the integrated square of this, which is $(\beta_8/120)^2$. Hence the geometric disk of radius 2, and $\beta_8 = 16$, will in reality give a Strehl Definition of 98.2 percent or the tolerance for 90-percent definition is $b_8 = 6.0\lambda$, i. e., four-sevenths of the optimum.

Results like this, and comparable ones for the lower aberrations, explain why practical designers never had trouble to remain within tolerances, as they were apt to adopt far too strict ones. The great use made of the graphical representation of the *longitudinal* errors, initiated by von Rohr, has even led to balancing so as to make the longitudinal aberration oscillate between narrow limits. Even this gross error does little harm because of the beneficial influence of destructive interference. In short, the peculiar properties of the waves of light are very favorable to the lens designer, even if he ignores them altogether.

Coming now to the off-axis errors, I can only give a brief survey of one of these, astigmatism. The calculations were made by Nijboer, in much the same way as above. Figure 1.6 gives the contour lines of equal intensity for a small amount of astigmatism, which according to geometric optics would give rise to a "circle of least confusion" of radius 2. In reality the Airy disk proper is changed very little, the error having a marked influence only on the first and second dark rings, which are each split up into four dark spots. Figure 1.7 shows the case of a four times greater astigmatism. Here the Airy disk itself is split up into four maxima and the cross-shaped figure is very pronounced. This is the largest amount for which the series development with circle polynomials could be used, the final formula occupying a whole printed page.

Large amounts of various aberrations were studied experimentally by Nienhuis. Figure 1.8 gives one of his results. What puzzled us was that the figure did not seem to approximate to the evenly illuminated circle of geometric optics. A tentative explanation is as follows. At any point Q of the edge of the diaphragm some light will be diffracted, inward and outward, in directions perpendicular to the edge. The direct ray through Q will intersect the receiving plane in a point S on the geometric optical circle. If Q describes the edge of the pupil, S will describe the geometric circle *in the opposite direction*, as
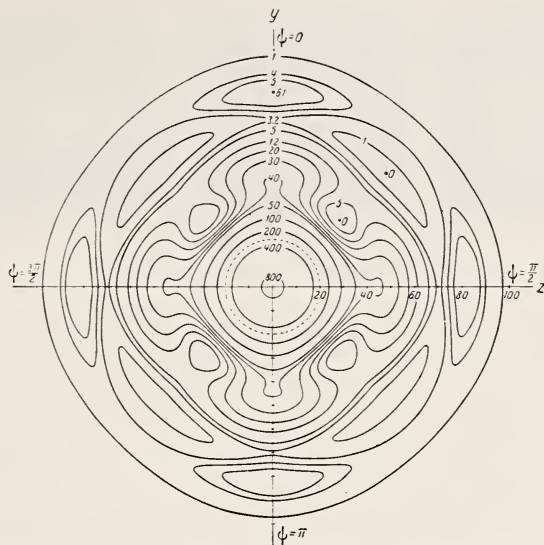
FIGURE 1.6. *Contour lines in the central plane for third-order astigmatism of amount β = 1.*

The numbers give the intensities corresponding to a value of 1,000 at the center of the Airy disk. The broken circle indicates the geometric disk.
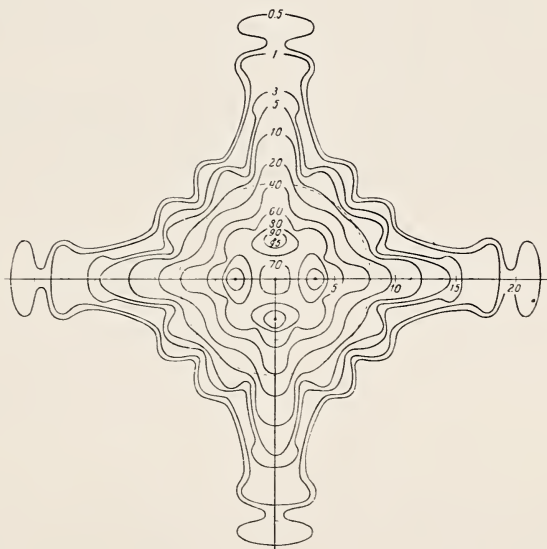
FIGURE 1.7. *Contour lines in the central plane for third-order astigmatism of amount β = 4.*

Maximum intensities of 0.095 at four off-axis spots. The broken circle should be evenly illuminated according to geometric optics.

FIGURE 1.8.   *Astigmatism, β=17.*
A, complete pattern; B, edge effect alone; C, pattern at focal line.

is well known.   As S moves, a narrow strip of diffracted light will therefore be seen radiating from S in a direction turning opposite to the radius through S.   These strips will all be tangent to a four-pointed curve, an asteroid, which indeed appears in figure 1.8,a, as the boundary of the illuminated pattern.   Nienhuis gave an experimental proof of this explanation by covering up most of the lens aperture, leaving only a narrow annulus at the margin. He thus obtained figure 1.8,b, in which the circle of least confusion has disappeared, leaving only an interference pattern caused by the overlapping of the diffracted rays. Practically the same is found in figure 1.8,c, where the geometric optical pattern has contracted into a focal line.

The last development in this matter is by N. G. van Kampen, who found asymptotic formulas for small wavelengths, which confirm these experimental results.

F. Zernike, Physica **1,** 689 (1934).
B. R. A. Nijboer, Thesis Groningen (1942).   Physica **10,** 679 (1943); **13,** 605 (1947).
K. Nienhuis, Thesis Groningen (1948).
F. Zernike and B. R. A. Nijboer, Contribution in La Theorie des Images Optiques, p. **227** (Paris, 1949).
K. Nienhuis and B. R. A. Nijboer, Physica **14,** 590 (1949).
N. G. van Kampen, Physica **14,** 575 (1949); **16,** 817 (1950).
E. Wolf, Reports on Progress in Physics **14,** 95 (1951).

2. The Contrast of Optical Images and the Influence of Aberrations

By André Maréchal [1]

Introduction

The estimation of the quality of optical images is mainly based on the knowledge of their contrast, i. e., the relative variation of the illumination in the image. In most cases improving the quality means increasing the contrast. Even in the vicinity of the resolution limit, any gain in contrast is helpful because it may slightly improve the resolution.

To begin with, we shall consider the mechanism of the formation of optical images in order to determine the contrast. Then we shall compute the losses of contrast produced by the presence of small aberrations in the images of some classical objects.

Theoretical Study of the Contrast of Optical Images

The Formation of Images

We generally define contrast as being the ratio $(I-i)/i$, I and i being the maximum and minimum illumination in the image. This definition does not hold for a bright point source or a bright line source, but it is very useful in the case of a dark point or a dark line in an extended background, or in the case of a periodic structure. The distribution of light in the image of a dark line is represented by the solid curve in figure 2.1, and the distribution in the object is shown by the dotted curve.

Consider an instrument where Ox is the axis, Oyz the object plane, and $O'y'z'$ the image plane (fig. 2.2). We can refer the exit pupil to angular coordinates β, γ with origin at O'. We can now determine the laws of formation of the images and study at least the two extreme cases of coherent and incoherent illumination.

Incoherent Illumination

Suppose that every point of the object acts as an independent source (as, for instance, in the use of astronomical instruments). The image of the point source will be a diffraction pattern, sometimes influenced by the presence of aberrations, and the distribution of light in the image will be obtained by adding the illuminations produced by the different points.

Consider a point source at $O(y=z=0)$. The distribution of light in its image is given by the laws of diffraction, i. e., the Huygens-
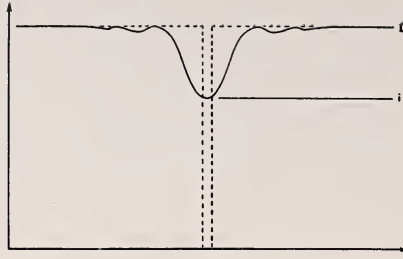
FIGURE 2.1. *Contrast reduction in the image of a dark line.*
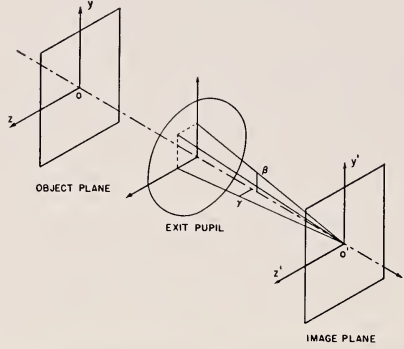
FIGURE 2.2. *Geometrical orientation of the object, lens, and image*

Fresnel principle, which can be expressed as follows: If $\Delta(\beta,\gamma)$ is the wave distortion produced by aberrations, the complex amplitude in y', z' will be

$$A(y',z') = \iint_P \exp\left[j\,\frac{2\pi}{\lambda}\,(\Delta + \beta y' + \gamma z') \right] d\beta d\gamma, \tag{1}$$

where λ is the wavelength.

The complex amplitude $A(y',\,z')$ appears as the Fourier transform of

$$\exp\left[j\,\frac{2\pi}{\lambda}\Delta \right].$$

If we denote now by

$$D(y',z') = |A(y',z')|^2$$

the distribution of light in the diffraction image, the resulting illumination for the whole object will be

$$I(y',z') = \iint_0 D(y'-y, z'-z)O(y,z)dydz, \tag{2}$$

where $O(y,\,z)$ represents the distribution of light in the object. Let us now consider some classical cases.

10

(a) *Dark point in a bright field:* $O(y, z)$ is unity everywhere, except in a small surface s around the origin O. The illumination will be

$$I(y', z') = \iint D(y'-y, z'-z)dydz - sD(y', z')$$

$$= \iint D(y', z')dy'dz' - sD(y', z'). \tag{3}$$

The first part is obviously a constant proportional to the total energy in the diffraction image of a point source, and we have to subtract from this constant the quantity $sD(y', z')$, as shown in figure 2.1. The contrast varies then as the function $D(y', z')$. This means that any improvement shown in it will be the same as for the case of the image of a single bright point.

(b) *Bright line:* Suppose that a bright line source of width ϵ is extended along the Oz axis. $O(y, z)$ is zero everywhere except for $0 < y < \epsilon$; the illumination is then

$$I(y', z') = \epsilon \int_{-\infty}^{+\infty} D(y', z'-z)dz = \epsilon \int_{-\infty}^{+\infty} D(y', z')dz' = \epsilon S(y'), \tag{4}$$

which is a function of y' only.

If we represent by a solid the function $D(y', z')$, the value at any point y' of the illumination $\epsilon S(y')$ would be proportional to the area of the plane section of that "diffraction solid" where it is intersected by a plane through the point and perpendicular to the y' axis.

(c) *Dark line:* The illumination will be obviously complementary to the preceding

$$I(y', z') = \iint D(y', z')dy'dz' - \epsilon \int D(y', z')dz'. \tag{5}$$

(d) *Edge of a bright area:*

$$O(y, z) \begin{cases} 0 \text{ if } y < 0 \\ 1 \text{ if } y > 0, \end{cases}$$

$$I(y', z') = \int_0^\infty dy \int_{-\infty}^{+\infty} D(y'-y, z'-z)dz = \int_0^\infty S(y'-y)dy.$$

We will define the contrast as the slope of the curve $I(y')$: $dI/dy' = S(y')$. This signifies that *when the illumination is incoherent, the contrast of the images is related either to $D(y', z')$ when the object is a point, or to $S(y')$ when the object is a line or the edge of an area.*

Coherent Illumination, Phase Contrast

Now suppose that the vibrations from the object are no longer independent of one another, but that they are related, the complex amplitude being a function $F_o(y, z)$. To determine the vibrations in the pupil, we have to compute in any direction from the object the resulting amplitude by applying the Huygens-Fresnel principle. The

11

mathematical expression is the transformation giving the amplitude in the pupil.

$$G(\beta,\gamma)=\iint_O F_0(y,z)\ \exp\left[-j\frac{2\pi}{\lambda}(\beta y+\gamma z)\right]dydz,$$

and the amplitude in the image will be

$$F_i(y',z')=\iint_P G(\beta,\gamma)\ \exp\left[j\frac{2\pi}{\lambda}(\Delta+\beta y'+\gamma z')\right]d\beta d\gamma$$

$$=\iint_P \exp\left[j\frac{2\pi}{\lambda}(\Delta+\beta y'+\gamma z')\right]d\beta d\gamma$$

$$\iint_O F_0(y,z)\ \exp\left[-j\frac{2\pi}{\lambda}(\beta y+\gamma z)\right]dydz.$$

By reversing the order for the integrations, the latter becomes

$$F_i(y',z')=\iint_O F_0(y,z)dydz\iint_P \exp\left[j\frac{2\pi}{\lambda}\left\{\Delta+\beta(y'-y)+\gamma(z'-z)\right\}\right]d\beta d\gamma$$

(6)

$$=\iint_{O_\xi} F_0(y,z)A(y'-y,z'-z)dydz.$$

This relation is very similar to (2) where the intensities are now replaced by the complex amplitudes.

Let us now examine various cases, beginning with the case of a uniformly illuminated object, in order to know the amplitude in a coherent background.

(a) *Uniform surface*: Let $F_o(y,z)\equiv 1$ everywhere;

$$F_i(y',z')=\iint A(y'-y,\ z'-z)dydz,$$

which can be written as

$$F_i(y',z')=\iint A(y',z')dy'dz'.$$

We have already noticed that $A(y',z')$ is the Fourier transform of $\exp j(2\pi/\lambda)\Delta(\beta,\gamma)$. We can write the reverse transform as follows:

$$\exp\left[j\frac{2\pi}{\lambda}\Delta(\beta,\gamma)\right]=\iint A(y',z')\ \exp\left[-j\frac{2\pi}{\lambda}(\beta y'+\gamma z')\right]dy'dz'.\quad(7)$$

If we choose $\beta=\gamma=0$ we obtain the value of

$$F_i(y',z')=\iint A(y',z')dy'dz'=\exp\left[j\frac{2\pi}{\lambda}\Delta(0,0)\right],$$

which is a constant over the field.

We can now use the convention that $\Delta(0,0)=0$, so that

$$F_i(y',z')\equiv 1.$$

(b) *Dark point in a coherent background:* We have to subtract from the preceding result the contribution of a point located at O and whose area will be $F_i(y',z') = 1 - sA(y',z')$.

The intensity will be

$$|F_i(y',z')|^2 = (1-sA)(1-sA^*) \simeq 1-s(A+A^*) = 1-2sR[A(y',z')], \quad (8)$$

where $R(A)$ means the real part of the complex amplitude $A(y',z')$. The contrast of the image will be expressed by $R(A(y',z'))$.

(c) *Bright coherent line:* Suppose that the line is along Oz, ϵ being its width. Then

$$F_i(y',z') = \epsilon \int A(y',z'-z)dz = \epsilon a(y'),$$

where

$$a(y') = \int A(y',z')dz'.$$

The intensity will be:

$$\epsilon^2 a(y')a^*(y'). \quad (9)$$

(d) *Dark line in a coherent background:* We must subtract the amplitude corresponding to the line from the uniform amplitude.

$$F_i(y',z') = 1 - \epsilon a(y').$$

The intensity will be:

$$1 - 2\epsilon R[a(y')],$$

and the contrast will be related to $R[a(y')]$.

(e) *Phase contrast objects:* Phase objects observed with phase contrast techniques behave like amplitude objects observed in coherent illumination. The results obtained in b, c, and d can be applied to them.

Expressions of Contrast Implying Various Pupillar Integrals (for Small Aberrations)

It is very useful to express the various quantities that govern the contrast of the image by integrations performed on the pupil. It will then be possible to study the effect of aberrations, whose nature will be given by the wave distortion $\Delta(\beta,\gamma)$. To obtain these expressions, we use general properties of the Fourier transform. Because of Professor Duffieux's work,[2] such transformations are rather easy. Furthermore, if we limit ourselves to small aberrations, we will be able to use the following expansions:

$$\exp\left[j\frac{2\pi}{\lambda}\Delta\right] = 1 + j\frac{2\pi\Delta}{\lambda} - \frac{2\pi^2}{\lambda^2}\Delta^2; \quad \cos\frac{2\pi}{\lambda}\Delta = 1 - \frac{2\pi^2}{\lambda^2}\Delta^2.$$

[2] P. M. Duffieux, L'intégrale de Fourier et ses applications à l'Optique (Société Anonyme des Imprimerie Oberthur, Rennes, 1947).

13

Points

a. *Bright point or dark point in an incoherent background:*
The illumination for a bright point will be

$$|A(y',z')|^2 = \left| \iint \exp\left[j\frac{2\pi}{\lambda}(\Delta+\beta y'+\gamma z') \right] d\beta\, d\gamma \right|^2.$$

We can conventionally introduce the linear terms $\beta\gamma'+\gamma z'$ into $\Delta(\beta,\gamma)$ and limit our study to $A(0,0)$, keeping in mind that Δ has to include a linear form in β and γ.
When Δ is small, and if $d\omega = \dfrac{d\beta\, d\gamma}{\iint d\beta\, d\gamma}$, then

$$|A(0,0)|^2 = \left[\iint\left(1+j\frac{2\pi}{\lambda}\Delta-\frac{2\pi^2}{\lambda^2}\Delta^2 \right)d\omega \right]\left[\iint\left(1-j\frac{2\pi}{\lambda}\Delta-\frac{2\pi^2}{\lambda^2}\Delta^2 \right)d\omega \right],$$

or

$$D(0,0) = |A(0,0)|^2 = 1 - \frac{4\pi^2}{\lambda^2}\left[\iint\Delta^2 d\omega - \left(\iint\Delta\, d\omega \right)^2 \right]. \quad (11)$$

b. *Dark point in a coherent background:* The real part of $A(0,0)$ will be expressed by

$$\iint \cos\frac{2\pi}{\lambda}\Delta\, d\beta\, d\gamma.$$

For small aberrations we will write

$$R(A(0,0)) = 1 - \frac{2\pi^2}{\lambda^2}\iint\Delta^2 d\omega.$$

Lines

a. *Incoherent illumination:* We evaluate

$$\int D(0,z')dz' = \int |A(0,z')|^2 dz'.$$

We notice that

$$A(0,z') = \int \exp\left[j\frac{2\pi}{\lambda}\gamma z' \right]d\gamma \int \exp\left[j\frac{2\pi}{\lambda}\Delta(\beta,\gamma) \right]d\beta$$

appears as the one-dimensional Fourier transform of

$$\varphi(\gamma) = \int \exp\left[j\frac{2\pi}{\lambda}\Delta(\beta,\gamma) \right]d\beta.$$

Applying the Plancherel theorem (which expresses mathematically the equality of the energy in the pupil and in the image when con-

14

sidered with two parameters) we can write

$$\int |A(0,z')|^2 dz' = \int |\varphi(\gamma)|^2 d\gamma.$$

For small aberrations this expression can be written

$$\int D(0,z')dz' = \int (\beta_2 - \beta_1)^2 \left[1 - \frac{4\pi^2}{\lambda^2} \left(\int \frac{\Delta^2 d\beta}{\beta_2 - \beta_1} - \left(\int \frac{\Delta d\beta}{\beta_2 - \beta_1} \right)^2 \right) \right] d\gamma,$$

$$(12)$$

where $\beta_1(\gamma)$ and $\beta_2(\gamma)$ are the lower and upper limits of the pupil.

(b) *Coherent line:* We evaluate $a(y') = \int A(y', z')\, dz'$.

Using eq 7 and making $\gamma = 0$,

$$\exp j\, \frac{2\pi}{\lambda}\, \Delta(\beta,0) = \int \exp\left[j\, \frac{2\pi}{\lambda}\, \beta y' \right] dy' \int A(y',z')dz',$$

which shows that $a(y')$ is the Fourier transform of

$$\exp\left[j\, \frac{2\pi}{\lambda}\, \Delta(\beta,0) \right].$$

The reverse transformation can be written

$$a(y') = \int \exp\left[j\, \frac{2\pi}{\lambda}\, \Delta(\beta,0) \right] \exp\left[-j\, \frac{2\pi}{\lambda}\, \beta y' \right] d\beta,$$

in which we will take $y' = 0$

$$a(0) = \int \exp\left[j\, \frac{2\pi}{\lambda}\, \Delta(\beta,0) \right] d\beta. \qquad (13)$$

According to the case to be studied (bright or dark line) we will use one of the two expressions:

$$|a(0)|^2 = \int \exp\left[j\, \frac{2\pi}{\lambda}\, \Delta(\beta,0) \right] d\beta \int \exp\left[-j\, \frac{2\pi}{\lambda}\, \Delta(\beta,0) \right] d\beta,$$

$$R\,[a(0)] = \int \cos \frac{2\pi\Delta(\beta,0)}{\lambda}\, d\beta.$$

For small aberrations these expressions can be written

$$|a(0)|^2 = (\beta_2 - \beta_1)^2 \left[1 - \frac{4\pi^2}{\lambda^2} \left\{ \int \frac{\Delta^2(\beta,0)}{\beta_2 - \beta_1}\, d\beta - \left(\int \frac{\Delta(\beta,0)}{\beta_2 - \beta_1}\, d\beta \right)^2 \right\} \right], \quad (14)$$

$$R\,[a(0)] = (\beta_2 - \beta_1) \left[1 - \frac{2\pi^2}{\lambda^2} \int \frac{\Delta^2(\beta,0)}{\beta_2 - \beta_1}\, d\beta \right]. \qquad (15)$$

15

Application to Determination of Tolerances for Various Aberrations

If the pupil is circular, polar coordinates may be useful:

$$\beta = \alpha \cos \phi$$

$$\gamma = \alpha \sin \phi.$$

Let h be the relative aperture α/α_m Then we may write the distortion of the wave front as follows:

$$\Delta = Dh^2 + Sh^4 + S'h^6 + (kh + ch^3 + c'h^5) \cos(\phi - \phi_0) + Ah^2 \cos 2(\phi - \phi_0),$$

in which D is the defect caused by errors in focusing, S is the third-order spherical aberration, S' is the fifth-order spherical aberration, k represents the linear expression in β and γ to be introduced in Δ (lateral displacement of the image in the ϕ_0 direction), C is the third-order coma in the ϕ_0 direction, C' is the fifth-order coma in the ϕ_0 direction, A is the astigmatism in the ϕ_0 direction, and h will vary from 0 to 1.

The results of the integrations are quadratic polynomials with respect to the preceding coefficients. The relative losses of contrast are the following:

<div style="text-align:center">INCOHERENT ILLUMINATION</div>

(a) *Point:*

$$\frac{4\pi^2}{\lambda^2}\left[\frac{D^2}{12} + \frac{DS}{6} + \frac{4}{45}S^2 + \frac{3}{20}DS' + \frac{1}{6}SS' + \frac{9}{112}S'^2\right.$$

$$\left. + \frac{k^2}{4} + \frac{kc}{3} + \frac{c^2}{8} + \frac{kc'}{4} + \frac{cc'}{5} + \frac{c'^2}{12} + \frac{A^2}{6}\right].$$

(b) *Line:*

$$\frac{3\pi^2}{4\lambda^2}\,[0.33962S^2 + 0.30542S'^2 + 2\times0.31992DS$$

$$+ 2\times0.28474DS' + 2\times0.31691SS'$$

$$+ \sin^2\phi(0.016549c^2 + 0.038678cc' + 0.023919c'^2)$$

$$+ \cos^2\phi(0.65160k^2 + 0.85639kc + 0.63653kc'$$

$$+ 0.31827c^2 + 0.50607cc' + 0.20991c'^2)$$

$$+ 0.32508(A\cos 2\phi + D)^2 + 0.812A^2\sin^2 2\phi].$$

<div style="text-align:center">COHERENT ILLUMINATION</div>

(a) *Point:*

$$\frac{2\pi^2}{\lambda_2}\left[\frac{D^3}{3} + \frac{SD}{2} + \frac{S^2+2DS'}{5} + \frac{SS'}{3} + \frac{S'^2}{7} + \frac{k^2}{4} + \frac{kc}{3} + \frac{kc'}{4} + \frac{c^2}{8} + \frac{cc'}{5} + \frac{c'^2}{12} + \frac{A^2}{12}\right].$$

(b) *Dark line:*

$$\frac{2\pi^2}{\lambda^2}\left[\frac{(D+A\cos 2\phi)^2}{5}+\frac{S^2}{9}+\frac{S'^2}{13}+\frac{2S(D+A\cos 2\phi)}{7}+\frac{2S'(D+A\cos 2\phi)}{9}\right.$$

$$\left.+\frac{2SS'}{11}+\cos^2\phi\left(\frac{c^2}{7}+\frac{c'^2}{11}+\frac{k^2}{3}+\frac{2cc'}{9}+\frac{2kc}{5}+\frac{2kc'}{7}\right)\right].$$

(c) *Bright line:*

$$\frac{4\pi^2}{\lambda^2}\left[\frac{4(D+A\cos 2\phi)^2}{45}+\frac{16S(D+A\cos 2\phi)}{15\cdot 7}+\frac{16S^2}{9.25}+\frac{8S'(D+A\cos 2\phi)}{9\cdot 7}\right.$$

$$\left.+\frac{48SS'}{5\cdot 7\cdot 11}+\frac{36S'^2}{13\cdot 49}+\cos^2\phi\left(\frac{k^2}{3}+\frac{2kc}{5}+\frac{c^2}{7}+\frac{2kc'}{7}+\frac{2cc'}{9}+\frac{c'^2}{11}\right)\right].$$

Various problems can now be solved. We have studied the following ones:

a. The effect of a given aberration alone, and the determination of the maximum value of that aberration such that the relative loss of contrast does not exceed the conventional value of 0.2 (average loss of contrast for a point in the case of the Rayleigh limit). If, for instance, we consider the effect of third-order spherical aberration alone (S term) on the contrast of an incoherent line, we write

$$\frac{3\pi^2}{4\lambda^2}\ 0.339S^2\leq 0.2\ \text{ or }\ |S|\leq 0.28\lambda.$$

b. For high-order aberrations we have to study the best way of balancing the high-order component with lower order terms. For example, in the case of "corrected" spherical aberration and its effect on the contrast of a coherent bright line, we write

$$\frac{4\pi^2}{\lambda^2}\left[\frac{4}{45}\ D^2+\frac{16}{15\times 7}\ DS+\cdots\right]\leq 0.2,$$

from which we can determine the proper values of D and S leading to the minimum value of the polynomial. These values are $D=(5/11)S'$ and $S=-(15/11)S'$. If we compute the minimum as a function of S' we obtain $S'<3.7\lambda$. The relation $S=-(15/11)S'$ determines the aperture of correction of the curve of longitudinal spherical aberration (fig. 2.3). By using the relations between longitudinal and wave spherical aberration, it can be shown that the aperture h_0 of correction will be $\sqrt{v_0}=\sqrt{2/3\cdot S/S'}$ times the maximum aperture h_m (v_0 is the parameter $(h_0/h_m)^2$).

Let us now consider the case of coma. As is well known, the distortion of the wave front is: Image magnitude times angular aperture times (offense against sine condition minus (longitudinal spherical aberration/distance pupil image). It is interesting to know what maximum value of the fifth-order term will be tolerable and what optimum aperture of correction is to be chosen. In the case of the

TABLE 2.1. *Values of various aberrations leading to a loss of contrast of 0.2*

		Incoherent illumination			Coherent illumination					
		Point	Line (φ=0)	Line (φ=π/2)	Point	Dark line (φ=0)	Dark line (φ=π/2)	Bright line (φ=0)	Bright line (φ=π/2)	
Defect of focusing	D |VI|	0.25 λ	0.29 λ		0.17 λ	0.22 λ		0.24 λ		
Third-order spherical aberrations {alone	S |VI‖VI|	0.25 λ S	0.28 λ S		0.22 λ S	0.30 λ S		0.27 λ		
{plus defect of focusing	D	−1.00 S	−0.98 S		−0.75 S	−0.71 S		−0.86 S		
	S	0.95 λ	1.04 λ		0.90 λ	1.00 λ		0.92 λ		
Third- plus fifth-order spherical aberrations, plus defect of focusing	r₀ ‖VI	1.00	1.00		0.89	0.85		0.91		
	S′	3.75 λ	4.00 λ		4.00 λ	4.50 λ		3.70 λ		
Astigmatism {φ=0 or φ=π/2	A |VI|VI|VI|	0.17 λ	-----	0.29 λ	0.25 λ	0.22 λ		0.24 λ		
	D±A	-----	-----		-----					
{φ=π/4	A |VI‖VI‖VI|	0.17 λ	-----	0.18 λ	0.25 λ					
Third-order coma {alone	C |VI‖VI‖VI|	0.20 λ	0.20 λ	0.86 λ	0.28 λ	0.27 λ	∞	0.19 λ	∞	
{plus lateral displacement	k C	−0.67 C	−0.67 C	?	−0.67 C	−0.60 C	?	−0.60 C	?	
	C	0.60 λ	0.58 C	0.86 λ	0.85 λ	0.66 λ	∞	0.47 λ	∞	
Fifth-order plus third-order coma plus lateral displacement {r₀		1.20	1.20	1.17	1.20	1.11	?	1.11	?	
{C′		2.50 λ	2.40 λ	3.10 λ	3.50 λ	2.70 λ	∞	1.85 λ	∞	

APERTURE

h_0

h_m

LONGITUDINAL SPHERICAL
ABERRATION

FIGURE 2.3. *Aperture of correction of longitudinal spherical aberration.*

coherent bright line perpendicular to the comatic flare ($\varphi_0 = 0$) we write

$$\frac{4\pi^2}{\lambda^2}\left(\frac{k^2}{3}+\frac{2kc}{5}-\frac{c^2}{7}+\frac{2kc'}{7}+\frac{2cc'}{9}+\frac{c'^2}{11}\right)<0.2.$$

The minimum value of the polynomial is obtained when $k = (5/21)C'$ and $C = -(10/9)C'(c_0 = 10/9$, so that the optimum aperture of correction will be $h_0 = (\sqrt{10}/3)t_m$ and the tolerance will be $C' = 1.85\lambda$. The results are collected in table 2.1. In the case of coma we have listed the tolerance for the two possible directions of the comatic flare (when $\varphi_0 = 0$ the flare is perpendicular to the object). In coherent illumination the coma has no effect when the flare is parallel to the line and the tolerance becomes infinite. In the case of astigmatism the best focus is located midway between the focal lines if the object is a point, and the tolerances are given for that position ($D = 0$). When the linear object is parallel to one of the focal lines, the best focus is located on that line and then the astigmatism has no effect. The maximum displacement with respect to that line is given by the tolerance for errors in focusing.

Let us now take an example. Suppose that the amount of astigmatism in an instrument is such that $A = \lambda/8$ (the distortion of the wavefront being $\pm 0.125\lambda$ on both sides of the reference sphere; altogether (0.25λ). The losses of contrast will be, for a dark point in a bright field, in incoherent illumination $0.2(0.125/0.17)^2 = 0.11$ and in coherent illumination $0.2(0.25/0.25)^2 = 0.05$. The losses for an incoherent line would be (midway between the two focal lines) when the line is parallel to a focal line $0.2(0.125/0.2)^2 = 0.04$ and, when the line makes a 45° angle with both focal lines, $0.2(0.125/0.18)^2 = 0.10$. The rates of losses of contrast can vary on a large scale (here the ratio of the extreme cases is about 3).

The Mechanical Computation of Diffraction Patterns

When the aberrations are either very small or very large, it is possible to compute the distribution of energy in diffraction patterns by using mathematical expansions.[3] We have already shown that the

---

[3] See paper 1, page 1, in this Symposium.

FIGURE 2.4. *Mechanical integrating device for computing diffraction patterns.*

FIGURE 2.5. *Repartition of energy represented by isophote curves, in the presence of "corrected" coma, balanced 3d÷5th order, for four values of the aperture of correction $(h_o/h_m) = 1^{1/2}$, $1.1^{1/2}$, $1.2^{1/2}$, $1.3^{1/2}$.*

losses of contrast can be expressed as functions of various pupillar integrals that can also be easily expressed in the case of small aberrations. When the aberrations are neither very small nor very large (in the transition between diffraction and geometrical optics) these mathematical procedures may fail. It is then useful to perform the integration expressed in eq 1 for a point source by means of a special mechanical device (fig. 2.4).

The principle of such a machine has already been described elsewhere.[4] The real and imaginary parts of the complete integral are developed by two integrating wheels, whereas the distortions of the wavefront (of any order) are given by cams, amplifier levers, adding tape, etc. The machine has been used for solving the following problems:

a. Distribution of energy in the presence of third-order aberrations. The goal of that study was mainly the knowledge of the transition between diffraction patterns and geometrical caustics. It seems that the transition is very rapid in the case of astigmatism when the focus-

[4] A. Maréchal, Thesis (Paris, 1947); Rev. opt. **26**, 257-277 (1947), **27**, 73-92, 269-287 (1948); J. Opt. Soc. Am. **37**, 982 (1946).

FIGURE 2.5.—Continued.
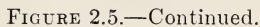
ing is done on a focal line. The diffraction pattern is already similar to the geometrical pattern even when the aberration is as small as $\lambda/4$. The transition is more gradual for coma, and still more gradual for spherical aberration.

b. Determination of the proper aperture of correction of geometrical aberrations when the aberrations are of the order of two or three times the Rayleigh limit.

c Special studies of practical cases of microscope objectives, Schmidt cameras, etc. The results can be represented by curves of equal illumination in the image plane.

Figure 2.5 shows the example of the study of comatic flares in the presence of corrected coma. The distortion of the wavefront is supposed to be conveniently represented by the superposition of third- and fifth-order terms; the fifth-order term is 5 λ (twice the tolerance for a bright-point source. See table 2.1); the third-order term is varied by steps as well as the aperture of correction $h_0 = v_0^{\frac{1}{2}} h_m$; the maximum illumination in the flare is 0.436, 0.46, 0.45, 0.40 for $v_0 = 1$, 1.1, 1.2, 1.3 so that the optimum aperture of correction is about $1.13^{\frac{1}{2}}$, whereas for a smaller aberration the maximum would be obtained for $1.2^{\frac{1}{2}}$. As a consequence the aperture of correction has to be decreased with increasing aberration. Extensive results will be published in the near future.

————

3. Diffraction Images Produced by Fully Corrected Objectives of High Numerical Aperture

By Harold Osterberg and Robert A. McDonald [1]

Introduction

The theory of diffraction by optical systems has been extended by R. K. Luneberg [1] [2] to include the effect upon the diffraction image of the geometrical coordination of the rays in the object and image space. An explicit form of the primary diffraction integral has been derived from Luneberg's formulation by Osterberg and Wilkins [2] upon the supposition that the objective is free of spherical aberration and coma. The main purposes of the present paper are to integrate and to discuss the primary diffraction integral under the additional supposition that absorption and reflection losses within the objective are negligible. A universal law governing diffraction phenomena with fully corrected objectives can be obtained only by ignoring the variable loss of light in the axial bundle of rays. We shall see that the universal law thus obtained differs in several interesting and important respects from the accepted classical laws governing idealized diffraction images.

Amplitude Variation Over the Converging Wavefront

Wavefronts expand (fig. 3.1. and 3.2) from an unresolvably small area centered about the axial point O and are converged without spherical aberration toward the conjugate point O' of the image space. Let

$$\rho \equiv \sin \vartheta; \qquad \rho_m = \sin \vartheta_m;$$

$$\rho_o \equiv \sin \vartheta_o; \qquad \rho_{om} = \sin \vartheta_{om}. \tag{1}$$

The numerical apertures of the objective with respect to its object and image space are $n\rho_m$ and $n_o\rho_{om}$, respectively, where n_o and n are the refractive indices of the object and image space, respectively.

If Abbe's sine condition is obeyed,

$$nM\rho = n_o\rho_o, \tag{2}$$

where M is the magnification ratio. Accordingly,

$$nM\rho_m = n_o\rho_{om} \equiv N.\,A., \tag{3}$$

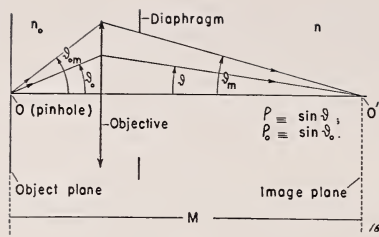where $N.\,A.$ denotes the numerical aperture of the system.

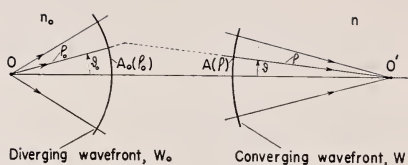FIGURE 3.1. *Convention with respect to the axial bundle of rays.*

FIGURE 3.2. *The amplitudes of A_o and A of the diverging and converging wavefronts, respectively.*

When the rays of the axial bundle obey (eq 2) and when these rays are transmitted equally from O to O' (fig. 3.2) it can be shown [3] that except for an unimportant factor of proportionality

$$A(\rho)/A_o(\rho_o) \equiv k(\rho) = (1-\rho^2)^{\frac{1}{4}}/(1-\rho_o^2)^{\frac{1}{4}}, \tag{4}$$

in which $A_o(\rho_o)$ and $A(\rho)$ represent amplitudes on the expanding wavefront W_o and on the converging wavefront W, respectively. For unpolarized light one may set

$$A_o(\rho_o) = 1 \tag{5}$$

and

$$A(\rho) = k(\rho) = (1-\rho^2)^{\frac{1}{4}}/(1-\rho_o^2)^{\frac{1}{4}}. \tag{6}$$

Microscope objectives have the property $|M| > 1$ and hence that $\rho_o^2 > \rho^2$. Accordingly, we observe from eq 6 that the amplitude $A(\rho)$ increases toward the outer portions of the converging wavefront. This increase becomes substantial when ρ_o can approach unity. The upper practical value of $\rho_{om} = N.A./n_o$ falls in the neighborhood of 0.95. The corresponding ratio $K(\rho_m)$ is approximately 1.8.

On the other hand, camera and telescope objectives have the property $|M| < 1$ and hence that $\rho_o^2 < \rho^2$. It follows from eq 6 that the amplitude $A(\rho)$ decreases toward the outer portions of the converging wavefront. This amplitude variation is ordinarily negligible with telescope objectives because they have low numerical aperture with respect to both object and image space.

Let $T(\rho)$ denote the amplitude transmission of the system for rays in the axial bundle. Reflection losses render $T(\rho)$ a decreasing function of ρ. The functions $T(\rho)$ and $k(\rho)$ counteract with microscope objectives but conjoin with camera objectives.

The Primary Diffraction Integral

The primary diffraction integral $U(r)$ for fully corrected objectives obeys [4] the relation

$$U(r) = 2\pi \int_0^{\rho_m} P(\rho) J_0(2\pi r n\rho)\rho\, d\rho; \tag{7}$$

$$P(\rho) \equiv \frac{A(\rho)T(\rho)}{(1-\rho^2)^{\frac{1}{2}}}. \tag{8}$$

24

Here, r is distance in number of wavelengths from the diffraction head (center of the diffraction image of an unresolved pinhole), $P(\rho)$ is the pupil function, J_0 is a Bessel function of the first kind and zero order, $A(\rho)$ is given by eq 6, and $T(\rho)$ is the amplitude transmission of the axial bundle of rays. For *lossless* objectives $T(\rho)=1$. The energy density in the diffraction image of a single, unresolvably small pinhole in an opaque slide is proportional to $U^2(r)$.

By combining eq 6, 7 and 8, we obtain explicitly

$$U(r) = 2\pi \int_0^{\rho_m} \frac{T(\rho)J_0(2\pi rn\rho)\rho d\rho}{(1-\rho^2)^{1/4}(1-\rho_o^2)^{1/4}} \tag{9}$$

in which ρ and ρ_o obey eq 2.

Airy's diffraction integral is obtained by setting $P(\rho)=1$ in eq 7. The familiar result is

$$U(r) = 2\pi \int_0^{\rho_m} J_0(2\pi rn\rho)\rho d\rho = 2\pi\rho_m^2 \frac{J_1(2\pi rn\rho_m)}{2\pi rn\rho_m}, \tag{10}$$

in which J_1 is a Bessel function of the first kind and first order.

Debye's diffraction integral is obtained by setting $A(\rho)T(\rho)=1$ in eq 8. It is

$$U(r) = 2\pi \int_0^{\rho_m} \frac{J_0(2\pi rn\rho)\rho d\rho}{(1-\rho^2)^{1/2}}. \tag{11}$$

The solutions of eq 10 and 11 are practically alike for small ρ_m. However [5], with respect to eq 11

$$\underset{\rho_m=1}{\text{Limit}}\ U(r) = 2\pi \frac{\sin(2\pi nr)}{2\pi nr}. \tag{12}$$

From a physical viewpoint, Debye's diffraction integral does not generate a single type of diffraction image.

Reversibility of the Primary Diffraction Integral for Fully Corrected Objectives

An objective and its primary diffraction integral are said to be reversible with respect to a pair of conjugate object and image planes when the diffraction images formed by interchanging the object point O and the conjugate image point O', figure 3.1, have similar distributions of energy density. Airy's diffraction integral of eq 10 is reversible. Debye's diffraction integral of eq 11 is not reversible.

The reversibility of the diffraction integral of eq 9 will now be demonstrated for the case $T(\rho)=1$, that is, for lossless objectives.

Let r_a denote Airy's limit of resolution measured as number of wavelengths in the image space. Then

$$2\pi r_a n\rho_m = 3.8317 \equiv \beta, \tag{13}$$

the first root of $J_1(2\pi r_a n\rho_m)=0$. Let x be distance measured in

Airy units r_a from the diffraction head to other points in the diffraction image.

$$x \equiv r/r_a. \tag{14}$$

Let also

$$w \equiv \rho/\rho_m. \tag{15}$$

By introducing the changes of variable of eq 13 to 15 together with ρ_o from eq 2 into eq 9, we obtain the primary diffraction integral for fully corrected, lossless objectives in the universal form

$$U(x) = 2\pi \rho_m^2 \int_0^1 \frac{J_0(\beta x w) w \, dw}{(1 - \rho_m^2 w^2)^{1/4} \left[1 - \left(\dfrac{n M \rho}{n_o} w \right)^2 \right]^{1/4}}. \tag{16}$$

$U^2(x)$ describes the distribution of energy density produced about point O', figure 3.1, by the light radiated from an unresolved pinhole centered on the object point O.

If the object point O and the conjugate image point O' are interchanged,

$$U'(x) = 2\pi \rho_m'^2 \int_0^1 \frac{J_0(\beta x' w) w \, dw}{(1 - \rho_m'^2 w^2)^{1/4} \left[1 - \left(\dfrac{n' M' \rho_m'}{n_o'} w \right)^2 \right]^{1/4}}, \tag{17}$$

in which the primed quantities refer to the new image space (formerly the object space). The equations of transformation from one space to the other are

$$n' = n_o; \quad n_o' = n; \quad M' = 1/M; \tag{18}$$

$$\rho_m' = \rho_{om} = nM\rho_m/n_o. \tag{19}$$

Therefore,

$$n'M'/n_o' = n_o/nM; \tag{20}$$

$$r_a' = 0.6098/n'\rho_m' = 0.6098/n_o\rho_{om} = r_a/|M|. \tag{21}$$

By introducing eq 18 to 20 into eq 17, one obtains in a straightforward manner the result

$$U'(x') = 2\pi \rho_m'^2 \int_0^1 \frac{J_o(\beta x' w) w \, dw}{\left[1 - \left(\dfrac{n M \rho_m}{n_o} w \right)^2 \right]^{1/4} (1 - \rho_m^2 w^2)^{1/4}}, \tag{22}$$

where x' is distance measured in Airy units r_a' in the new image space.

Comparison of the integrals of eq 16 and 22 shows that they are similar with respect to the role of x and x'. The primary diffraction integral of eq 16 is therefore reversible.

It follows from the reversibility of the primary diffraction integral that the diffraction images produced by a fully corrected objective operated as a microscope or as a camera objective between a fixed pair of conjugate planes are similar. This similarity holds in spite of the fact that the amplitude variation on the convergent wavefront is an increasing function of ρ, figure 3.2, with microscope objectives but a decreasing function of ρ with camera objectives.

Repetition of the above argument relative to eq 16 and 22 shows that the more general primary diffraction integral of eq 9 is reversible

provided that the amplitude transmission $T(\rho)$ for the axial bundle is the same whether the light passes from O to O' or from O' to O in figure 3.1. Functions $T(\rho)$ having this property will be called reversible. If, therefore, the reflection and absorption losses within a fully corrected objective lead to reversible amplitude transmission functions $T(\rho)$, the corresponding primary diffraction integral is reversible.

An Approximation with Respect to Microscope Objectives

In integrating eq 9 for microscope objectives, one may set

$$(1-\rho^2)^{-\frac{1}{4}}=1 \tag{23}$$

because $0\leqq\rho\leqq\rho_m$ with $\rho_m<0.04$. The primary diffraction integral for fully corrected microscope objectives may therefore be integrated in the alternate approximate forms

$$U(r)=2\pi\int_0^{\rho_m}\frac{J_0(2\pi r\rho)\,\rho d\rho}{(1-\rho_o^2)^{\frac{1}{4}}}, \tag{24}$$

$$U(x)=2\pi\rho_m^2\int_0^1\frac{J_0(\beta xw)\,wdw}{(1-N^2w^2)^{\frac{1}{4}}}, \tag{25}$$

in which $n=1$ and

$$\rho_o=nM\rho/n_o=M\rho/n_o; \tag{26}$$

$$N\equiv\frac{|M|\rho_m}{n_o}=\frac{N.A.}{n_o}. \tag{27}$$

$$r_a\equiv0.6098/N.A. \tag{28}$$

in number of wavelengths and

$$x\equiv r/r_a. \tag{29}$$

Phenomena at the Diffraction Head

The diffraction head corresponds to the center of the diffraction image of a single particle, that is, to the point $r=x=0$ in eq 24 and 25. From eq 25

$$\frac{U(0)}{\pi\rho_m^2}=2\int_0^1\frac{wdw}{(1-N^2w^2)^{\frac{1}{4}}};$$

$$=\frac{4}{3N^2}[1-(1-N^2)^{\frac{3}{4}}]. \tag{30}$$

N is defined by eq 27.

The energy density $G(0)$ at the diffraction head is given by $U^2(0)$ so that

$$\frac{G(0)}{\pi^2\rho_m^2}=\frac{16}{9}\frac{[1-(1-N^2)^{\frac{3}{4}}]^2}{N^4} \tag{31}$$

$G(0)/\pi^2\rho_m^4$ is called the normalized energy density. From eq 31

$$G(0)/\pi^2\rho_m^4 \rightarrow 16/9 \qquad \text{as} \qquad N \rightarrow 1; \tag{31, a}$$

$$G(0)/\pi^2\rho_m^4 \rightarrow 1 \qquad \text{as} \qquad N \rightarrow 0; \tag{31, b}$$

$$\therefore 1 \leqq G(0)\,\pi^2\rho_m^4 \leqq 16/9. \tag{31, c}$$

The usual classical value of $G(0)/\pi^2\rho_m^4$ is obtained from Airy's diffraction integral of eq 10 and is unity. We observe that the normalized energy density at the diffraction head can exceed the classical value appreciably.

It has been concluded by R. K. Luneberg [6] that among all primary diffraction images obeying eq 7 and producing *equal* total energy flux through the plane of the image, the primary diffraction image having the greatest normalized energy density at the diffraction head is the classical case of eq 10, that is, the case in which $P(\rho)$ is equal to a constant. This conclusion remains useful in considering the properties of diffraction integrals. However, one should not construe Luneberg's theorem to mean that the highest possible normalized energy density at the diffraction head is under *all* circumstances the classical value unity.

Primary Diffraction Images with Microscope Objectives

With respect to eq 25, let

$$F(w^2) \equiv (1 - N^2 w^2)^{-\frac{1}{4}} \tag{32}$$

be approximated by the power series

$$f(w^2) \equiv 1 + a_1 w^2 + a_2 w^4 + \ \ldots \ a_m w^{2m}, \tag{33}$$

in which the finite number m of coefficients a_j is chosen high enough so that $f(w^2)$ is a satisfactory approximation to the monotonic function $F(w^2)$ in the interval $0 \leqq w \leqq 1$.

In the following calculations the coefficients a_j have been determined from $f(w^2)$ and $F(w^2)$ by the method [7] of least squares.

From eq 25, 32, and 33,

$$\frac{U(x)}{2\pi\rho_m^2} = \sum_{j=0}^{m} a_j \int_0^1 J_0(\beta x w) w^{2j+1} dw. \tag{34}$$

It is convenient to introduce the change of variable $z \equiv \beta x w$ and thus obtain

$$\frac{U(x)}{2\pi\rho_m^2} = \frac{1}{\beta^2 x^2} \sum_{j=0}^{m} \frac{a_j}{(\beta x)^{2j}} \int_0^{\beta x} J_0(z) z^{2j+1} dz. \tag{35}$$

Then, because

$$\int_0^{\beta x} J_0(z) z^{2j+1} dz = \sum_{\nu=0}^{j} \frac{(\beta x)^{2\nu}}{(-4)^{\nu-j}} \frac{(j!)^2}{(\nu!)^2} [\beta x J_1(\beta x) + 2\nu J_0(\beta x)], \tag{36}$$

$$\frac{U(x)}{2\pi\rho_m^2} = \frac{J_1(\beta x)}{\beta x} + \sum_{j=1}^{m} \frac{a_j}{(\beta x)^{2j}} \sum_{\nu=0}^{j} \frac{(\beta x)^{2\nu}}{(-4)^{\nu-j}} \frac{(j!)^2}{(\nu!)^2} \left[\frac{J_1(\beta x)}{\beta x} + 2\nu \frac{J_0(\beta x)}{(\beta x)^2} \right]. \quad (37)$$

To calculate $U(x)$ for small x-values, we have preferred to expand $J_0(z)$ in eq 35 into the standard Bessel series and to integrate term by term. The resulting expression for $U(x)$ is

$$\frac{U(x)}{2\pi\rho_m^2} = \frac{1}{2} \sum_{j=0}^{m} a_j \sum_{\nu=0}^{\infty} \left(\frac{-1}{4} \right)^{\nu} \frac{(\beta x)^{2\nu}}{(\nu!)^2(j+\nu+1)}. \quad (38)$$

It is pointed out that the method of eq 32 to 38 can be applied to the exact integral of eq 16 by choosing

$$F(w^2) = (1 - N^2 w^2)^{-\frac{1}{4}} (1 - \rho_m^2 w^2)^{-\frac{1}{4}}. \quad (39)$$

The use of eq 39 instead of eq 32 is not justified unless one employs a higher number of coefficients than employed by the writers.

The normalized energy densities $U^2(x)/\pi^2\rho_m^4$ have been calculated from eq 37 and 38 for the cases $N \equiv N.A./n_0$ equal to 0.85, 0.92 and 0.95 and are plotted against $x = r/r_a$ as the solid curves in figures 3.4 to 3.6. The case $N = 0.95$ corresponds to "dry" microscope objectives having the numerical aperture 0.95 or to oil-immersion objectives having the numerical aperture 1.44.

The number and values of the coefficients a_j are listed in table 3.1. The agreement between $F(w^2)$ and $f(w^2)$ is illustrated in figure 3.3 for the case $N = 0.85$. The deviation of f from F is less than 3, 13 and 19 parts per thousand for $N = 0.85$, 0.92 and 0.95, respectively. The addition of one coefficient for the case $N = 0.95$ is indicated but would not, we estimate, alter appreciably the diffraction curve of figure 3.6.

TABLE 3.1 *Number and Values of the Coefficients* a_j

N	a_1	a_2	a_3	a_4
0.85	0.208353	−0.065553	0.231649	--------
.92	.206122	.203478	−.316991	0.491368
.95	.322624	−.328069	.362459	.400207

The normalized energy densities $U^2(x)/\pi^2\rho_m^4$ belonging to Airy's diffraction integral of eq 10 are included for comparison as the broken curve of figures 3.4 to 3.6. Airy's diffraction curve is unity at the diffraction head where $x = 0$ and reaches its first zero at Airy's limit of resolution, where $x = 1$.

The most conspicuous property of the solid curves drawn in figures 3.4 to 3.6 for fully corrected, lossless microscope objectives is the rapid increase of energy density at the diffraction head ($x = 0$) with increasing N. This property has been explained on page 28.

A second interesting property of the solid curves of figures 3.4 to 3.6 is that the first zero of the energy density occurs at x-values less than unity, that is, within Airy's limit of resolution. The data from which figures 3.4 to 3.6 have been plotted show that the first zeros occur near $x = 0.96$, 0.95, and 0.94 for cases $N = 0.85$, 0.92, and 0.95, respectively. Let x_0 denote the first root of $U^2(x) = 0$. Then with respect to the primary diffraction curves for fully corrected objectives $x_0 \leqq 1$. Furthermore, x_0 is a decreasing function of N.
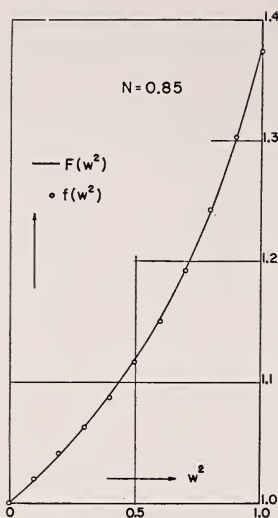
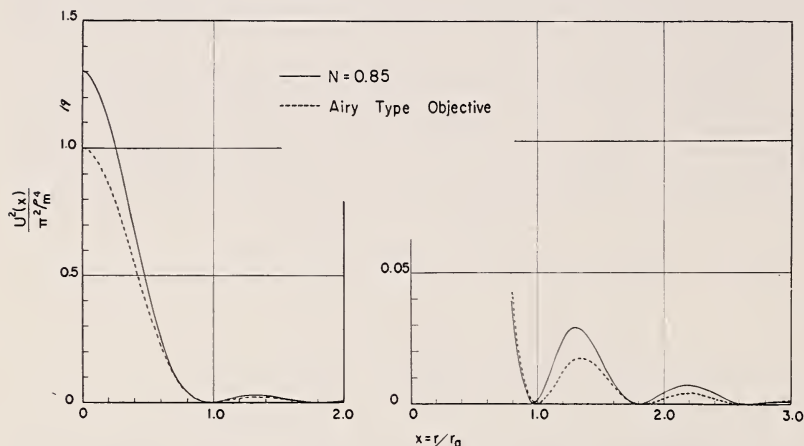FIGURE 3.3   *The fit of $f(w^2)$ to $F(w^2)$ for the case $N=0.85$.*

FIGURE 3.4.   *The primary diffraction curve for the case $N=0.85$.*

$U^2(x)$ is proportional to the energy density in the diffraction image.   Distance x from the center of the diffraction image is measured in Airy units r_a.

It is especially noteworthy that x_0 is a decreasing function of N while the energy density at the diffraction head is an increasing function of N.   The energy density at the diffraction head is therefore a *decreasing* function of $x_0(N)$.   The reduction of x_0 below $x_0=1$ by artificial means such as coating [8] the exit pupil of the objective is invariably accompanied by marked reduction in the energy density at the diffraction head.   In other words, when one attempts to decrease x_0 artificially, he is faced by the damaging fact that the energy density at the diffraction head becomes an *increasing* function of x_0.

Figures 3.4 to 3.6 show that the ratio of the energy densities at the subsidiary maxima to the energy density at the central maximum is higher with the fully corrected objective than with Airy type objec-
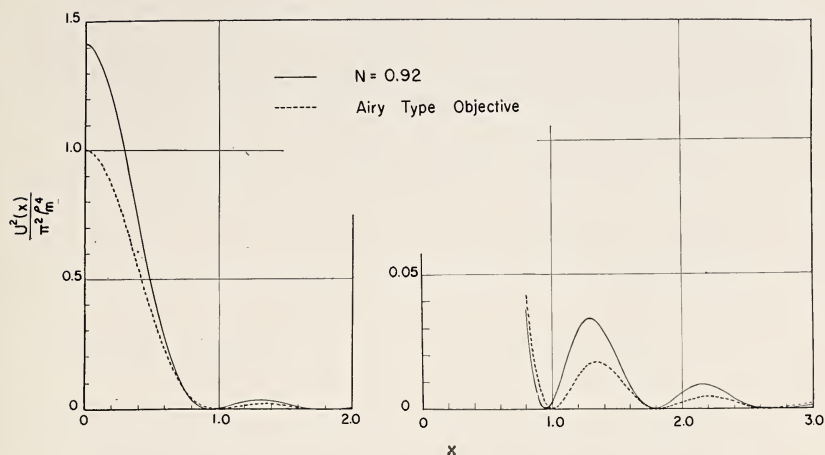
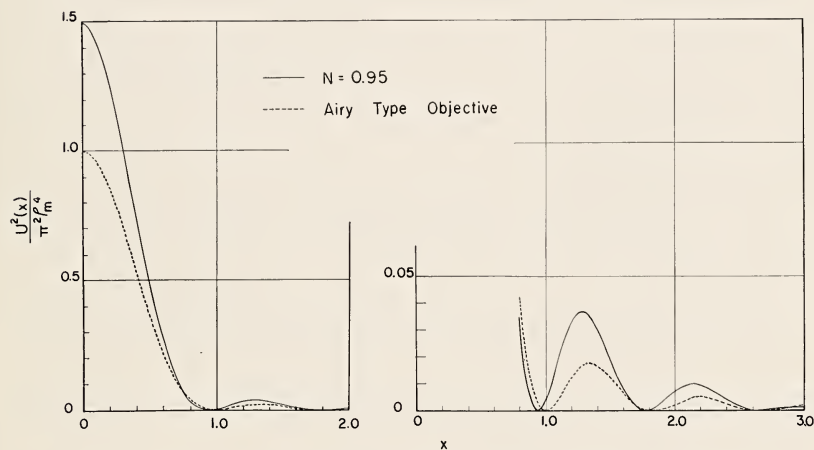FIGURE 3.5. *The primary diffraction curve for the case N = 0.92.*

FIGURE 3.6. *The primary diffraction curve for the case N = 0.95.*

tives (idealized objectives that obey Airy's diffraction integral). The relative increase in the brightness of the diffraction rings is however small.

Resolution of Two Like Pinholes in an Opaque Slide

We have seen from figures 3.4 to 3.6 that as N is increased the energy density at the diffraction head increases and the location x_0 of the first zero in the energy density decreases, that the normalized energy density at the diffraction head is greater than unity and that x_0 is less than unity. These observations lead to the expectation that fully corrected objectives will have greater resolving power than the idealized Airy-type objective. This expectation will now be justified.

Let two like pinholes having unresolvably small area A be located equidistantly from the optical axis and be illuminated as in a microscope by a substage condenser whose numerical aperture with respect to the object space of the objective is variable and is denoted by

$$N.A._{\text{condenser}} = n_o \rho_{cm}.\tag{40}$$

Let

$$s \equiv N.A._{\text{condenser}}/N.A._{\text{objective}};\tag{41}$$

$$\beta \equiv 3.8317;\tag{42}$$

$x \equiv$ distance measured in Airy units from the optical axis in the image space of the objective;

$2L \equiv$ separation in Airy units of the centers of the geometrical images of the two pinholes in the image space of the objective;

$$Z_1 \equiv \beta(x-L);\tag{43}$$

$$Z_2 \equiv \beta(x+L);\tag{44}$$

$$\psi \equiv \frac{2J_1(2\beta sL)}{2\beta sL};\tag{45}$$

$$u(Z) \equiv U(Z)/2\pi\rho_m^2,\tag{46}$$
where $U(Z)$ is to be calculated from eq 37 and 38 for fully corrected objectives;

and

$G(x) =$ the distribution of energy density in the sharply focused diffraction image of the two pinholes.

Then it can be shown by repeating for opaque backgrounds an argument by Osterberg and Wissler [9] that

$$G(x)/4\pi^3 A^2 \rho_m^4 \rho_{cm}^2 = u^2(Z_1) + u^2(Z_2) + 2\psi u(Z_1)u(Z_2).\tag{47}$$

Equation 47 is identical to an equation derived independently (and almost simultaneously) by Hopkins and Barham [10]. Equation 47 includes the effect of the numerical aperture of the condenser upon the diffraction image of the two pinholes.

The curves of figures 3.7 to 3.9 have been calculated from eq 47 for the case $s=1$, that is, for the case in which the numerical aperture of the substage condenser is equal to the numerical aperture of the objective. The curves $G(x)$ are symmetrical about the point $x=0$ and therefore have been plotted only for $x>0$.

The family of diffraction curves of figure 3.7 belong to fully corrected objectives for which $N=0.85$. A pronounced concentration of energy density occurs at the points $x=\pm 0.5$ Airy unit when the separation $2L$ is 1.0 Airy unit. The two pinholes are easily resolved when $x=\pm 0.5$, are undoubtedly resolved when $L=0.45$ but are probably
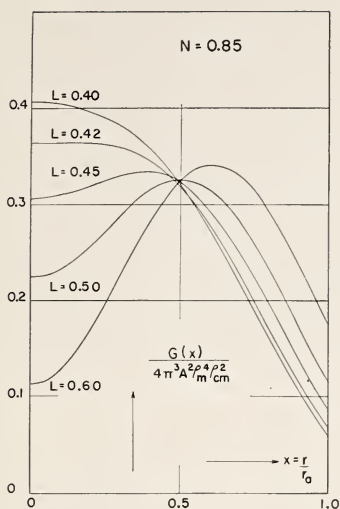
FIGURE 3.7. *Diffraction curves for two pinholes for the case N = 0.85.*

The pinholes are located equidistantly from the optical axis and x is measured in Airy units from the optical axis.
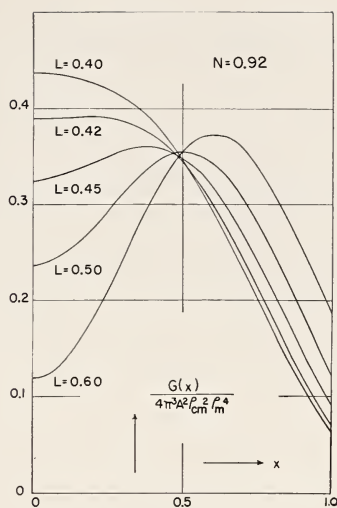

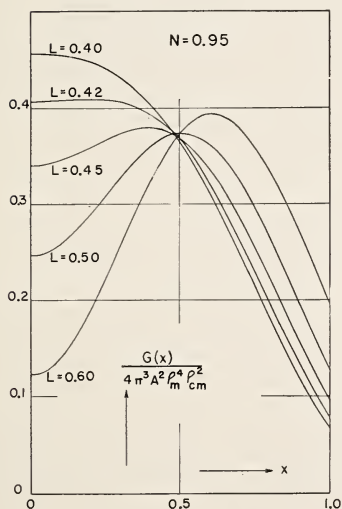FIGURE 3.8. *Diffraction curves for two pinholes for the case N = 0.92.*

FIGURE 3.9. *Diffraction curves for two pinholes for the case N = 0.95.*
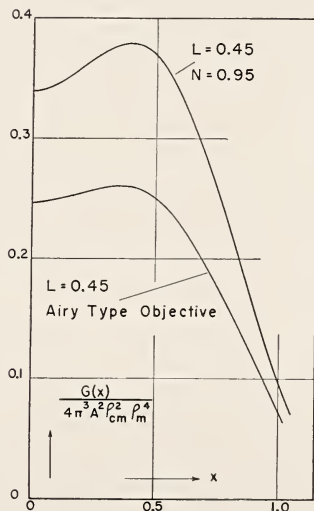

FIGURE 3.10. *Comparison of the diffraction curves for two pinholes as produced by an objective of Airy type and by a fully corrected objective for which N = 0.95.*

The pinholes are separated by $2L = 0.90$ Airy unit in the object space.

not resolvable when $L=0.42$. The *physical limit* [11] of resolution is the separation $2L$ for which

$$\frac{\partial^2 G(x)}{\partial x^2}=0 \quad \text{at} \quad x=0. \tag{48}$$

Evidently, the physical limit of resolution, $2L$, falls slightly below 0.84 when $N=0.85$.

The family of curves of figure 3.8 belong to the case $N=0.92$. A more definite dip in the energy density now occurs at $x=0$ for the half-separation $L=0.42$. The diffraction images described by this family of curves show definitely better contrast and resolution than do the diffraction images for the case $N=0.85$. This trend toward better contrast and resolution with increasing N is continued in figure 3.9 for the case $N=0.95$.

The physical limit $2L$ is estimated to be 0.82 when $N=0.95$. The physical limit $2L$ is [12] 0.843 with Airy-type objectives in viewing two, like, true pinholes in an opaque slide under the condition $s=1$. The physical limit of resolution is therefore lower (more favorable) with fully corrected objectives than with the idealized Airy-type objective.

The diffraction curves of figure 3.10 have been included for comparison of the fully corrected objective with the Airy-type objective. To compute $G(x)$ for objectives of Airy type, one determines $U(Z)$ in eq 46 from Airy's diffraction integral of eq 10. Whereas the half-separation $L=0.45$ is resolved by both objectives, resolution is more certain and contrast is better with the fully corrected objective.

Another interesting point of comparison is afforded by figure 3.10. According to the laws of geometrical optics, the concentration of energy density should occur at $x=\pm L$ Airy units when the separation is $2L$ Airy units. Figures 3.4 to 3.6 display the typical tendency that the concentrations of energy density do occur near $x=\pm L$ when $2L\geqq 1.0$, but that the concentrations depart from the locations $x=\pm L$ when $2L<1.0$. In figure 3.10 the concentration of energy density occurs near $x=0.4$ and $x=0.35$ for the fully corrected objective and the Airy-type objective, respectively.

The concentration of energy density lies closer to the geometrical location $x=0.45$ with the fully corrected objective than with the objective of Airy type.

In summary, a fully corrected objective that has negligible internal losses of light due to reflection and absorption is superior to the idealized objective of Airy type in practically all respects from the viewpoint of physical optics. The amplitude variation introduced over the converging wavefront by bending the rays from the object space into the image space in accordance with Abbe's sine condition is beneficial and should not be counteracted by the deposition of an absorbing coating over the exit pupil of the objective to render the objective of Airy type.

Conclusions

The following advantages over the idealized objectives of Airy type are possessed by fully corrected objectives that have negligible internal losses of light.

(1) The energy density at the diffraction head is higher. This energy density increases with $N \equiv N.A./n_o$ and approaches 16/9 times the classical value as N approaches unity. (2) The first zero of the energy density occurs nearer to the diffraction head. (3) The physical limit of resolution is lower (more favorable) in viewing two pinholes in an opaque slide. (4) Actual resolving power is better in viewing the diffraction images of two illuminated pinholes in an opaque slide. (5) Contrast in the diffraction image of two illuminated pinholes is higher.

The single advantage retained by the classical objective of Airy type is that the brightness of the diffraction rings is lower with respect to the brightness of the central maximum.

The primary diffraction integral for fully corrected objectives is reversible. Not all primary diffraction integrals have this property.

References

[1] R. K. Luneberg, Mathematical theory of optics, p. 344–386 (Brown University, Providence, R. I., 1944).
[2] H. Osterberg and J. E. Wilkins, Jr., J. Opt. Soc. Am. **39**, 553–555 (1949).
[3] H. Osterberg and J. E. Wilkins, Jr., J. Opt. Soc. Am. **39**, 554, eq 6 and 8 (1949).
[4] H. Osterberg and J. E. Wilkins, Jr., J. Opt. Soc. Am. **39**, 554, eq 3, 8, and 9 (1949).
[5] A. Gray, G. B. Mathews, and T. M. Macrobert, Bessel Functions, p. 211, eq 66 (Macmillan & Co., Ltd., London, 1931).
[6] R. K. Luneberg, Mathematical theory of optics, pp. 390–391, eq 50.45 (Brown University, Providence, R. I., 1944).
[7] I. S. Sokolnikoff and E. S. Sokolnikoff, Higher mathematics for engineers and physicists, pp. 536–542 (McGraw-Hill Book Co., New York, N. Y., 1941).
[8] H. Osterberg and J. E. Wilkins, Jr., J. Opt. Soc. Am. **39**, 556, figure 3 (1949).
[9] H. Osterberg and F. C. Wissler, J. Opt. Soc. Am. **39**, 558–561 (1949).
[10] H. H. Hopkins and P. M. Barham, Proc. Phys. Soc. [B] **63**, 737–744, eq 22 (1950).
[11] H. Osterberg, J. Opt. Soc. Am. **40**, 299 (1950).
[12] H. Osterberg, J. Opt. Soc. Am. **40**, 300, figure 5 (1950).

Discussion

Dr. F. Mooney, Bausch & Lomb Optical Co., Rochester, N. Y.: I would like to hear Dr. Osterberg redefine a fully corrected objective.

Dr. Osterberg: Insofar as my purposes are concerned, a fully corrected objective is without spherical aberration and is fully corrected for any design condition that is without color. With monochromatic light, you would not have to correct for color.

Dr. G. Toraldo, Instituto Nazionale di Ottica, Arcetri, Florence, Italy: I would like to remark on Dr. Osterberg's paper. I think that to treat the problem from a purely academic point of view is not sufficient when such large angles are considered. I have already discussed with Dr. Osterberg the question of polarization, but there is still another question. What is the meaning of the curves of the diffraction disks when one has such a large angle? One must specify what is intended. Also, it appears that the conservation-of-energy law is violated because either the other patterns or these are true. The other patterns contain more energy. Which is the correct one?

Dr. Osterberg: You have asked, I believe, two questions. One question you have asked really amounts to this: If we had an illuminated pin hole can we say that the wavefront that expands from it can be regarded as spherical?

The first remark to be made is this: In the classical viewpoint, of course, when considering Airy-type objectives we asked no question in this regard so I may stand my ground and say I take the same privilege. However, there is this to be said. I agree with Toraldo that the only way you can really solve these problems is to start with the atomic radiator in the source, trace your polarization vectors all the way through, and calculate the energy density at the end.

I have done this in connection with the building up of a theory of phase microscopy. I have started with the dipole radiators in the source and have traced the development through to the final energy densities and under certain approximations that are made we get the final result.

Now, the approximations in a theory like this are sometimes very difficult to analyze. Just what do they mean? If one supposes that the microscope illuminator has sufficient numerical aperture to focus the complete diffraction energy in the source then your conclusion would follow. That takes care of one question. What was the second?

DR. TORALDO: The energy consideration, the diffraction disk.

DR. OSTERBERG: Yes. Here is how I view that. Let us take a small cone of rays containing a certain amount of energy. On passing through a microscope objective, this energy flux converges as another, and longer, cone. As the angular aperture increases, the emergent flux is squeezed into a narrower and narrower cone relative to the incident cone, so the amplitude is greater on the outer portions of the wavefront.

Now, when you develop a theory for the Airy-type objective what do you say? You say that the energy on that wavefront is uniform. What have you done? You have thrown away energy. I have only included that energy and so I get the taller diffraction peaks.

CHAIRMAN: Dr. Herzberger.

DR. M. HERZBERGER, Eastman Kodak Research Laboratory, Rochester, N. Y.: I wish to make one remark with respect to the beautiful paper of Prof. Zernike. I enjoyed very much the mathematical treatment as well as the agreement shown between theory and practice. I wish to say that geometrical optics is not as bad as it has been made to seem. The circle of confusion of which Prof. Zernike spoke is based on an inaccurate approximation to geometrical optics. From more exact computation one gets a picture that is similar to your diagrams and that shows the diamond shape of the image though, of course, it does not show the interference pattern.

PROF. ZERNIKE: I want to make a remark on Dr. Osterberg's paper. I looked into the history of the subject myself about 5 years ago. I did not publish anything about it. A year and a half ago I visited H. H. Hopkins in England who has written a paper on this subject. Then we learned of a paper by "Ignatovsky" a Russian physicist who worked in Germany before 1917, but this paper was written in 1918 in Russian. He probably solved everything about it, including the directed vibrations of light and my impression is that because of that the influence is much less than Dr. Osterberg has shown. The Airy disk remains practically unchanged.

DR. OSTERBERG: With respect to Dr. Hopkins, I am under the impression that any of his papers that I have seen had to do with assumed vibrations on the converging wavefront. Generally speaking,

the amplitude is decreasing. I am not aware that Hopkins has studied the particular variation of this paper. Here we have included the variation of the amplitude produced by bending of the rays in accordance with the Abbe sine condition. Is it your impression that Hopkins has discussed this consideration?

Prof. Zernike: I could not say off-hand.

Prof. Maréchal. Hopkins has discussed the reverse. He considered an incident plane wavefront and studied the formation of the image assuming that the angular aperture was very great. It is exactly the reverse of the case studied by Dr. Osterberg.

In the line of Hopkins' work we have performed the computation for the diffraction image given by an aplanatic system of high numerical aperture (in the image space). The image is no longer circular, but elliptical, as Hopkins has shown. The direction of the longer axis of the ellipse representing the electric energy is also the direction of the shorter axis for the magnetic energy; now if we consider the Poynting vector we find a circular distribution. Experiments have been tried in order to show the ellipticity of the distribution of electric energy, but no conclusion has been obtained as yet.

Dr. Osterberg. Are you certain that he included the bending of the rays in his (Ignatovsky's) work according to the Abbe sine condition? If he did not he would not get this result.

Prof. Zernike. I should have added that he gave me a reprint of the paper to be translated but I have not referred to it for a year and a half.

Dr. Osterberg. Its contents should be interesting.

Dr. M. Golay, Signal Corps Engineering Laboratories, Fort Monmouth, N. J. We have seen many papers in which we are concerned with what happens to an image. We start with a point source, or if you like, from a perfect Airy disk, which I believe will give us the same results, and we arrive at an imperfect image. Of course, we are not interested in one lens. We are interested in many lenses. We are interested in what happens to the imperfect image and how much worse it is in the next step, and one question would be, what is there about these calculations that is additive and would permit one to take these steps individually and know what happens at the end of the system. I have a suspicion that some of the B^2 terms of Prof. Zernike have this additive property. If you look at it from the standpoint of thermodynamics, something is always lost and is never retrievable. Recently in communication engineering there have been beautiful applications of this notion of things that are lost and not recovered. I wonder if there is not an application in this field of that notion of entropy that has a good mathematical and useful significance, and can be applied to what happens in images and give us a good strong concept to discuss not just one lens but to discuss a system of lenses. This is merely a suggestion.

Dr. Herzberger. I would like to make a remark about what has been said. I think there is a law that governs the action of an optical system. The optical image formation can always be described as the result of a wavefront calculated from geometrical optics, diffracted at the exit pupil. There is ample experimental evidence to verify this statement, which is the basis of all calculations of the diffraction image.

Besides, there is some kind of an invariant, discovered by R. Straubel, that governs the connection between contrast and resolution in

optical systems. One can only improve one of these two factors at the expense of the other. Straubel especially investigated the influence of absorption on resolution and contrast, an investigation in which the theoretical conclusions were paralleled by experimental studies.

DR. R. C. SPENCER, Air Force Cambridge Research Laboratories, Cambridge, Mass. I am interested primarily in microwave optics and microwave optical systems. Although we have attempted to obtain some help from existing optical theory, in general we have been forced to attack directly certain problems that differ materially from the usual optical problems. These differences I would like to point out.

(a) The intensity illumination over the aperture of a reflector-type antenna or metal lens is not constant but reduced by as much as a factor of 10 at the edges. (b) The f/d ratios range from unity down to one-third. (c) The field, or the angle over which a horn will pick up the image, ranges in certain cases from 30 through 90° and at times 360°. (d) In some cases the feed is off-axis so that the optical system is not symmetrical. (e) We are working much closer to the limitation of the Rayleigh disk. For instance, a drop of 20 percent (one decibel) in the intensity of the center of the Rayleigh or Airy disk will raise the intensity of the side lobes (diffraction rings) to such an extent that the pattern becomes unusable.

Operational methods [1] have proved a powerful tool in evaluating the effect of aberrations on diffraction patterns. For a one-dimensional aperture with amplitude illumination $f(x)$, the diffraction pattern is the Fourier transform $g(u)$ of $f(x)$. It can be shown that if $f(x)$ is perturbed by multiplication by a function $H(x)$ then the diffraction pattern is $H(-iD)g(u)$ where $D=d/du$. Thus $H(-iD)$ is an operator that transforms or perturbs the normal diffraction pattern $g(u)$. This device gives a short-cut method for evaluating certain patterns. Thus the pattern for an amplitude illumination $(1-x^2)$ within the range ± 1 and zero without, is $(1+D^2)g_0(u)$ where $g_0(u)$ is the amplitude diffraction pattern for the uniformly illuminated aperture. For an aperture illumination in a series of powers of x over the range ± 1 the diffraction pattern is a series of derivatives of $(\sin \phi)/\phi$.[2]

When applied to the analysis of phase errors over the aperture caused by various aberrations, the amplitude perturbation function is $H(x)= e^{i\phi(x)}$ and the perturbed amplitude diffraction pattern is simply the normal pattern $g(u)$ multiplied by the operator $e^{i\phi(-D)}$, the rule being to replace x by $-iD$. This has been used successfully in evaluating the effects of defocusing and should converge for other aberrations. In general, the operator is expanded in powers of D. For example, in the case of Fresnel diffraction for a uniformly illuminated aperture the aperture amplitude $e^{i\beta x^2}f(x)$ is transformed over to the amplitude diffraction pattern $e^{\beta D^2}g(\phi)=(1+\beta D^2+\beta D^2+\beta^2 D^4/2!+ \ . \ .)g(\phi)$ where where β is the phase error at either end of the aperture and $g(\phi)=$

[1] MIT Rad. Lab. Series, **12**, 186 (McGraw-Hill Book Co. 1949): (a) Roy C. Spencer, Paraboloid diffraction patterns from the standpoint of physical optics, MIT Radiation Laboratory Report T-7 (Oct. 21, 1942); (b) Roy C. Spencer, Synthesis of microwave diffraction patterns with application to csc²θ patterns, MIT Radiation Laboratory Report 54–24 (June 23, 1943); (c) Roy C. Spencer, et al., Tables of Fourier transforms of Fourier series, power series, and polynomials, MIT Radiation Laboratory Report S-58 (Aug. 30, 1945); (d) Roy C. Spencer, Fourier integral methods of pattern analysis, MIT Radiation Laboratory Report 762-1 (Dec. 5, 1945); (e) Roy C. Spencer and Pauline M. Austin, Tables and methods of calculation for line sources, MIT Radiation Laboratory Report 762-2 (March 30, 1946).

[2] Harvard Computation Laboratory Series, vol. 22, Tables of the function $(\sin \phi)/\phi$ and of its first eleven derivatives, Introduction—Section II "Application" by Roy C. Spencer.

$(\sin \phi)/\phi$ is the amplitude diffraction pattern for a uniformly illuminated aperture.

A special aspect of this problem is the fractional loss in gain (or brightness at the center of the Rayleigh disk) due to phase errors. If the phase front is approximated by a plane so that the square of the error between the phase front and the plane is minimized by the method of least squares, using as a weight function the amplitude illumination $F(x, y)$, then the fractional loss in gain[3] is given by

$$\frac{\Delta G}{G} = \frac{\int \phi^2 F(x, y)\,dxdy}{\int F(x,y)\,dxdy} = \overline{\phi^2}.$$

Thus, neglecting powers of ϕ above the second the fractional loss in gain is the *weighted-mean squared phase error*, with ϕ in radians. Needless to say, if the loss in gain is held down to a few percent, this sets an upper limit on both the increase in the width at half power and the energy increase in the side lobes.

Dr. J. G. Baker, Research Associate of Harvard College Observatory, and Optical Consultant to the Perkin-Elmer Corp., Norwalk, Conn.: I should like to ask Dr. Zernike if in his treatment he can include the primary astigmatism in terms of the higher order angular departures from the optical axis. When you go far off-axis, large aberrations have to be considered, and I am wondering whether your expressions will not have to be modified. You have normalized to unity over the aperture in each case and I am wondering if corrective factors are needed when you get to large oblique aberrations? The image may show hybrid symmetrical and unsymmetrical aberrations.

The question is of interest to me because in some systems I have designed, the inclinations run to 30 degrees off-axis and yet the requirements on image quality are comparable to those on-axis. Vignetting is another thing to consider.

Dr. Zernike: We have applied the theory only to circular pupils.

Dr. Maréchal: There is no essential difference in the off-axis theory at moderate angular apertures for a circular pupil.

Dr. Baker: I have one system of 52-degrees total field at f/0.65 where these considerations are of importance. The vignetting arises from film-holder shadows as well as from lens and mirror cells, and the pupil is far from circular.

Mr. A. J. Lipinski, University of California, Los Alamos Scientific Laboratory, Los Almos, N. Mex.: I wish to address a question to Dr. Zernike. How did you consider the geometrical as compared to the diffraction images? You mention four times better just by the diffraction theory and then you later mention this can be made by balancing the aberrations to 28 times what actually appears geometrically. Then you say in a practical case you reduce it to 5.6 times. I am a little uncertain as to what you mean.

Prof. Zernike: I had to abridge my talk somewhat. I had really in my paper a somewhat different thing. The 5.6 or 28 refer to the difference between balancing the error or by fourth and sixth orders or by not doing that. This polynomial (illustrating) would mean

---

[3] Roy C. Spencer, A least square analysis of the effect of phase errors on antenna gain, AFCRC Report E5025 (January 1949).

that you want the ray to go up and down so that you have this circle of confusion. You have the answer to this in the paper. I find that by doing it geometrically, making a transverse aberration of the rays you get an approximation that is not too far from the diffraction theory treatment, which means that judged by the geometrical optics you are always rather far out. This disk that would be calculated from the rays is much larger than the relative function that you get.

If you look at the diffraction properties, it would be only 1.75 times less than the correct one. I think that answers your question.

MR. G. H. CONANT, JR., Harvard College Observatory, Cambridge, Mass.: Dr. Zernike, would you restate your ordinate and abscissa on the diagram of the isophotes? Axial sections near the focus were shown.

PROF. ZERNIKE: It was just near the focus of the aberrations. You would have to rotate the entire diagrams in order to get a three-dimensional figure.

PROF. MARÉCHAL: The question of resemblance between the effective diffraction pattern and the geometrical approximation is rather intricate. It seems that it depends entirely upon the nature of the aberrations. In the case of astigmatism if you focus on one focal line the diffraction pattern is very similar to the geometrical focal line even if the aberration is very small (half a wavelength). If you increase the aberration you will rapidly obtain the aspect predicted by geometrical optics.

In the case of coma it seems that you have to go up to three or four wavelengths at least to have some resemblance between diffraction image and the geometrical one, and for spherical aberrations, it seems to be still more. In fact, it is quite difficult to tell exactly how we can define the resemblance.

In the case of coma you have still a high maximum when the aberration is of the order of magnitude of 2 or 3 wavelengths. Instead of rings you have horseshoes.

DR. H. R. J. GROSCH, General Electric Co., Lockland, Ohio: I would just like to comment upon noise. I suppose we could interpret band width, which, after all, is the limitation of the channel over which the information is coming as some function of N, but what are we going to do about the question whether the noise introduced by the imperfect lens system is what the communication boys call "white" or not? I think the trouble is that there are highly selective effects. I have a feeling that we are not anywhere near close enough to information theory from the communication engineering viewpoint yet.

MR. GOLAY: As I understand it, we have a loss due to the poor quality of the system. This poor quality may be something that cannot be avoided but it is not noise. We may have a system that is as good as we can make it. But can we have a measure of the poor quality that is additive so that we can predict what we will have at the end of the system?

DR. BAKER: This is a short question. Are these orthogonal functions necessary and sufficient to represent any possible aberrations of the wavefront to all orders?

PROF. ZERNIKE: Certainly.

4. Bases for Testing Photographic Objectives

By L. E. Howlett [1]

General Considerations

It is a matter of more than passing interest that in spite of the fact that optical instruments have for a very long time been widely used as tools in scientific research, there is a great deal of disagreement as to the way in which their image quality should be judged. An explanation for this condition, which is at least plausible, may be found in the historical development of the uses to which lenses have been put. Whatever the real validity of such an explanation it serves to bring out the erroneous trends of thought on optical performance that have been largely responsible for the present diversity of opinion.

There have been three distinct periods in the development of the usefulness of optical instruments to meet scientific needs.

Until the early years of the nineteenth century optical instruments found their major, if not exclusive, scientific use as an aid in making visual observations. Astronomical research combined with other applications of telescopes was probably the most important scientific use.

Shortly after the opening of the nineteenth century the development of photographic science was very rapid. For pictorial reasons a demand arose for lenses that would cover a reasonably large field of view. To attain this end a diminution of image quality below that of the astronomical requirement was considered quite acceptable because the intrinsic resolution of available emulsions was certainly inferior to that obtained visually on the axis of telescopic systems.

By the twentieth century lenses came to be combined not only with photographic emulsions but with many other kinds of energy receivers such as thermopiles, photocells, etc. They became an important element in television where the character of their optical imagery is of great importance in the performance of electrical circuits and the over-all quality of a system as a transmitter of information in the form of electrical signals.

It is a most unhappy circumstance that during the expansion of the use of lenses from the stage where they were an aid to visual observation to where they were used to make observations in combination with photographic emulsions, thermopiles, photocells, etc., designers ignored the necessity of reconsidering and adjusting for validity the design criteria, which had proved wholly acceptable for the design of good objectives for visual use. It seems they had become generally satisfied with an erroneous conception that if a lens performed well in combination with the eye it would perform equally well when combined with any other type of energy receiver. Those who would not expect a microscope objective designed for the infrared to excel in the

[1] Division of Physics, National Research Council, Ottawa, Canada.

ultraviolet nevertheless assumed that a good visual lens must be a good photographic lens. The thought that a lens performing in only an adequate way with the eye might be superior to a first-quality visual lens when combined with a photographic emulsion or other receptor did not receive appropriate consideration. It seems that only currently is the condition underlying such a concept being given any attention by designers. If all lenses over their reputed field of view were limited only by diffraction theory no such confused thinking would have developed in the study of lens-film and lens-electrical combinations. The fact that all lenses with significant fields of view have their imagery limited by aberrations makes possible a very varied performance of the same lenses when combined with different types of observing receptors. Therefore, it cannot be too strongly emphasized or frequently repeated that when the image of a point source is an irregular spot in which the distribution of energy is a function of specific residual aberrations, a variety of interpretations of its form and size can be made by energy receivers of different chromatic and contrast sensitivity and that hence a number of different standards of performance are possible.

In the beginning, astronomy presented the most important demands and the requirements of this field together with limitations of theoretical understanding and available optical materials governed the type of systems developed. In observing stars it is not essential, even though sometimes desirable, that the telescope cover a large field of view. In any case, a very large amount of astronomy can be carried out by instruments having quite limited fields. At the same time it is of prime importance that the quality of the image be of the very best. It so happened that with these conditions it was possible to develop lens systems of quite simple design with available materials that would form on the axis an image limited only by diffraction effects. Because of the use it was natural enough to use simple visual experiments to determine the quality of the image formed by the finished product and to judge in large measure the success of the design by the ability of the system to separate visually the images of two closely placed point sources in the object space. Diffraction theory permitted the mathematical prediction of the intensity distribution in the image of a point source when no aberrations were present and an ultimate standard of performance could thus be set. The ability of an optical system to separate point sources came to be called resolution and Lord Rayleigh's name is associated with the standard mathematical definition of this property when diffraction theory alone determines image quality. Resolution, in the absence of aberrations and except for its dependence on the wavelength of the light employed, came to be considered as an intrinsic property of the geometry of the system. Since the range of wavelengths used in such optical systems is generally not large, the persistent tendency to slip erroneously into thinking of visual resolution as an intrinsic property of optical systems was not too calamitous. However, it was most unfortunate that this mode of thought had become a habit when consideration had to be given to photographic objectives in which images are limited by residual aberrations.

When aberrations are present the distribution of the energy in the image, both as to intensity and wavelength, is dependent on the particular residual aberrations left by the designer and the quality of the light forming the point source. This energy distribution will be a

function of the particular design and it is difficult to define the function completely and accurately on a theoretical basis. The distribution will vary with position in the field. The position of best focus or the position of the smallest size of image will vary with the position in the field of view and the position can be different for different means of observation—eye, photographic emulsion, photocell, etc. When the lens is used in photography, it is practically essential that the field of view be recorded on a flat plane. Since the position of best focus varies with the angular position in the field of view, a problem arises as to how to select as the focal plane for a particular lens-emulsion combination the one that will give the best compromise between the size of spots in different positions in the field of view. The interpretation of the energy in the image plane for a photographic emulsion will not be the same as that of the eye because of different response characteristics. It may see either a larger or smaller spot. In consequence the lens-emulsion combination may be able to separate two nearby point sources with less or more facility than the lens-eye. The position of maximum brightness of the spot may appear in different places to the eye and to the photographic emulsion. Consequently photographic records of distortion might differ from visual ones. There are emulsions with different sensitivities and each will give its own interpretation of the image as to size and apparent energy distribution. Other types of receptors give other interpretations of the same energy distribution because of their own peculiar responses. It becomes immediately apparent from these considerations why a lens of high visual performance is not of necessity equally satisfactory for photographic emulsions or for forming a signal to be transmitted electrically.

The foregoing shows that the essential requirement for judging the quality of images in the presence of aberrations is that the judgment be based, not on any particular type of energy receiver for observing the images, but on that particular energy receiver with which it is proposed to couple the lens for some definite scientific task. It is not to be taken from this that visual observations are to be outlawed in making studies of photographic objectives. It does require, however, that if for convenience a visual method is desirable a series of experiments must first be done to show the equivalence of the visual methods to those of practice. The same applies to the substitution of any other kind of receptor for the observations. The equivalence once established holds only for one lens type.

One can point out that for well-designed photogrammetric objectives no difference has yet been shown between the distortion as measured visually and the distortion as measured photographically. On the other hand, it is reasonable to suppose that it would be possible for a significant difference to exist in the case of lenses of inferior design or poor workmanship. Accordingly, if the purpose of the test is to establish acceptability of the new lens for photogrammetric purposes, it is more reasonable because it is safer to determine photographically the distortion of a lens of unknown performance.

The plane of best photographic focus can differ from the position of best visual focus even in a good lens. Consequently this plane must be initially selected for a particular lens on the basis of a photographic investigation. Such an investigation is tedious, and it is fortunate that for a particular type of lens it seems that the plane of

best photographic focus once determined can be located for other lenses of the same type by reference to the position of best visual focus on the axis. Preferably, several lenses of a type should be used to determine the relationship.

Inasmuch as some earlier visual methods of image assessment are shown to be generally inappropriate for lenses intended for photography and other specialized purposes, it is well to consider whether the long used resolution criterion is a suitable one for grading the performance of different types of photographic objectives. We must make the decision in the light of the use. In the case of lenses used for photography that will be the basis of photogrammetric routines or reconnaissance the collection of information is of prime importance. For this purpose the resolution criterion seems very apt. The ability of the lens to reveal information must be very intimately and directly related to its ability to resolve adjoining fine detail. There may be some question as to whether the relationship is a linear one. Indeed, there are good grounds for thinking that in air photography *useful* information does, to some extent, come in size ranges. Thus, a 10-percent increase in resolution might be relatively unimportant in small scale photographs for forestry interpretation since such an increase does not make possible the recording of another size group of useful detail. On the other hand, at large scale when a built-up area is being photographed an increase of 10 percent in resolution might very materially increase, by a much larger percentage, the amount of useful information available for interpretation. Although accepting the validity of resolution as a method of evaluating a lens for air photography, it can still be conceded that experiments designed to relate such measurements to photographic interpretability may be of considerable general interest. Nevertheless, it seems unlikely that these experiments will either disprove that a lens of higher resolution will give more information than a lens of low resolution or disturb the continuing practical demand for lenses of higher and higher resolution.

An Acceptable Testing Procedure for Aerial Photographic Objectives

Keeping the previous discussion in mind we can propose suitable tests for photographic objectives intended for aerial photography. Different but similarly conceived tests can be proposed for lenses intended for other tasks. The procedures set forth in the following are those currently used for aerial photographic objectives in the National Research Laboratories of Canada.

A good quality collimator having a focal length at least four times that of the objective to be tested is used to place the resolution target at infinity since the object distance in aerial photography is always the equivalent of infinity. The excess focal length of the collimator guarantees that its aberrations will be insignificant compared with those of the lens under test. The quality of the illumination of the target is mean noon sunlight since this is characteristic of aerial-photographic operations. In front of the collimator is a sturdy beam that swings about an axis near the end closest to the collimator. The test lens is mounted on the beam over this axis so that its entrance pupil is flooded with light at all angular positions of the beam. The

latter also carries a plate holder that can be placed at varying distances up to 60 inches from the test lens. The test lens carries the filter or filters that will generally be used with it in practical operations. This will include not only colored filters but any filter that is used to equalize the intensity across the field of view of the test lens. The lens can be adjusted so that the film or plate plane is parallel to the machined surface of its mounting flange. Provision is made for maintaining the film flat either by a register glass or a suction back. The plate holder covers a narrow strip of the field of view from corner to corner of a 9- by 9-inch format. Recordings of the collimator target are made photographically at a number of angular positions in a series of parallel planes in the neighborhood of the visual focus. The emulsion to be used with the lens will be the one to be employed in air operations. The film strips are processed according to the recommended procedure under careful sensitometric control. The exposure time at each position of the field is the same. It is chosen so that the density produced at each position favors the negative material yielding the maximum resolution. If vignetting is severe, positions representing smaller areas of the format are sacrificed. After processing the film strips the resolution is assessed with the aid of a binocular microscope. The magnification and illumination are so adjusted that the reading of the maximum possible resolution is favored. From these data through-the-focus curves are plotted for each angular position in the field. From values taken from smooth curves based on these points off-axis graphs are plotted for the several planes studied. The resolution is weighted on an areal basis. The resolution at each position of the field is assumed typical of the annulus in which it occupies the central position. Appropriate allowance is made for the fact that the outer annuli are incomplete because the camera format is rectangular. A final curve can be drawn, which relates average resolving power over the rectangular format for a series of planes parallel to the mounting flange. From this curve the plane of best average photographic focus can be selected.

Visual observations are taken to determine the position of best visual resolution on the axis. The separation of this position from the plane of best average photographic resolution is used to permit routine focusing of lenses of the same type by a simple visual routine.

The Shape and Contrast of the Resolution Target

No reference has so far been made to the shape and contrast of the target best suited to lens resolution studies. Both these factors have a bearing on the actual numerical quantity that will be obtained for the resolution. Therefore it is highly desirable to obtain some measure of general agreement between laboratories on the exact shape and contrast of the target so that values obtained by each may be immediately compared. Unfortunately, there is very wide disagreement on these points. Most targets will place lenses in the same order of merit, but there will be exceptions. These are important ones for performance testing because the proper design of a target will be more severe on inferior lenses. Contrast of target for certain uses of lenses is more important than shape of target. For performance tests in aerial photography, if testing is done at only one contrast, low-contrast targets should certainly be used. This is essential

because of the very great preponderance of low-contrast detail in scenes from the air. However, it is certainly of advantage to do tests also at high contrast, providing the significance attached to these results from a performance point of view as compared with results from low-contrast targets is in proportion to the relative frequency of high- and low-contrast detail in an air photograph. Perfect lenses would be placed in the same order by all shapes of targets at all contrasts and it is again the residual aberrations that complicate the problem of assessment.

The correct basis of choosing a target can be simply stated and there is unlikely to be serious disagreement on the point. The form and contrast of the target should conventionalize the task in the practical application of the lens. It should be as simple as essential requirements allow to make testing an easy routine. In air photography this means that it should attach equal importance to edges in any direction and that it should take account of the mutually destructive effect at the limit of resolution of neighboring edges at any orientation with respect to each other. It should be possible to associate some degree of recognizability with the limit of resolution. It should not make demands for qualities that will serve no useful purpose in the application of the lens or production difficulties may be caused unnecessarily. Important requirements must take precedence over lesser ones.

Some interest has been shown in letters and other targets calling for a greater degree of recognition than the more conventional geometrical forms. It is argued that they have the merit of presenting the observer with an unknown target. The value of this is questioned, since in very extensive observations on resolving-power targets in our laboratories practically no discernible tendency to cheat unconsciously has been noted. The establishment of the limit of resolution of adjoining edges is after all the prime requirement whether the edges constitute a conventional target, a letter, or some natural object in the landscape. It therefore seems unprofitable to complicate the testing for no profitable result. In using letters there is also the further complexity or uncertainty that the recognizability of a given letter will vary with its orientation in the field of view because of the asymmetry of the spot-like image of a point source, which due to aberrations is formed by a photographic objective. It therefore appears that there is no advantage in going to great complexity. A single, simple target conventionalizing to a reasonable degree the practical task will fully suffice.

We think we have met the target requirements outlined in a satisfactory way for routine lens performance tests in the laboratories of the National Research Council of Canada. We have chosen a bright annulus form of target on a darker background. The thickness of the annular ring is equal to the diameter of the central area. In the recorded images of this target resolution is considered to exist as long as the image appears as a line of density enclosing a central light area. If the line of density is broken by astigmatism or if the central area disappears, the limit of resolution is exceeded. This accepts resolution with some distortion as having adequate recognizability. The target has the advantage that it does not place any particular importance on edges in any given direction such as the radial or tangential. This, as stated earlier, is certainly the condition in actual operations where equally important detail will be lying so that its edges are in

any direction at all. The simple single form of this target as compared to lines in the radial and tangential directions, or in more directions, permits an experimental averaging of the various factors. It thus avoids the necessity of making an average on some arbitrary arithmetic basis of the resolving power in two or more specific directions. The mutually destructive effect at the limit of resolution of nearby edges lying in different directions, which obtains in any aerial photograph, is rather adequately represented by the mutually destructive effect of small sections of the inside of the annular ring. Resolution measurements on lines in specific directions take no account of this important factor which is particularly important in the case of poorer lenses. At the limit of resolution the aberrations in all directions at the inside edge of the annulus add up to destroy resolution and cause the disappearance of the central area. It has been definitely shown in the case of lenses of inferior design that the annulus type of target records a relatively more rapid fall-off of resolution towards the corner of the photograph than radial or tangential lines. Under such circumstances it is reasonable to suppose that the resolution picture given by the annulus target is more representative of what will occur in air photography than the performance suggested by the resolution on radial and tangential lines. This revealing property is of obvious value in testing new lenses of unknown performance.

Selwyn and his coworkers first suggested during the last war that because the great majority of aerial detail is of low contrast it would be sensible to adopt a low-contrast target for the assessment of aerial-photographic objectives. Our laboratories in Ottawa have accepted this proposal and our annulus targets have the log contrast of 0.2 suggested by Selwyn. Recent work by Carman and Carruthers has given extensive experimental authority to Selwyn's proposal and has further shown that the typical average log-contrast ratio for a great variety of scenes is considerably lower than 0.2. A log contrast of 0.1 might be considered more typical. It is not thought that any great increase of validity will be obtained by substituting a 0.1 target for a 0.2 target in resolution tests. The difficulty of making the tests would be much increased. Nevertheless, tests are under way in our laboratories to make certain that this conclusion is warranted.

It is sometimes argued that high-contrast targets reveal differences between lenses more readily than low-contrast targets. This contention is undoubtedly true. The fact in itself is not of great importance unless the differences revealed are of real significance in the work to be assigned to the lens. In an aerial scene the amount of high-contrast detail is but a small fraction of the low-contrast detail. It seems unreasonable to evaluate the results obtained with high-contrast targets in any other proportion but this fraction. For aerial photography a lens that has 120 lines per millimeter on a high-contrast target and 15 lines per millimeter on a low-contrast target is not by any manner of means twice as good as a lens that has 60 lines per millimeter on high-contrast target and 15 lines per millimeter on a low-contrast target. Indeed, it is only very slightly better. On the other hand, if the low-contrast figure for the lens performing in the poorer way at high contrast were 25 lines per millimeter instead of 15 lines per millimeter it certainly would be nearly 66⅔ percent better than the other on an overall basis. Such a lens as that postulated is highly desired at the present time.

The Distinction Between Two Types of Testing

The relative merits of performance testing and tests that are designed to study one or more of the various physical factors that contribute to overall performance are quite often vigorously argued. Such arguments are quite pointless since the two kinds of tests serve quite different purposes. The typical designer requires, for the development of his science, physical information specified by him, which will help him to improve his methods and his designs so that his product can meet performance tests. The choice of information must be his. The user of the lens on the other hand is not concerned with these design problems. He wants to know how well a particular lens will cope with the problems with which it will be faced in a particular task in which he is interested and not some other task. Since all optics is generally a compromise one cannot usually meet all working requirements equally well. The two types of experiments may well be quite different in form since they are different in purpose. One is really an investigation and the other is a test. Neither type of work should have its usefulness limited by confusing the two separate requirements and the user must be left as free as is the designer to specify the testing requirement that best suits his problem. In our laboratories we are interested in both types of work. Experiments of the second type are at present going on, but no reference will be made to them here since the concern of this paper is to emphasize the proper general requirements for performance tests and to treat more specifically those required for photogrammetry and air reconnaissance. Agreement on performance tests does not need to await the solution of the long term problems of improving image formation. Performance tests must judge the end result in the most informative way and need not contribute directly to improvement of design.

Although performance tests do not of necessity contribute aid to the designer in his efforts to meet the user requirement, those outlined here have in fact made such a contribution. It is a general one but of the greatest importance. Because it does not point a detailed path it does not seem to have commended itself to many designers.

By contact printing a resolution target on to Aero Super XX it can be shown that Aero Super XX will resolve about 25 lines per millimeter with a line target having a log contrast ratio of 0.2. For the many aerial photographic objectives that have been tested in our laboratories in combination with Aero Super XX, the effective maximum resolution of the combination turns out with monotonous regularity as about 15 lines per millimeter. This is approximately 60 percent of what the emulsion itself can do. The apparent visual resolution may be 150 or more lines per millimeter and will vary greatly from lens to lens and type to type. The high-contrast photographic results may also be quite high and varied. It is suggested by these results that the designer is endeavoring to get some fraction of the energy from a point source into a very compact spot in the image space and letting the rest fall where it will outside this spot. Visually this central spot is sufficiently concentrated, compared with its surroundings, to yield quite high resolutions. It is often of the order of

hundreds of lines per millimeter. The resolution may even be quite impressive photographically for high-contrast targets on Aero Super XX. However, for low-contrast targets on Aero Super XX, the light that falls outside of the central compressed spot is very significant, and in consequence very substantially lower resolutions are obtained. Much better results, and a far more useful lens, would be obtained if the designer ceased from making a concentrated spot from a relatively small fraction of the light for visual interpretation and concentrated his attention on packing nearly all the light to which the emulsion is sensitive into a spot of quite significant size but one that would correspond to a resolution of 25 lines per millimeter on Aero Super XX with a low-contrast target, mean noon sunlight and a minus blue filter. In such a case we would have a lens ideally suited to work with this particular emulsion in air photography. Present lenses have characteristics that fight with those of the emulsion. Since there is little hope of changing the emulsion the lens designer must change his design criteria.

It is of interest to note that some time ago a number of metrogon lenses were tested by the methods outlined earlier and that their performance on low-contrast targets were very uniform. Visually and photographically at high contrast there were considerable differences. Nevertheless, no distinction has ever been made between these particular lenses by those continually using photographs from them for photogrammetric purposes over a period of years. This would seem to be practical subjective evidence to justify concluding that the low-contrast results are judging the important factors of performance in so far as the mappers are concerned. It has also been found in our laboratories that designs said by designers to be sensitive to certain curvatures, separations or surface figures are in point of fact quite insensitive to one or all of these factors with Aero Super XX, low-contrast targets, mean noon sunlight and minus blue filters. Again, this suggests that there is a straining to attain something that gives little profit in the practical photographic problem although the designer's concern about close tolerances is shown to be fully justified on visual performance of the same lenses.

Conclusions

The foregoing discussion leads to some important conclusions. Performance tests of lenses must be based on the condition of use. They must not be confused with experiments that the designer may require to obtain information on ways to improve his product. Performance tests for aerial photographic objectives must take account of the quality of the light and the characteristics of the instrument that will record the image. Resolution tests are generally an adequate means of testing the performance of these lenses if the main purposes is to secure information. A target for resolution tests must conventionalize, both in form and contrast, the task to be assigned to the lens in practice. Results of such testing suggest that we would get much better lenses for air photography if all the energy from a point source could be concentrated in a spot in the image space that would give, at any position in the useful field of view, 25 lines per millimeter on a

target having a log contrast of 0.2. On the basis of tests carried out on most of the available aerial photographic lenses this is obviously not being done at the present time although enough energy is being sufficiently concentrated into a smaller spot to lead the eye to observe a much larger resolution because of its ability, by reason of its response curve, to ignore to a sufficient degree radiation that falls outside the main concentration.

5. Quality Aspects of the Aerial Photographic System

By Duncan E. Macdonald [1]

As the work at the Boston University Optical Research Laboratory is primarily concerned with the field of aerial photography, the considerations of this paper have been directed to that field. However, it is felt that aspects of this discussion may be considered for more general application.

In aerial photography, pictures are taken of ground objects through long air paths by cameras, lenses, and associated equipment mounted on unstable and translating platforms. The resulting photographic negatives, consisting of nonhomogeneous clumps of silver, arranged in numbers and sizes as to approximate the geometric pattern of reflectivities of the ground objects, are then printed. These positive prints are viewed by trained observers known as photointerpreters, generally under low-power stereoscopes. These interpreters then perform the important transduction process—that of converting the message coded in silver clumps into a verbal report. This system is shown in figure 5.1. This paper will discuss the role of several of these components in the system performance.

Aerial photography has as its objective the gathering of information and, thus, the capacity of this photographic system for revealing information is taken to be a meaningful performance criterion. It is perhaps in order to review briefly the development of a common quality criterion that is now in general use, namely that of resolution. In 1850 Dawes [1] [2] reported on the result of his experiment on the conditions for separation of two stellar images, and in 1879 Rayleigh [2] presented a theoretical criterion for the resolution of two point images. Inasmuch as point images are difficult to treat experimentally, it has generally been the custom of the experimentalist since that time to employ test objects consisting of a periodic array of finite geometric details (usually lines) and then to score the resolution of the system in the appropriate coordinate in terms of the fineness of detail that it records (lines per millimeter).

In such fields as astronomical and spectroscopic photography, the original concepts of Rayleigh and Dawes appear to coincide in many instances with the technical objectives. Although at the present time a resolution score is used with some measure of success as a quality criterion for aerial photography, the message presented to the system is of a character quite dissimilar to that of the usual test object. It therefore appeared in order to examine how this and other criteria related to the performance of the system.

In the many approaches to the problem of photographic performance, there is, in general, one point in common—the consideration that the limit of resolution of the system is that point where contrast
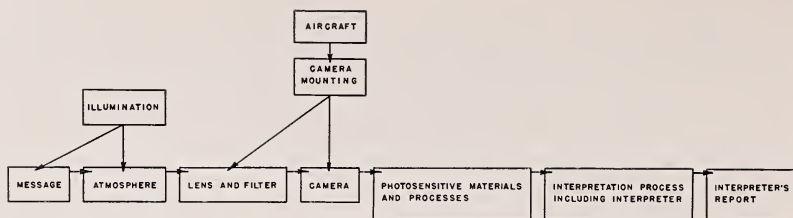
Figure 5.1. *The aerial photographic system.*

is degraded to a value below that of the visual contrast threshold. This implies that resolution is a function of the degradation of contrast of a system and suggests that a contrast-reduction function might provide some insight for the judgment of aspects of the system performance.

Consider the resolution limit of a lens-camera-film combination determined in the laboratory as being the limit of a band-pass filter. The nature of a vast majority of man-made objects, and indeed of the majority of natural objects of aerial photographic interest, is such that perfect imaging and recording would reveal square waves of intensity. (This is one characteristic of the message.) However, if square-wave detail is incident on the filter as is observed in the laboratory when employing a parallel-line test object, the shape, amplitude, and mean power of the waves become more and more altered as one approaches the band-pass limit.

The filter concept coincides with the one-dimensional aspect of the resolution target. For example, in the case of scanning pictures in a line-scan, video-type system, the limit of frequency of the system is VR, where V is the velocity of scan in millimeters per second and R the resolution in lines per millimeter. Whether the frequency characteristic is spatial, as in the case of a between-the-lens shutter, or is temporal, as in the line-scan application, the resolution number is directly proportional to the limiting frequency for a given set of parameters.

A photographic system may, then, be described in terms of its response to (a) detail that is of ordered occurrence, as in the case of the resolution target, (b) detail that is isolated, and (c) detail that is random, as, for example, the spatial frequency characteristics of the objects on which the system is to be employed as a detecting device. Because of the mutual influence of areas in proximity due to the spreading of the energy by the system components, particularly by the emulsion, the contrast-reduction functions differ for each of the above classes. (Analogy may be made to network theory, where in linear systems the steady state, transient or random responses uniquely determine the characteristics of the system.) Schade [3] has treated optical elements as low-pass filters and discussed the use of a "transfer characteristic" to describe the performance of a system component in terms of temporal responses for a given set of operational parameters.

Such characteristics in the aerial photographic system have been recognized as contrast-reduction functions. Contrast-reduction characteristics of a 6-inch $f/6.3$ Bausch and Lomb Metrogon, a 12-inch $f/5.0$ Eastman Kodak Aerostigmat, and a 24-inch $f/6.0$ Eastman Kodak
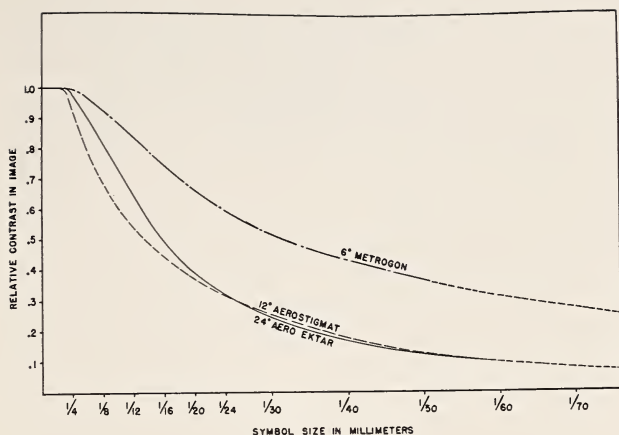
52

FIGURE 5.2. *Relative contrast in image.*

Aero-Ektar have been studied in the laboratory. Image contrast has been defined as $C = (T_b/T_l) - 1$, where T is the transmission as measured on the negative, with subscript b denoting background, l denoting line or object. Data were obtained from photographs of line patterns, employing a photographic plate inclined to the focal plane, as discussed by Zschokke [4] and later modified by Wetthauer [5]. The length of the target lines falls normal to the axis of tilt.

Figure 5.2 shows the contrast-reduction characteristics as measured on Super-XX emulsion for the case of isolated symbols, that is, isolated white lines on a dark field. The contrast is expressed as a function of image size.

Because of the nature of reduction in contrast, it is clear that in order to detect a symbol, more contrast is required in the object as the image becomes smaller. The limiting contrast expressed as a function of image size may be thought of as the contrast-detection threshold of the lens-emulsion system.

The quotient, visual contrast threshold (under the conditions of viewing) divided by the percentage contrast reduction for a given image dimension, is the object contrast threshold for this dimension of image. Assuming conservatively a visual threshold of 0.1 the data from figure 5.2 are taken to show the detection thresholds of these lens-emulsion systems. These are presented as figure 5.3. It is clear that this concept breaks down when the image becomes effectively a point. In the same figure the systems are compared for the case of symbols occurring at regular spatial positioning (resolution charts). It should be noted that this function for periodic symbols approximates that for the case of isolated dark lines on a bright field. In this latter case we have assumed that the visual contrast threshold varies linearly with the reciprocal image size from 0.1 at 2 symbols per millimeter to 0.4 at 22 symbols per millimeter. The 0.4 threshold appears in reasonable agreement with laboratory measures. The validity of the assumptions bears little on the main point. It is clear that there is a marked difference in the lens-emulsion performance for the two types of symbols.
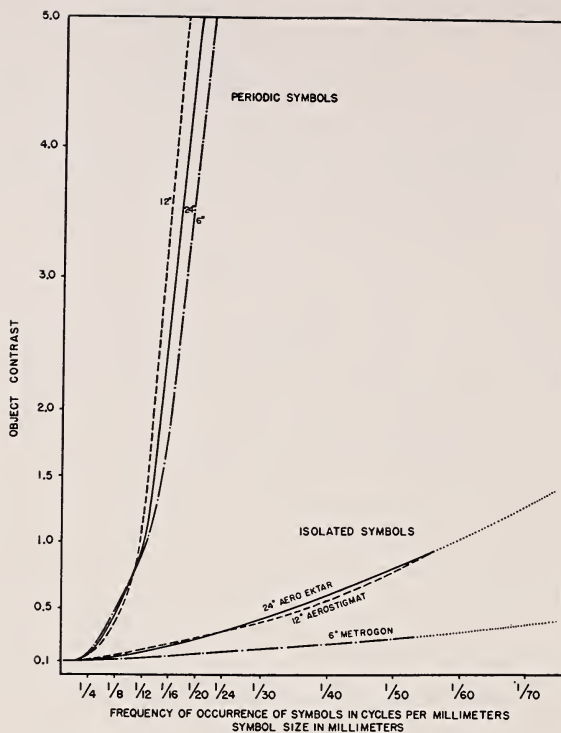
FIGURE 5.3. *Object contrast thresholds.*

The curves shown in figure 5.3 may be taken as representative of the maximum contrast capacity of the lens-emulsion combination for the typical aerial scene, as background and object are located on the toe and straight-line portions of the photographic characteristic, respectively. The test-object brightness ratio is of the order of 30 : 1.[3]

Let us now assume for purposes of discussion that the object space consists of a universe of isolated symbols. Figure 5.3 describes the response of a lens-emulsion system to this type of detail. We may assume these symbols are distributed such that for any given region $(\Delta x)(\Delta c)$ the same number of symbols occur. Then, in a single exposure the system explores a portion of the detail universe within the area Δ,

$$\Delta = kC_R(f\theta - x_L), \tag{1}$$

where C_R is the contrast range in the scene,[4] f the focal length, θ the angular coverage of the system, x_L the limit of detail size passed by the system, and k a constant that includes a reciprocal scale factor.

[3] The maximum brightness range encountered in an aerial scene, that is, highlight to shadow, is about 150:1. This maximum range is reduced by atmospheric haze to less than 30:1 by only a few hundred feet of air path.

The data employed for this presentation have been taken from film processed to a gamma of 0.98 ±.02. If the sensitometric data are known, it is straightforward to predict the alterations in the form of the curves (fig. 5.3) for any changes in median exposure level for other conditions of processing, or for any specific object if its photometric characteristics are known.

[4] $C_R = C_0 e^{-\beta h}$, where C_0 is the maximum contrast in the object, β the attenuation coefficient, and h the flight altitude.
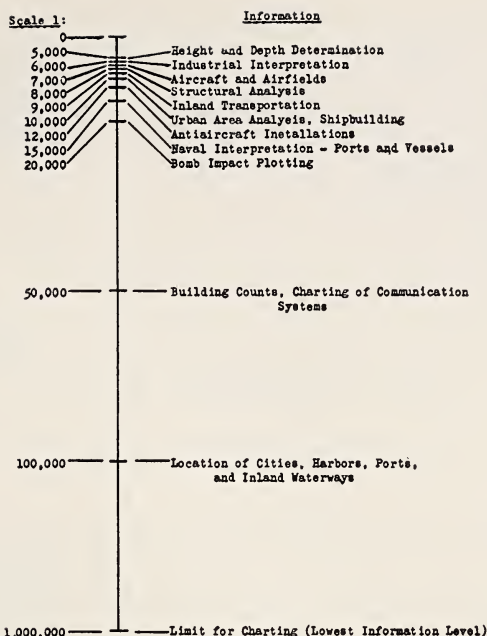
54

```
Scale 1:                              Information
     0
   5,000                    Height and Depth Determination
   6,000                    Industrial Interpretation
   7,000                    Aircraft and Airfields
   8,000                    Structural Analysis
   9,000                    Inland Transportation
  10,000                    Urban Area Analysis, Shipbuilding
  12,000                    Antiaircraft Installations
  15,000                    Naval Interpretation - Ports and Vessels
  20,000                    Bomb Impact Plotting



  50,000                    Building Counts, Charting of Communication
                                       Systems



 100,000                    Location of Cities, Harbors, Ports,
                                     and Inland Waterways



1,000,000                   Limit for Charting (Lowest Information Level)
```

FIGURE 5.4. *Information-level diagram for aerial photographic interpretation.*

It should be noted that the system may be tuned with respect to the detail universe. The tuning is accomplished by varying the altitude of flight, that is, by varying k, which enables the system to explore any preselected range of detail size.

If the contrast threshold of a particular lens-emulsion system (fig. 5.3) is described by $\psi(x)$, then the probability P that any particular unit of detail within the region will occur above the threshold may be written

$$P = \frac{\Delta - \int_{f\theta}^{x_L} \psi(x)dx}{\Delta}. \tag{2}$$

P may then be thought of as a measure of the range of symbols subject to detection through the use of the system.

At the present time, the type of significant detail for a given form of reconnaissance is of such a nature that military photointerpreters state their requirements for photographs solely in terms of scale. Whatever the significant quality criterion may be, the interpreter has achieved, through his experience with many analyses, a concept of a quality criterion that is meaningful to him. In limiting his requirement to a statement of scale, the interpreter tacitly assumes that the other quality aspects will not differ greatly from the present-day average. Thus, it appears that in military aerial photography, as in many other kinds of photography, levels of information are quantized, that is, changes in emission of information from aerial photographs appear to occur at discrete scale levels. This quantized aspect of the aerial photographic problem, as shown in figure 5.4, should not be
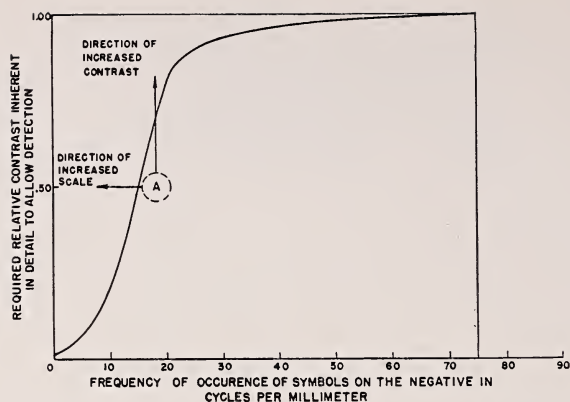
FIGURE 5.5. *Location of significant symbols, "A", in a photograph of not quite adequate quality for a particular purpose.*

taken to be more than an illustration of the gross character of information levels. These levels really define band limits and not discrete lines.

It is more efficient, for all types of intelligence information, to work near the upper limit of the allowable scale, since decreasing the scale reduces the area coverage per exposure. It is not, however, without sound basis that scale has evolved as a quality criterion. Aerial photography has been restricted, at present, almost totally to the use of Super-XX emulsion for work in the visible spectrum. A limiting ratio, detail size/grain size, thus has a constant denominator. An improvement of the detection characteristic for a given system is thus achieved by increasing detail size. This same conclusion holds in comparisons between present systems because of the emulsion dominance of the contrast-reduction characteristic in these systems. It has been suggested that those symbols of significance to a particular photo-interpretation task must be so located as to have a high probability of being above the incident contrast thresholds [6]. Thus, one might expect that very similar distributions in size and contrast apply to those symbols that are significant for many types of interpretation. In photographs that are not quite adequate for a particular job these symbols, as illustrated by A in figure 5.5, are submerged below the threshold. The most common method to assure bringing them above the threshold in the repeat photography is to tune the system by increasing the scale. This translates the threshold-characteristic origin to the right with respect to the fixed position of the symbols in the detail universe. The translation can also occur in the Y direction by refocusing (or a change in processing) to drop the threshold characteristic below the significant details.

An effort has been made to study the statistics of aerial photographic targets and, in particular, to attempt to isolate characteristics of the significant symbols for a photointerpretation task. In one approach to this problem, many of the using agencies of aerial photography were contacted. For example, agencies employing aerial photography for forestry purposes have been asked to submit representative prints of aerial photographs taken for their work, which they found to be
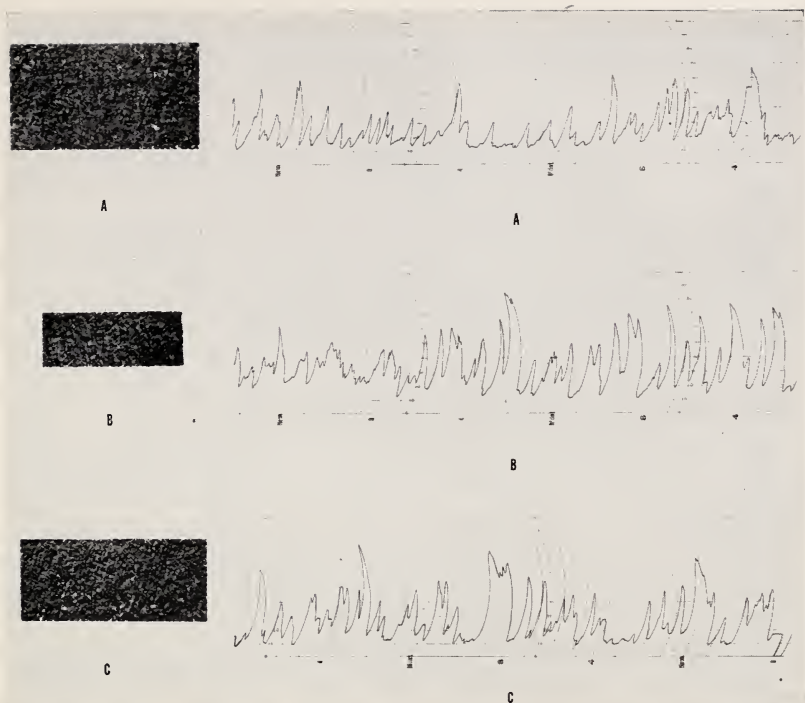
FIGURE 5.6. *Enlarged sections of photographs used to determine classification of forested land as to species, stand height, and stand density.*

Classified by the using agency as: A, Not quite good enough; B, just good enough; C, excellent. (Corresponding microphotometer traces to the right.)

not quite good enough (A), just good enough (B), and excellent (C) for a particular purpose in forestry survey. The purpose of the photography was also requested. By this approach, it was felt not only that the statistics of this type of message could be found, but that some insight might be gleaned as to the nature of those significant symbols that the photointerpreter requires in order to achieve success in the particular job. Figure 5.6 shows areas from three prints submitted by the Great Northern Paper Co., Bangor, Maine, and classified by the user in terms of A, B, and C. Scale of the photography was 1:15,840. The photographs were employed by GNP to determine species, stand height, and stand density. Six additional photographs, reported on here, were submitted by the Air Survey's Engineer, Land and Forest Department, Victoria, B. C., obliques taken with a 3.25-inch lens from altitudes of 17,400 to 20,000 feet. This photography was used to extend control over large areas of unsurveyed country. Analysis was conducted in the region of a 60-degree depression angle.

Microphotometer traces were made over three portions of each print. Samples of these traces are shown in figure 5.6, to the right. The scanning aperture was set an order of magnitude below the limit of resolution under normal viewing conditions. Although the scale of the photography differed from set to set, the areas traced were in all

cases over the image of wooded portions of the terrain. It is assumed by this restriction that this type of photography involved a message with circular symmetry; thus, the resultant plots should be similar for any coordinate of trace.

The analyses were made from the traces. An edge was scored for each case where a change in brightness (reflectivity) equal to or greater than 10 percent occurred (that is, $\Delta B/B=0.1$). A marked change of slope was requisite before the next edge was recorded. The 10-percent threshold value was chosen as representative of a weighted threshold after consideration of the limit of visual contrast threshold on the photographic material at the resolution limit, the range of brightness adaptation levels encountered in the interpretation process, and Blackwell's [7] liminal contrast data. If absolute rather than comparative values were required, it would be necessary to correct the data to allow for the ability of the eye to see edges at lower contrast in the case of larger details.

The close agreement from print to print within any user classification, regardless of scale or purpose of photography, seemed to justify the lumping together of the data from all prints under any one classification.

A plot of the relative frequency of occurrence of edges as a function of distances between them shows no significant difference with picture quality (fig. 5.7, A, not quite good enough; B, just good enough; C, excellent). In this and the following figures each plot deals with nearly 4,000 edges.

The data on reflectivity differences across edges, shown in figure 5.8, indicate a marked difference with picture quality. The better the picture the more skewed is the distribution of reflectivity differences toward the greater differences. From this one infers that the symbols are recorded with greater contrast (that is, with less contrast reduction by the system) in the better pictures.

Figure 5.9 shows the relative frequency distribution of edge gradients, dR/dx, R being reflectivity. Here the excellent photographs stand out, while the difference between A and B categories is slight, the just-usable photographs having relatively more steep-edge gradients than the non-usable pictures.

Additional determinations of the relative frequency of $\Delta R/R$ and dR/Rdx have been made. In each case the excellent photographs stand out from the others as would be expected, but the difference between classes A and B is again less obvious, apparently dealing with some 10 percent of the total information. Coincidentally, this figure is in good agreement with estimates obtained from experienced scientist-photointerpreters who have expressed the feeling that it is the appearance or nonappearance of some 5 or 10 percent of the symbols that differentiates between usable and nonusable photography.

The atmosphere plays two distinct roles in the degradation of the message—haze and turbulence. Atmospheric haze reduces the contrast. Duntley [8] has treated the theory of atmospheric contrast reduction and reports that all attempts to observe the so-called "ground glass plate" effect or the "edge" effect failed. Carman and Carruthers [9] have reported on their experiments treating the brightness characteristics of the aerial photographic message. Because of this fine work there seems to be little need to devote time here to the role of haze in aerial photography. The flight-test data show the expected reduction in contrast with increasing altitude.
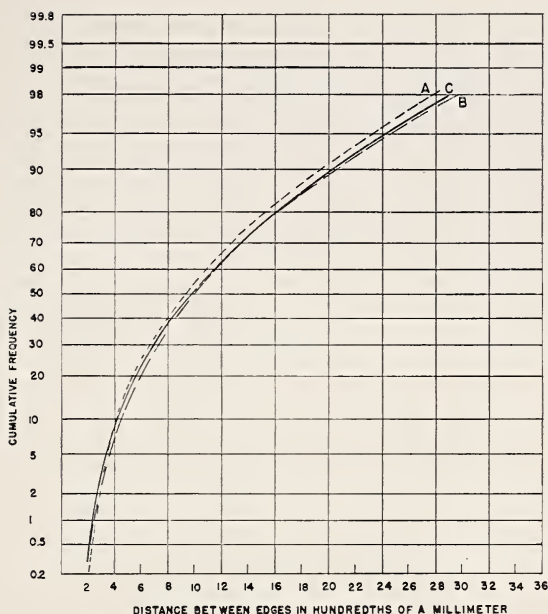
58

FIGURE 5.7. *Cumulative frequency distribution of distance between edges.*
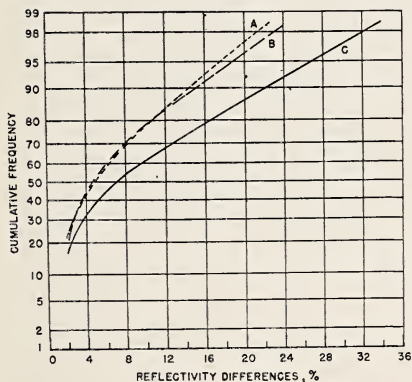

FIGURE 5.8. *Cumulative frequency distribution of reflectivity differences.*
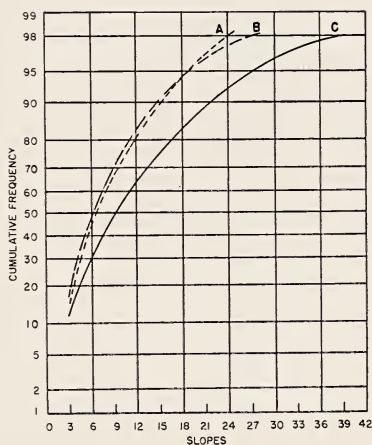

FIGURE 5.9. *Cumulative frequency distribution of slopes.*

On the other hand, air turbulence scatters energy out of the central image, which causes a lower edge gradient and a resultant reduction of image contrast. It appears logical to assume that the limit of definition obtainable in aerial photography will be set by the optical homogeneity of the atmospheric path. The inhomogeneities may be considered to be local motions of small masses of air. These motions are caused by micro-weather conditions and the motion of the aircraft through the atmosphere.

59

The average condition of the turbulence field introduces effects similar to those caused by inhomogeneities in the lens and window glass, effects which deteriorate the image through increasing flare and diffraction (that is, decreasing the image contrast). This average is not stationary for a number of reasons. One, a nontransient phenomenon, is the migration of vortices and masses resulting in changes in size, number, and density (air density) of the vortices in the turbulence state from moment to moment. These changes are about a mean condition. This phenomenon results in a change in image quality about an average. A second, a transient phenomenon, is the change in the orientation of pressure gradients with respect to the optical axis of the camera. This introduces an image motion during exposure which will, in general, cause a deterioration of the image.

Wind-tunnel tests have shown a deterioration of the image with increasing air velocities in a free tunnel [10]. Other tests which perhaps have more significance here have been discussed in that same report. Aerial photographs have been made over a point source of light, a 1-inch-diameter, krypton-filled, helical-coiled tube of an Edgerton flash lamp, cycled at 60 flashes per second. Photographs were taken with a shutterless camera from an altitude of 6,000 feet in a B–17 at an indicated air speed of 170 miles per hour. Point images selected for study were all located in the same position as referred to the optical axis of the system. Inasmuch as exposure times were in the order of 1/6,000 of a second, aircraft translation and other aircraft motions and transitory turbulence effects can be neglected. Thus, the variations in point dimensions from point to point were assumed to be attributable to the changes in the condition of the turbulence state from instant to instant and to variations in the grain distribution in the photographic emulsion. The point images, when photomicrographed, appeared to consist of a high-density nucleus and a low-density surround. Measures were made of the extent of the nucleus and the surround in two arbitrarily selected coordinates.

It was then necessary to determine which part of this variation of dimensions from point to point was due to the inherent emulsion properties, inasmuch as, at this limit, one is dealing with the effect of individual grains and their distribution. This consideration involved a laboratory test employing the same light source, a moving film magazine, the same lens as employed in the air, and an image scaled down by the same factor as in the aerial test and located at the same position with respect to the optical axis. The resulting distribution of point diameters, both for the nucleus and the surround, is shown in figure 5.10.

The distribution of the laboratory nucleus diameters shown in curve 3 affords some insight to the role of grain distribution in the emulsion. It is clear, as one compares these results with those shown in curve 1 (obtained from an aircraft in flight), that not only is the diameter of the central nucleus increased slightly but the dispersion of the diameters is greater than that obtained in the laboratory, due to the presence of turbulence. Curve 4 treats the laboratory case for the diameters of the low-density surrounds, and curve 2 the same diameters for the in-flight case. From these latter curves it is seen that turbulence causes a spread of the surround energy over a wide range. It is estimated that some 60 to 70 percent of the total energy of the point is in this surround.
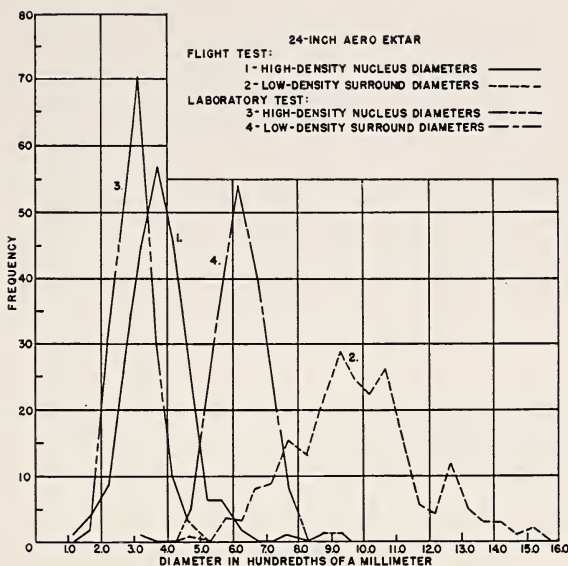
FIGURE 5.10. *Frequency distribution polygons of point-source image diameters on Super-XX emulsions.*

Although lens-emulsion characteristics have been treated previously (figs. 5.2 and 5.3) in introducing the treatment of the message, it is in order to touch briefly on one other aspect of these components. A replot of the data on image contrast, using focal setting as a parameter, indicates a shift in the position of minimum contrast reduction as a function of detail size (fig. 5.11). The conditions of this experiment included full-field illumination. This shift becomes more pronounced as the scattered light is reduced (as may be seen by comparison with previous results obtained with a collimated target [11]) or as one goes to higher-resolution systems.

This is a manifestation of the well-known fact that the focal settings for best resolution and minimum flare (that is, maximum sharpness) do not necessarily coincide. Presented in terms of figure 5.3, this would result in the intersection of the threshold curves for different focal settings of the same lens. It may be stated that focal setting should be selected to provide the minimum threshold, that is, minimum contrast reduction, for the size of symbol under observation. In the case of aerial photography, if the size distribution of the symbols required for operational success be known, then that focal setting which provides the lowest weighted contrast threshold over that size distribution is the best setting. From this it is seen that the system may be tuned in the coordinate of contrast for any given image size by varying the focal setting.

Concerning the photographic emulsion, the concept of microscopic and macroscopic photographic contrast has been discussed previously by Baker [12]. To treat this concept, assume a scattering function
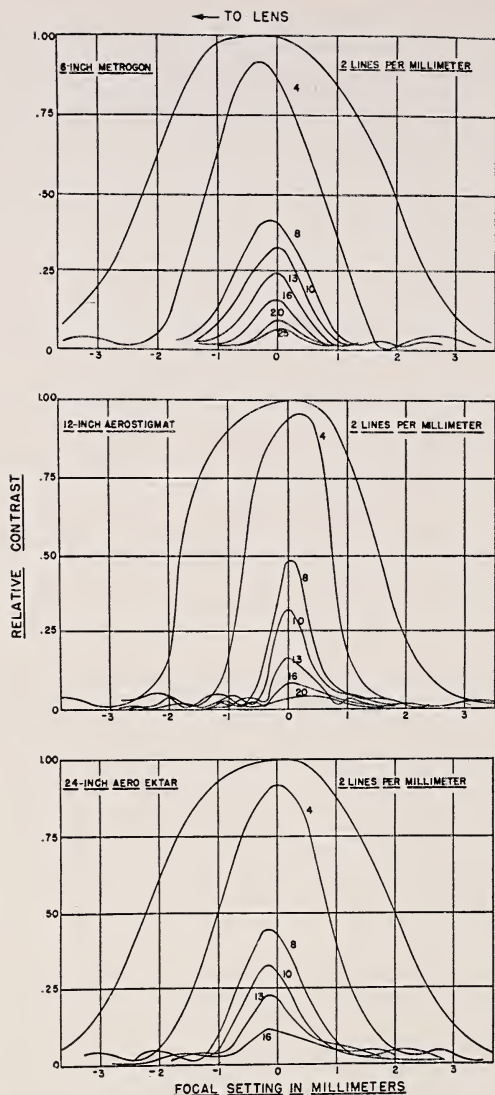
$$i(x) = I_L f(x), \tag{3}$$

61

FIGURE 5.11. *Contrast in the image as a function of focal setting, with image size as parameter.*

where $i(x)$ is the intensity at a distance x units from the edge of a line of intensity I_0 as imaged on the emulsion. The intensity that remains in the line image after scattering is I_L. This assumption implies that the intensity over an infinitesimal line of width Δx may be taken as a constant and the fall of intensity regarded as starting from the edge of the line. Figure 5.12a illustrates this condition for a line of width Δx. The dashed line indicates a distribution of intensity as imaged by a perfect optical system. The solid contour indicates the distribution of that energy due to the scattering of light in the emulsion. The
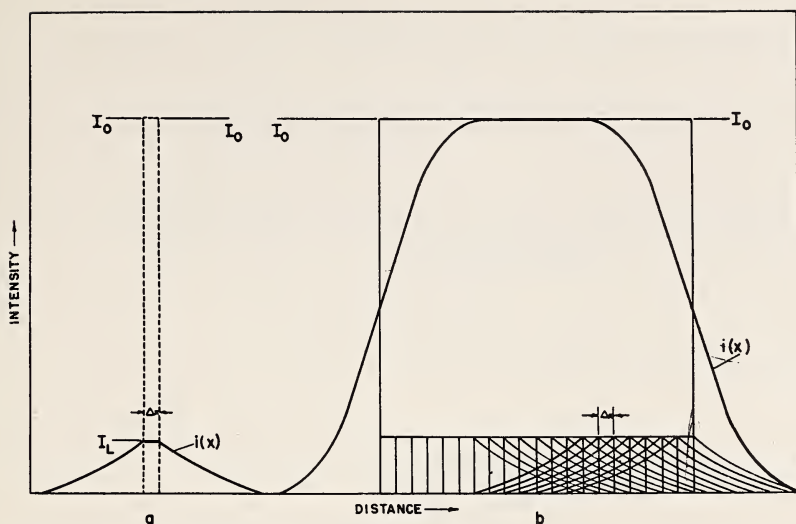
FIGURE 5.12. *Distribution of light intensity from a line assuming exponential scattering.*

intensity remaining in the line, I_L, is considerably less than the incident intensity, I_0.

Figure 5.12b shows the same case for larger detail where the behavior of each line element of width Δx is shown to the same scale as in 5.12a. The ordinates have been added to give the resulting contour of the energy distribution in the emulsion for the gross detail. For a line of this width, $I_L = I_0$ for the central portion. It is seen that there is a critical detail size below which the central portions of the detail image are no longer reinforced by energy scattered from other regions of the detail in sufficient quantity to compensate for the loss of energy from that central portion. Experiments show that for Super-XX emulsion this critical detail size for isolated symbols is about 60 microns. The presence of other symbols in proximity and/or the insertion of any degrading component (for example, the lens) in conjunction with this emulsion increases this critical dimension. The previous statistics, shown by figure 5.7, indicate that some 20 percent of the symbols in the aerial photography examined were recorded below the critical size for the emulsion alone. These symbols are, therefore, no longer in the same sensitometric reference framework, "photographic density: object brightness", as is observed for the macroscopic detail of the photograph.

Redefining I_L as the intensity of the midpoint of the line of width x_0, $\overline{I_L}$ as the average intensity over the line, and taking the origin of coordinates at the point $I = 0$ and the mid-point of the line, then

$$I_0 x_0 = \overline{I_L} x_0 + 2 \int_{\frac{x_0}{2}}^{\infty} I_L f(x) dx. \tag{4}$$

Having made the approximation $(\overline{I_L}/I_L) = 1$, the intensity I_L may then

be compared to the incident intensity I_0:

$$\frac{\overline{I_L}}{I_0}\,(x_0)=\frac{1}{1+\dfrac{2}{x_0}\displaystyle\int_{\frac{x_0}{2}}^{\infty}f(x_0)dx}.\tag{5}$$

The reduction in contrast for symbols occurring at regular spatial frequencies may be treated by an expression of the form

$$\frac{I_L}{I_s}\,(x_0)=[2f(x_0)]^{-1}\tag{6}$$

and for isolated symbols

$$\frac{I_d}{I_0}\,(x_0)=\frac{1}{1+\dfrac{1}{x_0^2}\displaystyle\int_{\frac{x_0}{2}}^{\infty}x_0^2\phi(x_0)dx},\tag{7}$$

where the subscripts s and d imply space and detail, or symbol, respectively, and where, in going to the case of two-dimensional scattering, the original function $f(x)$ is modified as $\phi(x)$.

The shutter [13], aircraft translation and camera vibration [14] have been previously discussed. They reduce contrast in a manner that may be treated by simple geometric considerations and need not be reviewed here.

To assess the problem relating resolution, scale, and contrast criteria to the detection and recognition of photographic detail, a laboratory test setup has been constructed to provide simulated aerial photography. A report has been made on the method and the preliminary results discussed [15]. It is noted that this experiment concerns itself solely with the information on the photograph as presented to the interpreter. The equipment employs a 20-millimeter $f/1.5$ Biotar lens stopped down to $f/5.6$, used in conjunction with a Fairchild Oscillo-Record Camera. As this camera has provision for moving the film during exposure, it allows opportunity to introduce image motion effects common to most aerial photographs. The camera is aimed in a horizontal direction into a 30-inch-diameter, plane, front-silvered 45-degree mirror located directly over a target array. To vary resolution conditions, a turret head containing 13 ophthalmic lenses and a clear aperture is located in front of the Biotar objective. By this means, 14 different resolution values, from about 4 to 70 lines per millimeter, in approximately equal steps of angular resolution, are obtainable on Super-XX emulsion. It is assumed throughout that this experiment represents photography employing perfect optics used at different focal settings. The validity of the assumption is substantiated by viewing microphotometer traces across the edge of two adjoining photometric areas for each resolution condition (fig. 5.13).

The target array is shown in figure 5.14. The test objects consist of squares and circles. The squares are $\frac{1}{4}$ of an inch on a side; the circles are of the same area as the squares. Thus, according to Duntley's work [16], the objects are of equal visibility. The objects are of
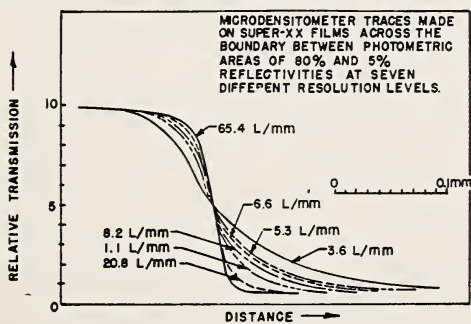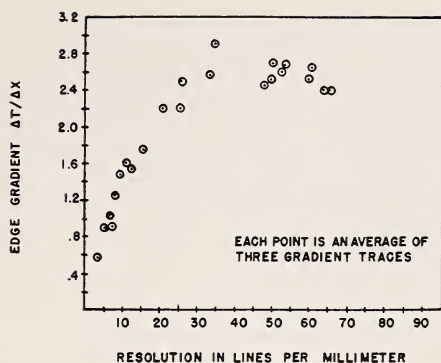
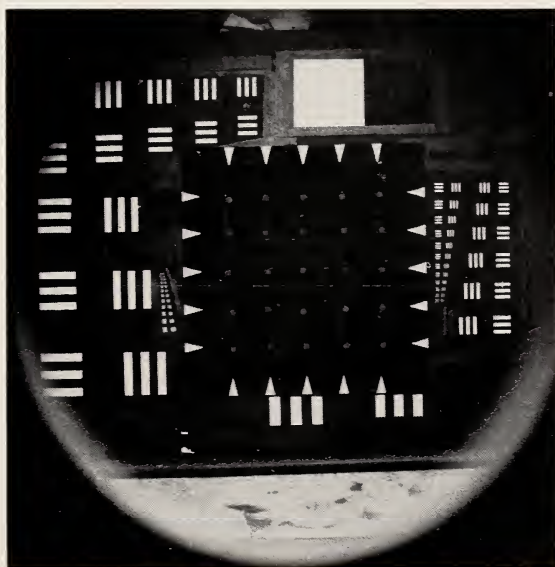FIGURE 5.13. *Edge traces and edge gradient as a function of resolution.*

FIGURE 5.14. *The target array.*

three different diffuse neutral-gray reflectivities, 5, 13, and 40 percent, and, in each case, the background for the objects is also one of these same three reflectivities. Thus, six different combinations of object-background reflectivities are available if the zero contrast conditions are excluded.

For scoring purposes the test objects are arranged on a background that is divided into 25 sections, a 5- by 5-unit grid, numbered down the sides and lettered across the top. The observers record the location of the test object by the proper designation, for example, 2C, 4C, 1A, etc. Forced discrimination is employed in all cases. For detection, a single object is located in each of 10 randomly selected positions of the 25 possible. In the analysis the photo reader is then required to indicate the 10 positions that most probably contain objects. For differentiation, or recognition, 12 circles and 13 squares are located on the target, one in each of the 25 positions. The analysis now calls for indicating the 13 positions that contain squares. This work was performed under 7X viewing with binocular microscopes. Throughout the work, the dimensions are normalized in such a way that the test object is taken as a unit square. Photographic scales are then recorded in terms of object dimension, and resolution is expressed in terms of lines per object. No generality is lost by this scaling of test objects. It is only necessary to observe in any transformation that the ratio of grain size to the detail image size be held constant. It is thus possible to translate these results to any desired dimension only by applying the same factor to both the object dimension and the photographic scale.

The form of a typical object differentiation plot as a function of resolution is shown for the scale of 1:3700 in figure 5.15. Incident brightness ratio is the parameter. As all films have been processed to a gamma of 0.98 $\pm$.02, the image contrast, as presented to the observer, may be tabulated for the case of macroscopic detail. The loss of contrast due to decreasing detail size and/or reduced resolution may then be calculated from system functions or measured on the film. The data of this experiment are tabulated below for the macroscopic case.
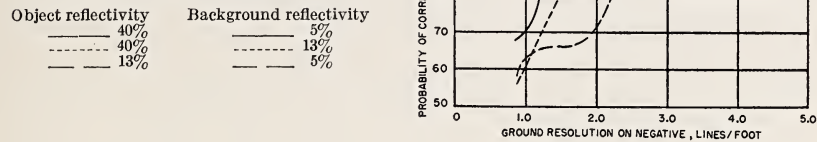
TABLE 5.1. *Contrast readings made at a resolution of 35 lines per millimeter on the negative and at a scale of 1:3700*

Reflectivities		Contrast
Object	Background	
%	%	
40	5	5.8
40	13	2.4
13	5	1.0

It is apparent from the experimental evidence at hand that count or detection is insensitive to resolution and, in fact, may be considered by treating only the energy distribution in the image and noting if sufficient contrast occurs to bring the object above the visual contrast threshold. Expressed in terms of angular resolution α, the number of objects per resolved spacing, the detection threshold on Super-XX emulsion over the range explored appears to follow the form

$$C - a\alpha - b = 0, \tag{8}$$

FIGURE 5.15. *Object differentiation as a function of resolution.*

Object reflectivity	Background reflectivity
———— 40%	———— 5%
- - - - - - 40%	- - - - - - 13%
— — 13%	— — 5%

where C is the image contrast and a and b are constants. a is about 0.09 and b is the contrast threshold that provides the same probability of detection at peak resolution of the photographic material (that is, same graininess conditions) for the same size symbol viewed under the same conditions.

Preliminary data from the recognition experiment are shown in table 5.2.

TABLE 5.2. *Resolution and contrast on the negative for various probabilities of recognition of isolated unit cubes*

Scale	Emulsion					
	Super-XX			Pan-X		
	Resolution		Contrast	Resolution		Contrast
	Lines/ object	Objects/ line		Lines/ object	Objects/ line	
50% probability						
1:2500	2.22	0.45	0.6	1.69	0.59	0.5
	1.67	.60	2.7	1.35	.74	1.5
	1.53	.65	4.5	1.27	.79	3.0
1:3700	2.56	.39	2.6	----	----	----
	1.81	.55	3.0	----	----	----
1:5000	2.00	.50	1.2	5.55	.18	0.6
	1.72	.58	2.0	2.85	.35	1.8
	----	----	----	2.27	.44	2.3
80% probability						
1:2500	3.12	0.32	0.74	2.22	0.45	0.5
	1.78	.56	2.8	1.81	.55	1.6
	1.56	.64	5.0	1.58	.63	3.1
1:3700	2.22	.45	3.6	----	----	----
1:5000	2.85	.35	1.6	3.70	.27	1.8
	2.50	.40	2.5	3.33	.30	2.7
95% probability						
1:2500	5.26	0.19	0.75	3.33	0.30	0.6
	2.38	.42	2.8	2.38	.42	1.6
	2.00	.50	4.8	2.22	.45	3.1
1:3700	3.33	.30	4.8	----	----	----
100% probability						
1:2500	7.14	0.14	0.8	5.00	0.20	0.75
	5.00	.20	3.2	2.85	.35	1.6
	4.00	.25	6.0	2.50	.40	3.1

A portion of these data is plotted in figure 5.16. These show image contrast as a function of resolution. Probability of recognition is a parameter. It is clear that the threshold for this form of recognition is sensitive to resolution but, broadly speaking, insensitive to contrast.

It is perhaps of interest to note in conjunction with these results the mechanism of the visual process. Peripheral vision provides for the detection of objects due to contrast (if we limit ourselves to a static gray scale, as in the case of photography). If the stimulus is sufficient, a reflex feed-back mechanism brings the object onto the fovea. Thus, the presence of the object is detected at low resolution and recognized (or studied with the intent of recognition) in the high-resolution portion of the eye.

It is possible to compare, by means of figure 5.16, two different emulsions that present the same type of data under the same conditions of resolution or contrast, the only difference being grain. In that class of photography where the purpose of the picture is to reveal information, it appears that consideration should be given to a graininess factor defined in terms of the role of the grain obscuring or obliterating information. In figure 5.16 the effect of graininess is manifest by the shifting of the curves in the direction of better resolution or increased contrast, for any given probability of recognition, as one goes from Pan-X to Super-XX emulsion.

Neglecting the psychological components (experience, recall, etc.) we may regard recognition as a higher-order detection. It is by means of detection of the presence of substructure within a symbol, or by detection of characteristic symbol groupings, or by detection of markings that reveal texture that recognition is achieved.

In the experiment described, the detection criteria measure the ability to detect the presence of, that is, to count, symbols. The result is not an absolute threshold for, as in actual interpretation, a search factor is involved. In this same experiment the recognition criteria correspond to detection of substructure. This is related to the decision as to whether a continuity or discontinuity in edge contour occurs. To render this decision the interpreter requires sufficient definition to observe that a finite portion of the object is bounded either by an arc or a discontinuity and/or a straight edge. Most probably the decision is based upon judgments of curvature.

From the particular results shown in figure 5.16, it is seen that when the unit square is imaged at 0.12 millimeter on a side on Super-XX emulsion the significant dimension over which a density difference must be detected appears to be about 15 microns at low contrast and 8 microns at higher contrast. Thus, at these higher contrasts the decisions are rendered by detection of brightness fluctuations occurring over spatial dimensions of the same order of magnitude as those of the individual grain. We may infer that the manner in which the resolution must improve as the image contrast decreases depends upon the symbol shape, and, more important, in the general case, upon the ratio of this significant symbol size to the grain size. From the experiment it is clear that in passing to better resolutions image contrast alone determines whether or not the sub-symbol is detected and, therefore, whether or not the gross recognition is achieved.

For simplicity, let us assume that the photographic interpreter makes his decisions or gains his information through scanning the photograph from blur point to blur point in a manner analogous to an
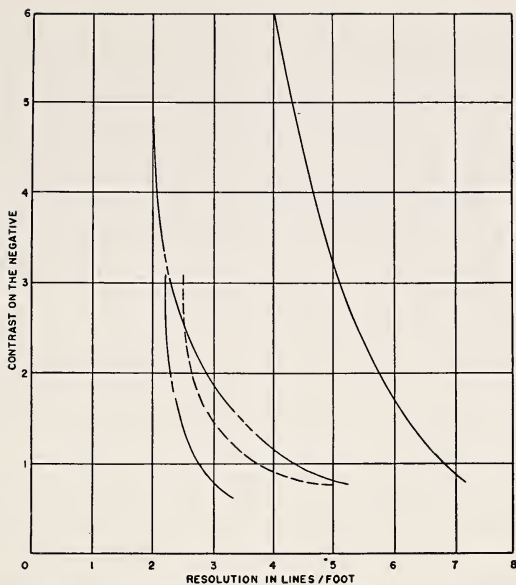
Scale 1:2500. 100-percent-probability: Super-XX, ———; Pan-X, — — — —. 95-percent probability; Super-XX, — - - —; Pan-X, — - —.
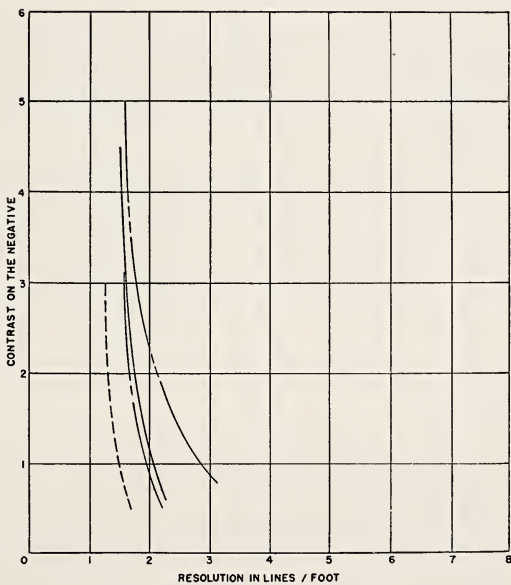
Scale 1:2500. 50-percent probability: Super-XX, ———; Pan-X, — — — —. 80-percent probability; Super-XX, — - - —; Pan-X, — - —.

FIGURE 5.16. *Contrast on the negative as a function of resolution to provide various probabilities of recognition.*

electron beam sweeping the mosaic of an orthicon tube. In this procedure the interpreter faces a basic decision as he proceeds from one blur point to the adjacent blur point: either "yes, the density does differ" or "no, the density does not differ." If there is no detectable change, the interpreter then concludes that there is no change from point to point in the object space as represented by the two image points. On the other hand, if a change does occur, then the interpreter must conclude that there is a change, a boundary, an edge, or a point of a brightness extreme within the corresponding area in the object space. According to the accepted view of the visual process the magnitude of this change is not important, nor are the levels at either side of the boundary. To quote Zoethout [17]:

> When a certain area of the retina is illuminated we are very little concerned with the absolute intensity of the light falling upon this area or the absolute intensity of the resultant sensation; but the ability to discern the difference between the light in this area and that in the neighboring portions of the retina is of prime importance. Upon *visual discrimination acuity* depends practically all our seeing.

Although the magnitude of the change in brightness over an edge is not important in the basic judgments that enter into photointerpretation, the brightness range over which a system can differentiate between small brightness differences enters into any measure of the information capacity of the system. Thus, although figure 5.3 (threshold contrast as a function of image size) shows the relative area of a one-dimensional detail universe that occurs above the system threshold, the one-coordinate case, to be completely represented, must be measured by a volume, the other coordinate being the median exposure level or its equivalent.

The role of the edge in photography is well recognized. Howlett [18] has long emphasized the importance of edges in the photographic problem and has designed his resolution targets with this in mind. More recently, Higgins and Jones [19] have treated subjective picture sharpness and in their work have correlated sharpness judgments with a gradient function.

The conditions of maximum sharpness and minimum flare (maximum symbol contrast) are apparently coincident. However, it is seen in figure 5.11 that, for those lenses examined, no one focal setting provides maximum image contrast for all symbol sizes. This implies that to achieve maximum information the focal setting must be chosen with consideration as to the size of the significant detail in the image. As a consequence, it turns out that the focal setting that provides maximum information in photographs that are studied under unit magnification does not necessarily provide maximum information under any other power.

Therefore, in such cases as aerial photography, when the picture is subject to study under several magnifications in an effort to extract the maximum information, it becomes necessary to consider the distribution of symbol sizes to be recorded. From this distribution it should be possible to determine the focal setting that provides the optimum (weighted) contrast threshold over the format.

Summary

A photographic system is called upon to record symbols of different shapes and sizes and with different spacings and arrangements. These symbols also occur at different contrasts.

This paper considers only a one-dimensional aspect, that is, size not shape. By viewing isolated symbols as well as symbols occurring at regular spatial frequencies two extremes of spacing or arrangement are considered. Because the present photographic emulsion dominates the contrast-reduction characteristic, the threshold curves will be of the same general shape for all present aerial photographic systems.

The statistics of certain photographic messages were analyzed; photographs were classified by the using agencies as not quite good enough, just good enough, or excellent, for a particular purpose. These photographs were taken of natural objects to lend validity to the assumption of similar distributions of distances between edges and inherent contrasts in the object space. The prints were analyzed and similar normalized distributions in the x-coordinate were observed on all pictures. On the other hand, the distributions of brightness differences across edges were different and related to usability of the photography. The results indicate that the greater the contrast reduction by the system the poorer was the picture as classified by the using agency. The presence of symbols on aerial prints, which are finer than the air-borne resolution generally recorded even with high-contrast, line-pattern targets, indicates that the nature of the detail universe is not comparable to that of our target structure. From these analyses there is an indication that the statistics of the recorded messages are quite similar, although the target and/or purpose may differ. This consideration also involves cursory examination of photographs for urban analysis.

The contrast-reduction function of the lens-emulsion system has been interpreted in terms of the object contrast that is required to render a detectable symbol on the emulsion for the two spatial arrangements considered. It is held that the shape of this threshold curve is important in determining the ability of the system to record symbols. Until a statistical weighting can be applied to the size, shape, and spatial distribution in the object space, probably no more meaningful expression of the system capabilities can be given than through an interpretation of eq 2, which considers a weighted mean performance over the format

$$P = \frac{\Delta - \int_{f_\theta}^{x_L} \psi(x)\,dx}{\Delta}.$$

This expression may be considered as giving the probability that any symbol in a random detail universe will be located above the system threshold. It becomes then a measure of the capacity of the system.

When symbols are in close proximity in the object space (approximated by the data for symbols occurring at regular spatial frequencies) it is improbable that these symbols possess sufficient inherent contrast to be recorded at the high-contrast resolution limit of this system. This fact, coupled with the observation that peak-contrast rendition in the image does not occur at the same focal setting for all image

sizes, implies that judgment of relative performance of aerial photographic systems on the basis of high-contrast resolution scores in the laboratory is not infallible. It further emphasizes that the best operational focus cannot be determined in the laboratory from a maximum resolution setting on a high-contrast target.

The experimental work on the psycho-physical detection-recognition criteria on the emulsion, as a function of resolution, scale, and contrast indicates that the fundamental problem is providing sufficient contrast in the image to permit its detection. For a given shape and size of symbol a certain minimum level of sharpness is requisite. At any point above this level of sharpness, however, the criterion is again one of contrast.

In figure 5.16 it is seen that at a given image contrast the same level of recognition may be achieved on Pan-X emulsion at a lower resolution than on Super-XX. This fact is introduced to point out the significant role that graininess plays in obscuring photographic information and the resulting need for directing more attention to this important factor. With the emphasis directed on achieving a large resolution number we may have overlooked this factor and others of equal significance that have major bearing on the information that the picture can reveal.

[1] W. R. Dawes, Mem. Roy. Astron. Soc. London **35**, 158 (1865–66).

[2] Lord Rayleigh, Phil. Mag. **8**, 266 (1879).

[3] O. H. Schade, Electro-optical characteristics of television systems (in four parts), RCA Rev. **9**, 5–37, 245–286, 490–530, 653–686 (1948).

[4] W. Zschokke, Phot. Korr. **36**, 131 (1899).

[5] A. Wetthauer, Z. Instrumentenk **41**, 148 (1921).

[6] D. E. Macdonald, Calibration of survey cameras and lens testing, Photogrammetric Eng. **17**, 383–389 (1951).

[7] H. R. Blackwell, Contrast thresholds of the human eye, J. Opt. Soc. Am. **36**, 624 (1946).

[8] S. Q. Duntley, The reduction of apparent contrast by the atmosphere, J. Opt. Soc. Am. **38**, 179 (1948).

[9] P. D. Carman and R. A. F. Carruthers, Brightnesses of fine detail in air photography, J. Opt. Soc. Am. **41**, 305 (1951).

[10] D. E. Macdonald, A preliminary consideration of air turbulence effects on definition in aerial photography (Boston Univ. Opt. Research Lab. Tech. Note 54, June 21, 1949).

[11] D. E. Macdonald, Calibration of survey cameras and lens testing (Boston Univ. Opt. Research Lab. Tech. Note 54, June 21, 1949).

[12] J. G. Baker and J. S. Chandler, Equipment for aerial photography, chap. 1, Summary Tech. Rep. NDRC, Div. 16.1 (Optical Instruments) 1946.

[13] A. H. Katz, Camera shutters, J. Opt. Soc. Am. **39**, 1–21 (1949).

[14] D. E. Macdonald, Progress report on aerial camera motions (OSRD Rept. 5178, June 1945). Contractor: Massachusetts Institute of Technology (OEMsr–203).

[15] D. E. Macdonald, Criteria for detection and recognition of photographic detail, Part I—Resolution, scale and contrast conditions for isolated detail, Part II—System performance (Boston Univ. Opt. Research Lab. Tech. Notes 69 and 72, July 10, 1950 and October 15, 1950). Also, paper presented at the Winter Meeting of the Optical Society of America, March 9–11, 1950, Criteria for detection and recognition of photographic detail (Abstract) in J. Opt. Soc. Am. **40**, 258 (1950).

[16] S. Q. Duntley, The visibility of distant objects, J. Opt. Soc. Am. **38**, 237 (1948).

[17] W. D. Zoethout, Physiological optics, 4th ed., p. 167 (The Professional Press, Inc., Chicago, 1947).

[18] L. E. Howlett, Photographic resolving power, Can. J. Research **24**, 15 (1946).

[19] G. C. Higgins and L. A. Jones, The nature and evaluation of the sharpness of photographic images, J. Soc. Motion Picture Engrs. **58**, 277–290 (1952).

6. A Mathematical Model of an Optical System [1]

By Max Herzberger [2]

The new methods of rapid computation (IBM equipment, electronic devices) give the possibility of investigating systematically the validity and practicability of optical methods. The author has started an ambitious enterprise. He is tracing a set of more than a thousand rays through each of three optical systems: (1) An aerial lens with an aperture of $f/5.6$ and a half-field angle of 22 degrees; (2) a wide-angle lens with an aperture of $f/6.3$ and a half-field angle of $37\frac{1}{2}$ degrees; (3) a lens with an aperture of $f/7$ and a half-field angle of 9 degrees.

The first problem is to see whether it is possible to find a mathematical model for the lens in question, i. e., to find a mathematical formula with not too many constants with which object and image rays can be coordinated analytically with an accuracy of a few units of the fifth decimal (for focal length equal to one). The second problem is to find out how many of the thousand rays are necessary and sufficient to derive a function with this accuracy. Figure 6.1 shows that the desired accuracy was obtained.

This mathematical model is used to analyze the optical images. The intersection points of a set of rays equally distributed over the entrance (or exit) pupil, (a) with a set of image planes, (b) with the plane through object point and system axis (the meridian plane), are plotted (fig. 6.2, 6.3, and 6.4).

The plot of these last points, which are called the *diapoints* of the object, lends itself particularly well to a graphical analysis of the system and permits one to investigate the aberrations of each single ray.

In this paper, a preliminary study is made of an aerial lens having an aperture of $f/5.6$. The agreement between the model, the exact calculation, and the photomicrograph of the images is shown, and an analysis of the diapoint diagram is given.

Once a mathematical model of an optical system is given, it can be investigated. In a future paper an attempt will be made to follow up recent ideas in order to compute the influence of diffraction, to analyze the image errors as function of the aberrations at the single surfaces, and to study different approaches to determine the quality of the optical image for the three lenses mentioned in the introduction.

The analytical approach taken in the present paper rests on the mixed characteristic function of Hamilton and Bruns.

Let the object and image origins be put at the entrance and exit pupils of the optical system (in this case, the nodal points). Let the optical axis coincide with the $z(z')$ axis of our coordinate system.
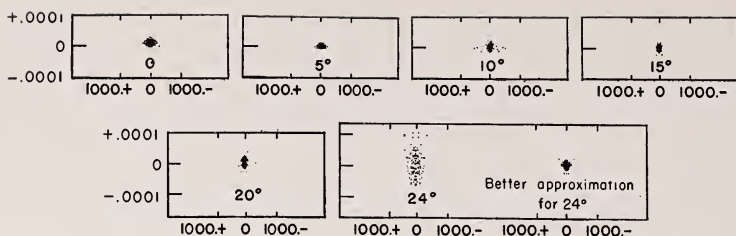
FIGURE 6.1. *Deviation of the data obtained by ray tracing from the mathematical model of the optical system.*

Let object and image rays be given in each case by the intersection point of the ray with the plane $z,(z')=0$, i. e., by x, y, (x', y'), respectively. The optical direction cosines may then be designated by ξ', η', ζ' and they will be normed in such a way, that

$$\xi^2+\eta^2+\zeta^2=n^2, \qquad \xi'^2+\eta'^2+\zeta'^2=n'^2, \tag{1}$$

n, n' being the refractive indices of object (image) space.

Since the optical system has symmetry with respect to the system axis, the characteristic function, V, is a function of the symmetric functions of the coordinates, i. e., of

$$u=\tfrac{1}{2}(\xi^2+\eta^2), \qquad v=\xi x'+\eta y', \qquad w=\tfrac{1}{2}(x'^2+y'^2), \tag{2}$$

and for a system of parallel rays with an inclination of 0, 5, 10, 15, 20, 24 degrees from the axis, respectively. V is a function of v and w, and the direction cosines of the image ray are

$$\xi'=\frac{\partial V}{\partial v}\,\xi+\frac{\partial V}{\partial w}\,x'_N, \qquad \eta'=\frac{\partial V}{\partial v}\,\eta+\frac{\partial V}{\partial w}\,y'_N. \tag{3}$$

From (3) the coordinates of the diapoint can be computed, i. e., the intersection point of the rays with the meridian plane, as

$$x'_D=(V_2/V_3)\xi, \qquad y'_D=(V_2/V_3)\eta, \qquad z'_D=-1/V_3, \tag{4}$$

using the abbreviation,

$$\frac{\partial V}{\partial v}=V_2, \qquad \frac{\partial V}{\partial w}=V_3.$$

Let V be given by a fifth-order development, i. e.,

$$V=a_0+a_2v+a_3w+\tfrac{1}{2}(a_{22}v^2+2a_{23}vw+a_{33}w^2) \\ +\tfrac{1}{6}(a_{222}v^3+3a_{223}v^2w+3a_{233}vw^2+a_{333}w^3), \tag{5}$$

and therefore,

$$V_2=a_2+(a_{22}v+a_{23}w)+\tfrac{1}{2}(a_{222}v^2+2a_{223}vw+a_{233}w^2),$$
$$V_3=a_3+(a_{23}v+a_{33}w)+\tfrac{1}{2}(a_{223}v^2+2a_{233}vw+a_{333}w^2).$$

74

Inserting these into eq 3, the coefficients a_i, a_{ik}, $a_{ik\lambda}$ can be calculated from a small number of rays traced through the system by the method of least squares. (At least five meridian rays and five skew rays should be traced.)

Figure 6.1 gives the fit of the mathematical model. The system had a focal length of 3, and the deviation for ξ', η', ζ' calculated by (5) from the ray-trace data for the rays up to a 20° field angle is, in general, smaller than 1×10^{-5}. Only for the field angle of 24° did the deviation amount to a few units in the fifth decimal. Introducing a few more coefficients reduced the errors to insignificant size.

The coefficients of the functions are given by table 6.1.

TABLE 6. 1.

	0°	5°	10°	15°	20°	24°
a_2	----------	+1.000005	+0.999999	+1.000008	+1.000216	+1.000776
a_3	+0.326322	+0.325219	+.321972	+.316264	+.307558	+.297747
a_{22}	----------	−.351799	−.346312	−.335021	−.313910	−.278191
a_{23}	----------	−.087948	−.088330	−.091106	−.094766	−.094814
a_{33}	+0.089396	+.091207	+.087397	+.067329	+.050488	+.030552
a_{222}	----------	−.06390	−.07172	+.02139	−.00011	−.03586
a_{223}	----------	−.00680	−.03740	−.06315	−.06494	−.28434
a_{233}	----------	−.39280	−.45960	−.49228	−.58776	−1.09514
a_{333}	−4.15539	−4.26160	−4.44175	−3.88328	−4.41015	−4.72242

Inspection of table 6.2 shows that the coefficients are slowly changing functions of $u = 1/2 \sin^2 \sigma$. They can be replaced by

$$a_i = a_i^{(0)} + a_i^{(1)} u + 1/2\, a_i^{(2)} u^2 + 1/6\, a_i^{(3)} u^3, \tag{6}$$

thus obtaining V as a function of u, v, w.

When V is obtained, the desired intersection points are calculated for a large system of rays, evenly distributed over the exit pupil (the points were ordered in equilateral triangles evenly distributed over the vignetted exit pupil). Thus, the intersection can be plotted (a) with a set of planes perpendicular to the axis (in this case, three planes were chosen, one through the Gaussian focal plane, and two 0.75 and 1.5 mm, respectively, in front of it); and (b) with the meridian plane (calculation of the diapoints).

Figure 6.2 shows these spot diagrams. On the right will be seen the size of the vignetted apertures, this giving a measure of the light going through the system for different points of the field. The lines $y_N = $ constant are drawn on these apertures, and they contain the points used for the spot diagrams. The symmetry of the problem with respect to the meridian plane makes it only necessary to scan the rays of one-half of the aperture.

TABLE 6.2.

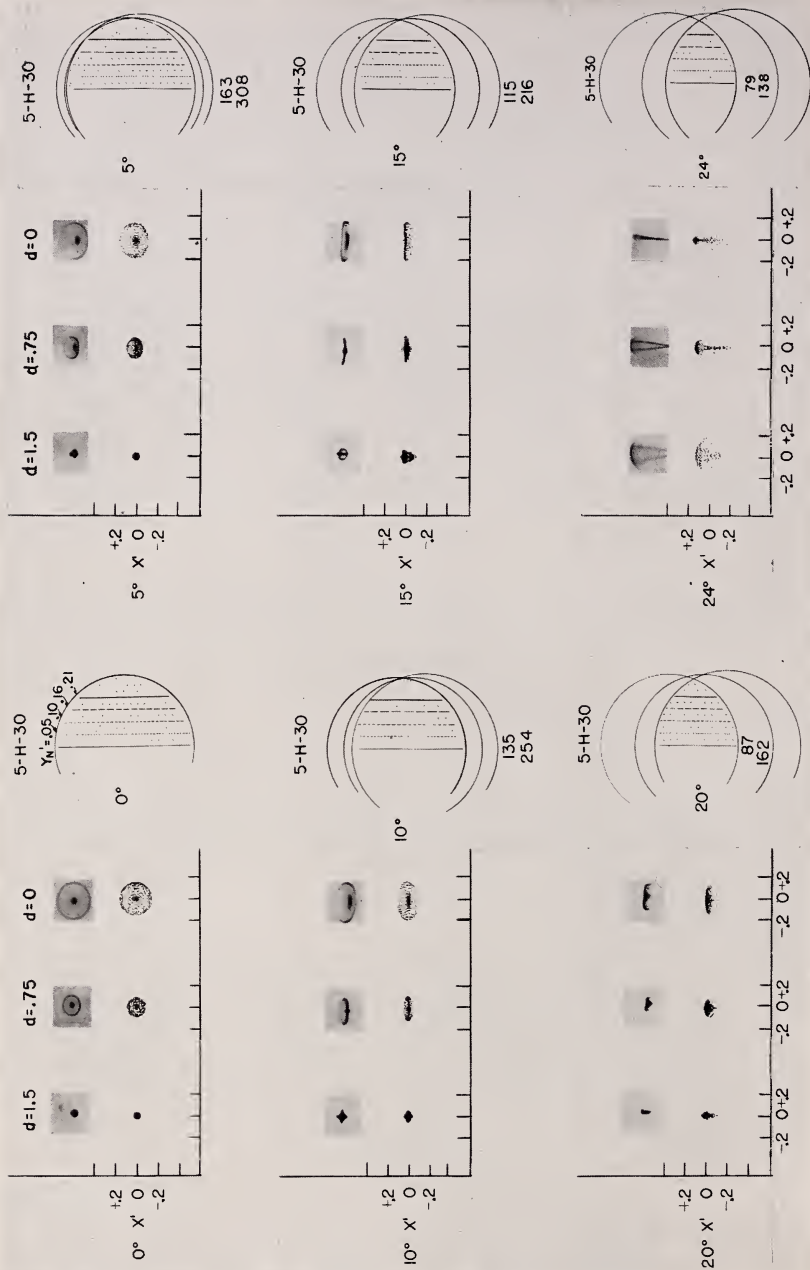|	24°		24°		24°
a_2	+1.000731	a_{222}	+.17443	a_{2223}	−11.1725
a_3	+0.0297783	a_{223}	−.11766	a_{2233}	−4.4855
a_{22}	−.283235	a_{233}	−.58800	a_{2333}	−24.0728
a_{23}	−.096540	a_{333}	+2.62365	a_{3333}	−501.5666
a_{33}	−.009835	a_{2222}	+4.7495		

FIGURE 6.2. *Comparison of computed images with microphotographs.*

76

Figure 6.2 shows the quality of the optical image in detail. Obviously, the best image plane is not the plane through the Gaussian focus, but the plane about 1.5 mm in front of it. The field is very slightly curved forward and comes back at about 20 degrees and then curves rapidly backward. The image at 20 degrees looks slightly more compact than at 15 degrees, and at 24 degrees there is no longer any image to speak of.

Also, Figure 6.2 contains reproductions of photomicrographs taken by L. A. Jones and R. N. Wolfe[3] of these laboratories printed side by side with the spot diagram, showing the agreement between computation and photographic image.

The difficulty of analyzing the spot diagrams lies in the fact that near the plane of best focus there is a heaping of singularities of the wave surface. This is illustrated in Figure 6.3. The image is analyzed by computing the intersection points of the lines, $y_N'' = \text{constant}$, going through a set of parallel lines in the exit pupil.
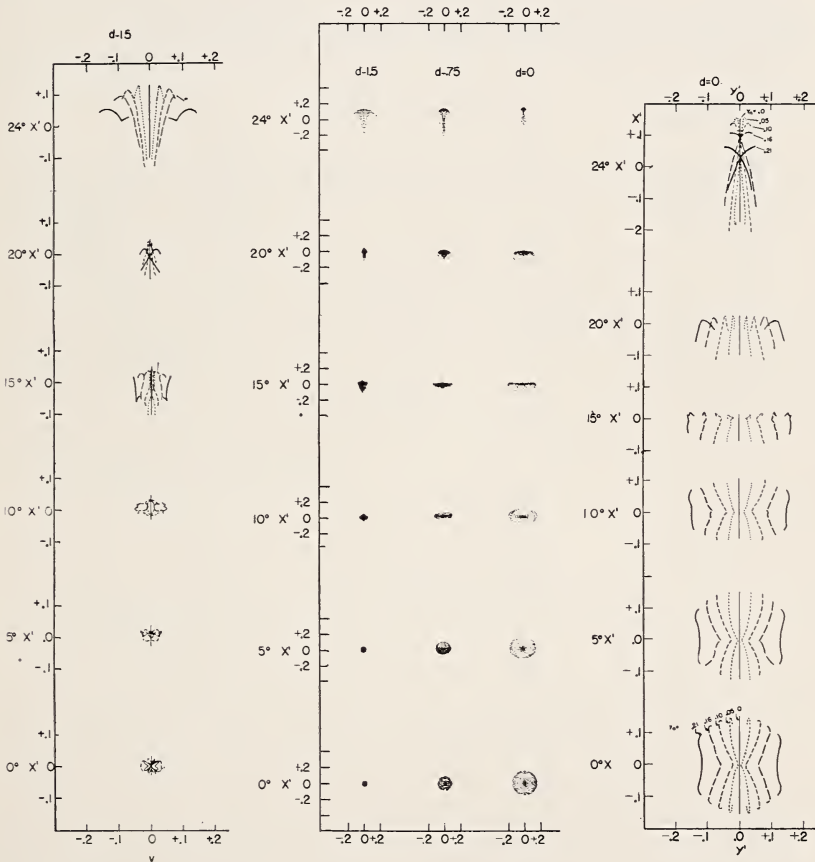
FIGURE 6.3.   Analysis of the optical images obtained by scanning the vignetted aperture.

[3] L. A. Jones and R. N. Wolfe, J. Opt. Soc. Am. **35**, 559 (1945).

The figures are drawn for the Gaussian focal plane and the best image plane 1.5 mm in front of it, and they are four times magnified with respect to the spot diagrams, thus showing where the condensation of light occurs. Inspection of these figures shows that these curves are highly irregular and multivalent, which makes it very obvious why it is so difficult to compute the intersection points in the image plane as simple functions of the direction of the ray and the intersection height in the aperture of the plane.

The author has, for a long time, drawn to the attention of the optical designer the importance of computing the diapoints, i. e., the intersection of the rays (from a given object point) with the meridian plane, i. e., the plane through the object point and the axis (for a meridional ray, the diapoint coincides with the sagittal focus). It can be proven that the knowledge of the diapoints gives complete information about the optical image and, moreover, that the diapoints allow a simple analysis of the image qualities. If the diapoints form a straight line, the image rays form a symmetric image with the line in question as the axis symmetry. The deviation of the diapoints from the best straight line gives a measure of asymmetry.

Figure 6.4 is a plot of the spot diagrams and, on a large scale to
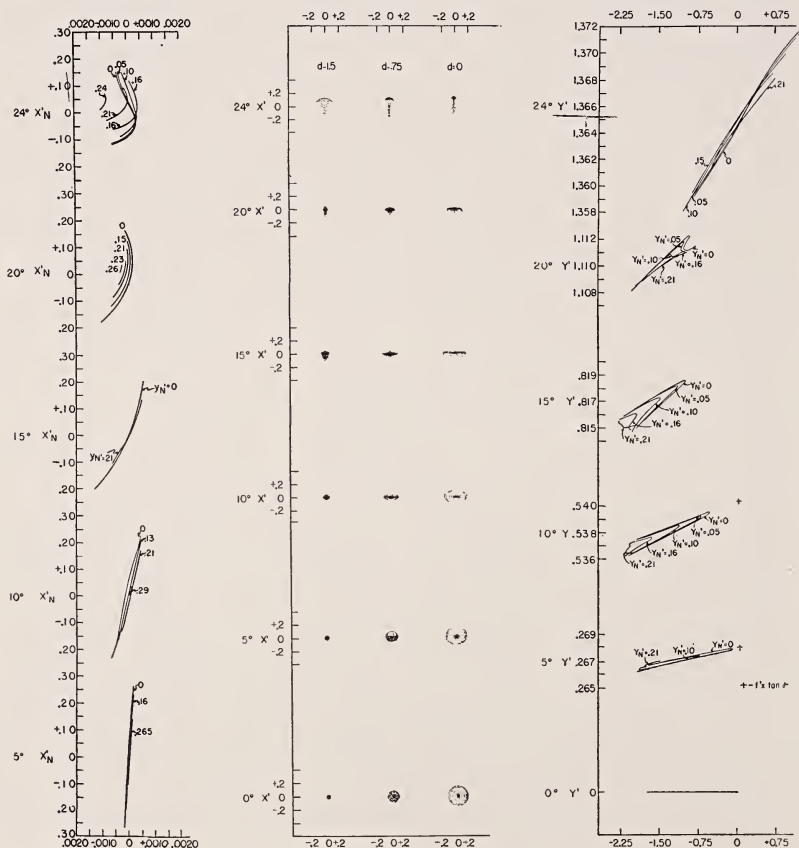
FIGURE 6.4. *Diapoint analysis of an optical system.*

the right, the diapoints for each angle, the lines hereby corresponding to lines y'_N =constant. On the left is plotted for each of these lines the deviation from symmetry. The attention of the reader is called to the fact that these lines are simple curves that can easily be analyzed mathematically.

Even a superficial inspection of these curves shows that the spherical aberration remains practically constant up to 20 degrees, that a slight forward curvature exists which is corrected at about 20 degrees, and not much of an image remains at 24 degrees. These curves show the effect of vignetting, and also that the image at 20 degrees is slightly better, though slightly less symmetric, than at 15 degrees off-axis.

The mathematics of the problem in question is best attacked when the angle characteristic, W, is employed and used as a function of the variables,

$$a = \zeta, \qquad b = \xi\xi' + \eta\eta', \qquad c = \zeta'. \qquad (7)$$

The diapoint coordinates then become, for an infinite object point (a = constant),

$$x'_D = \frac{\partial W}{\partial b}, \qquad z'_D = \frac{\partial W}{\partial c}, \qquad (8)$$

thus determining W completely. Moreover, distance K of the points of the caustic from the diapoints can be computed by the expressions,

$$n'^2 K^2 - [\{ n'^2(n^2 - a^2) - b^2 \} W_{22} - 2bc W_{23} + (n'^2 - c^2) W_{33}]$$

$$+ [(n^2 - a^2)(n'^2 - c^2) - b^2](W_{22} W_{33} - W_{23}^2) = 0 \quad (9)$$

with

$$W_{22} = \frac{\partial^2 W}{\partial b^2}, \quad \cdots \,,$$

and the values of the function

$$W_{22} W_{33} - W_{23}^2 \qquad (10)$$

give a measure of the asymmetry of the image.

It will be shown in a future paper how these diapoint aberrations can be split up into the contribution of the single surfaces, according to the methods already described.[4, 5]

Diffraction effects also can easily be computed from a knowledge of the characteristic function, V, since V gives, for each point of the exit pupil, the light path for a plane-entering wave. By using Fresnel's integrals, the amplitude and phase of the resulting light vector can then be computed at any point in image space.

Furthermore, the geometrical optical data will be coordinated with the recent attempts to find an analytic measure for the sharpness of an optical image.

[4] M. Herzberger, J. Opt. Soc. Am. **38**, 324–328 (1948).
[5] M. Herzberger, J. Opt. Soc. Am. **37**, 485–493 (1947).

7. Methods and Apparatus for Measuring Performance and Quality of Optical Instruments

By A Arnulf [1]

Introduction

From the user's point of view, and for a given point in the field, the value of an optical instrument is completely described by two figures. The first is its performance (Leistung), which, after Löhle, is the number showing how much the instrument multiplies visual sharpness. In France this number is called the "amplification" of the instrument. The second figure is the quality, which, in its simplest form, is the ratio of the performance of the instrument under test to the performance of a perfect instrument with similar characteristics.

Visual Instruments

Theory

Let us first consider the case of visual instruments in which the aerial image is observed with an eye-piece. The work on resolving power relative to perfect instruments, which I reported a few years ago, allows us, first, to find an absolute criterion of the quality of the instrument, and secondly, to find a simple connection between its quality and its performance. [2]

We define the perfect instrument as an instrument in which the distribution of light in the image is determined by diffraction alone. For such an instrument, the resolving power is given by the equations

$$S \cdot \Omega = 2n \sin U \cdot T = s_\omega \cdot \omega, \tag{1}$$

where S and T are the angular and linear resolving powers of the instrument, Ω and $n \sin U$ are the linear aperture and the numerical aperture of the objective, and s_ω is the resolving power of the eye for the diameter ω of the pupil.

These equations may be written

$$S \cdot G = T \cdot P = s_\omega, \tag{2}$$

where G is the angular magnification, and P the power of the instrument.

[1] Institut d'Optique, Paris, France.
[2] A. Arnulf, Compt. rend, ac. sc. **200**, 52 and 306 (1937); La Vision dans les Instruments, Edit. Rev. opt. (1937).

Let us consider an imperfect instrument with the same exit pupil and the same magnification as the perfect instrument. S', G', T', P' are corresponding terms for the defective instrument. We know that $S' > S$. Therefore,

$$S' \cdot G > S \cdot G \quad \text{or} \quad s'_\omega > s_\omega. \tag{3}$$

Furthermore,

$$\frac{S}{S'} = \frac{s_\omega}{s'_\omega}. \tag{4}$$

Therefore, if we compare a defective instrument with a perfect one with identical optical characteristics, the ratio of the resolving powers, or of the performances, is given by the ratio of the resolving powers of the eye in the image field. The quality of the image is then defined by

$$E_i = \frac{s_\omega}{s'_\omega}, \tag{5}$$

which is called *instrumental efficiency* because it depends only upon the effect caused by the defects of the image.

Let us compare the quality of two different instruments, for example, the big telescope in Yerkes Observatory and a low-magnification microscope. The above method will give for each one its instrumental efficiency and the ratio of these two efficiencies will give the relative instrumental efficiency of the two instruments.

We may sum up the principle of this method as follows. The resolving power of the eye in the image field of a perfect instrument is the same as the resolving power of the eye alone with the same diameter of the pupil. The defects of the instrumental image degrade this resolving power, and this degradation determines the quality.

Total Efficiency

The criterion defined above gives the loss of quality compared with a perfect instrument of the same type, and it is practically the most important fact. However, where the user is to be considered, there is a loss of the effective quality caused by the stopping of the pupil, which increases the resolving power of the eye. Therefore, it is necessary to consider this stopping effect in order to compare the quality and the performance. Similar considerations to those previously used show that when we include both the stopping of the eye and the defects of the image, the quality is given by

$$E_t = \frac{s_n}{s'_\omega} = \frac{s_n}{s_\omega} \cdot \frac{s_\omega}{s'_\omega} = E_p \cdot E_i, \tag{6}$$

S_n being the lowest resolving power obtainable with the eye, corresponding to a pupil diameter of about 2.5 to 3 mm. This is, for all practical purposes, the resolving power of the eye in daylight. $E_p = s_n/s_\omega$ is the efficiency of the stopping process. This can be obtained by measuring visual sharpness.

Radius of pupil (mm)	0.2	0.4	0.75	1.0	2.0	4.0
Pupil efficiency	0.07	0.23	0.50	0.64	0.88	1.0

According to what has been said, one conclusion that may be reached is that an optically perfect instrument might be a visually moderate instrument. This is the case especially with instruments operating at the magnification giving the best resolving power.

Relation Between Efficiency and Performance

In the case of visual instruments, there is a precise relation between the performance and the efficiency. The performance is given by the amplification A, where $A \equiv s_n/S$. This means that

$$A = G \cdot E_t. \tag{7}$$

In this equation G is the angular magnification for telescopes, or the conventional magnification for microscopes or such instruments. For a perfect instrument, $E_i = 1$ and $A = G \cdot E_p$.

Experimental Procedure

Measuring the efficiency requires: (1) The determination of the resolving power of the tested instrument from which is deduced the resolving power of the eye in the image field by using the formula

$$s'_\omega = S\, \frac{\Omega}{\omega} = S \cdot G,$$

or

$$s'_\omega = T \cdot \frac{2n \sin U}{\omega} = T \cdot P;$$

(2) the resolving power of the eye without the instrument, for one or several apertures of the pupil, which gives the relation to the perfect instrument.

The quality of an instrument is a function of the experimental conditions. Therefore, it is necessary to choose conditions such that the result might be valid in as many cases as possible, and also to duplicate as nearly as possible the conditions under which the instrument is to be used. The conditions of testing once defined will be taken as a constant standard.

1. The test object is seen against a background of uniform brightness, which covers the entire field of the instrument. This is necessary if one is to introduce all the stray light that the instrument might give under use. The effect eventually produced by the sun or moon will be realized by sources situated in precise points in the field.

2. No optical system will be interposed between the instrument and the test object in order to avoid any effect caused by its defects and stray light it might introduce.

3. No telescope is placed behind the eye piece to increase the magnification for the same reasons given in item 2, and also because the increased efficiency makes the test unnecessarily strict and does not duplicate the conditions of use. It is not reasonable to expect the instrument to have a better quality than that given by a perfect instrument with the same magnification.

4. The standard resolution test object is the Foucault test object, or a test object consisting of two parallel lines of infinite length and of varying contrasts. For simplicity one can use a contrast of 1 and a very low contrast, 0.01 or 0.02. Tests with the higher contrast are more sensitive to astigmatism or focusing defect, whereas with the lower contrast they are more sensitive to stray light, spherical aberration, and chromatism. It is desirable to use several brightnesses representing daylight, twilight and night. Night instruments tested under daylight conditions might give results completely different from the desired results.

Description of the Apparatus

Figures 7.1 and 7.2 show the arrangement of the apparatus. The background with a uniform brightness is obtained by placing the nose of the instrument I in a lighted sphere S, and by using several plane screens E_1, E_2 which illuminate the field. The black test object is seen against the last screen.

The variation of the contrast is obtained by using a rotating white disk D, of the same brightness as the backround, and consisting of two similar sectors between which the aperture can be varied while the disk is rotating. The contrast is given by $C = \alpha/360$, (α in degrees), α being the total aperture of the disk.

Furthermore, the angular size of the test object T can be varied continuously by projecting its image on a fixed plane, the objective O and the test object T being moved simultaneously with suitable cams. The apparatus is simple to use with any form of test object (fig. 7.3 and 7.4).
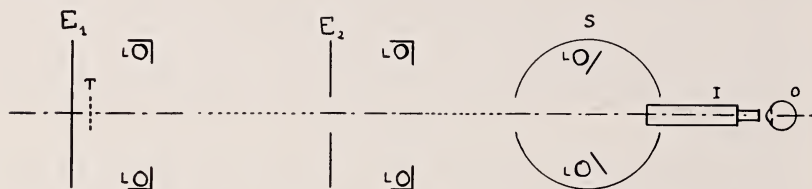

FIGURE 7.1. *Testing apparatus for visual instruments.*
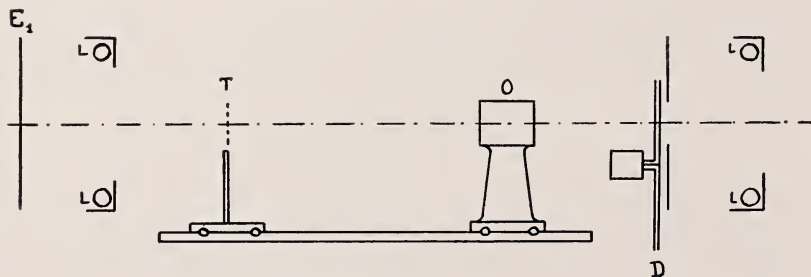(General arrangement.)

FIGURE 7.2. *Testing apparatus for visual instruments.*
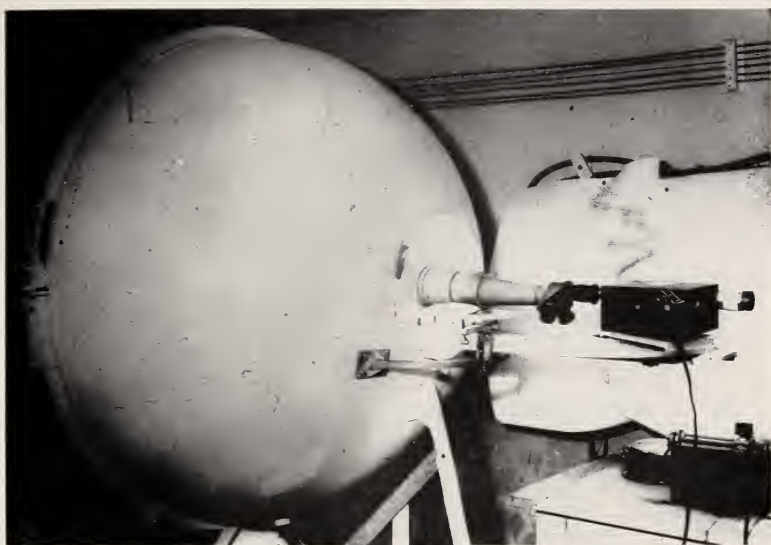Control of the contrast and the angular size of the test object.

84

FIGURE 7.3. *Testing apparatus for visual instruments.*
Instrument to be tested and sphere S. For efficiency measurements, the photometer behind the eye piece is removed.
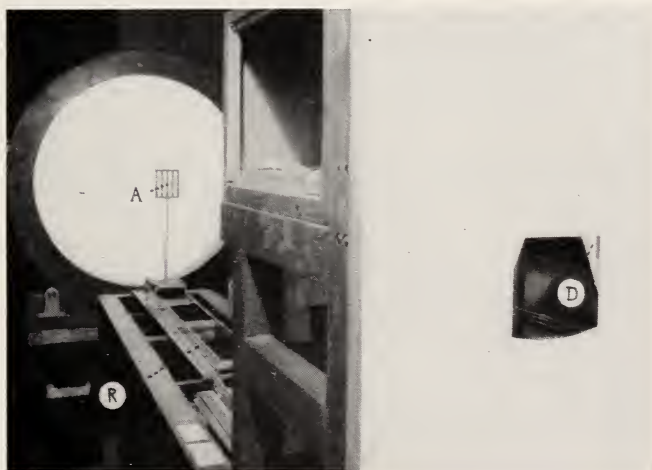

FIGURE 7.4. *Testing apparatus for visual instruments.*
Rotating disk D; test object A; optical bench R; round bright screen for the background.

Standard characteristics for the examination of visual instruments are as follows:

Brightness of the background.

$$5 \cdot 10^{-3}, \ 10^{-4}, \ 2 \cdot 10^{-6}, \ 5 \cdot 10^{-7} \ 10^{-8}, \ 5 \cdot 10^{-10} \text{ Stilb.}$$

The values most used are in italic.

Nature of the test object. Foucault resolution test object, or equivalent test object with two parallel lines.
Contrasts.

| 1.0 | 0.5 | 0.25 | 0.10 | 0.06 | 0.03 | 0.015 | 0.01 |

Generally we use two contrasts, contrast 1.0 and the contrast nearest the threshold for the brightness used.

Points tested in the field. The center, a circle equal to two-thirds of the radius of the field, a circle 2° from the edge of the field.

The instrument is always tested in the same conditions under which it is used, without any accessory optical aids.

Examples of the Effect of Isolated Defects

Decrease in brightness. Figure 7.5 shows the variation of the efficiency with the brightness for an optically perfect instrument with a transmission factor 0.35, using a Foucault test object of unit contrast. The efficiency is related to the slope of the curve of visual sharpness and brightness. It is equal to 1 for daylight and very nearly 1 for night vision (the slope is zero or very small). It is least in the brightness region where visual sharpness varies most rapidly with the brightness, that is, in twilight, and again near the threshold.

Stray light. Stray light produces a decrease in perception for low contrasts, which can be computed in every case if we have previously measured the stray light. Figure 7.6, curve 1, is the curve for an instrument the correction of which is nearly perfect. Table 7.1 gives the values of efficiency versus contrast for two binoculars in the center of the field. Lines A refers to an instrument without blooming, lines B to the same instrument with bad blooming which increases the scattering.

Monochromatic spherical aberration. This effect has been reported by M. Françon for both theoretical and experimental methods.[3] The
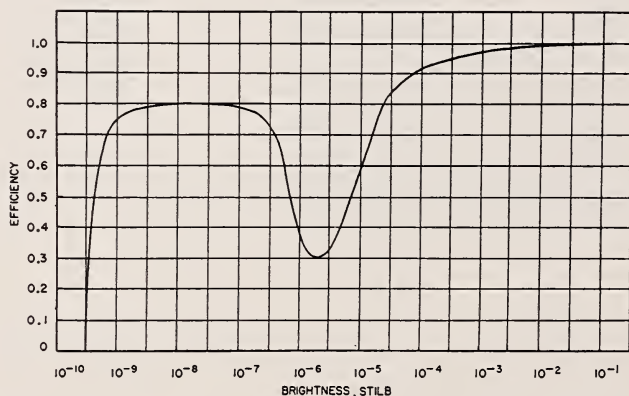

FIGURE 7.5. *Efficiency versus brightness of a perfect instrument having a transmission factor of 0.35.*

Test of resolution with full contrast.

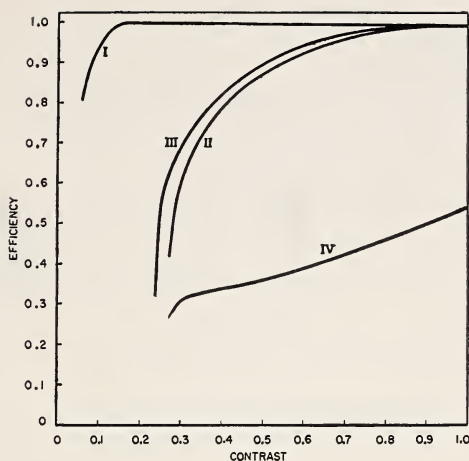[3] M. Françon, Rev. opt. [10] **26**, 354 (1947); [3] **27**, 157 (1948).

FIGURE 7.6.   *Efficiency versus contrast.*

TABLE 7.1.   *Efficiency versus contrast for two binoculars, before and after a diffusing blooming, showing the effect of stray light*

(Center of the field)   Lines A: Before blooming.   Lines B: After blooming.

Contrast	1	0. 6	0. 3	0. 1	0. 08	0. 05
1 { A	0. 73	0. 91	0. 97	0. 98	0. 97	0. 90
B	. 73	. 90	. 91	. 91	. 15	0
2 { A	. 85	. 85	. 95	. 95	. 85	0
B	. 85	. 85	. 94	. 60	. 25	0

curve (fig. 7.7) gives, as a function of the third-order spherical aberration given in phase differences, the values of efficiency for various contrasts in the test object, when the exit pupil of the instrument corresponds to the best resolving power (magnification giving $\omega = 0.6$ mm).   We can see that for small contrasts and the Rayleigh tolerance ($\lambda/4$) the fall in efficiency is considerable.   Figure 7.8 shows experimental results obtained with a contrast of 0.03, varying the diameter of the pupil of the eye.   An aberration corresponding to the Rayleigh tolerance has no effect on pupils larger than 3 mm.   For smaller pupils the efficiency decreases quickly.   Figure 7.6 III and IV shows the results for a very large spherical abberration.

Astigmatism.   The results in table 7.1 are those of the very common case of an instrument with a little astigmatism.   The efficiency is better for medium contrasts than for high contrasts.   Figures 7.9 and 7.10 show how the resolving power varies with the diameter of the pupil when astigmatism is present, and this explains the above results.

Chromatism.   This question is being studied both theoretically and experimentally.   Curve II figure 7.6, shows the variation of efficiency with contrast for a very large chromatism (aperture f/12, no chromatic correction).
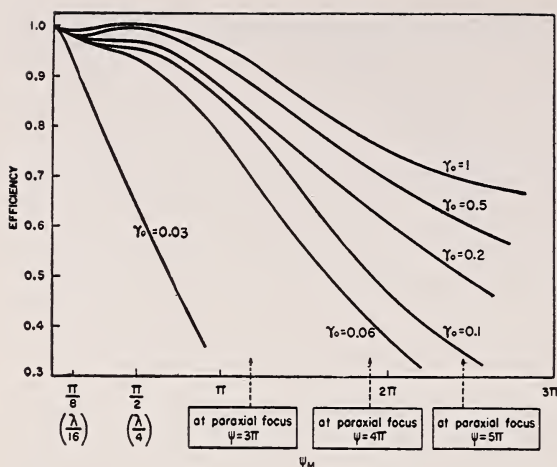
FIGURE 7.7. *Efficiency versus spherical aberration expressed in phase differences for various contrasts.*

FIGURE 7.8. *Efficiency versus pupillar diameters for various amounts of spherical aberration expressed in phase differences.*
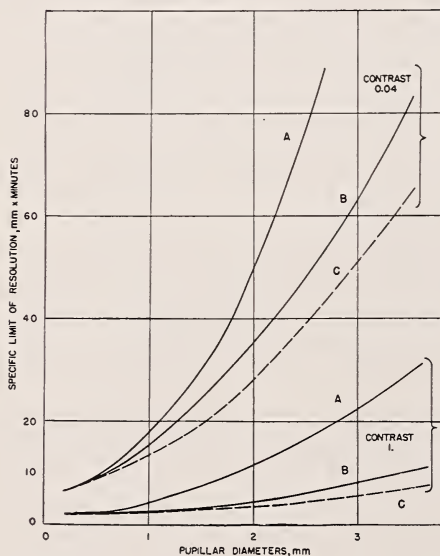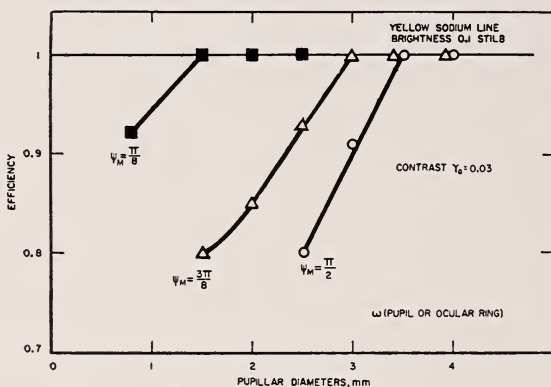

FIGURE 7.9. *Specific limit of resolution of the eye versus the pupillar diameter.*

A, without astigmatism; B, astigmatism (4 diopters, parallel to the lines of the test object): C, astigmatism (4 diopters, perpendicular to the lines of the test object).
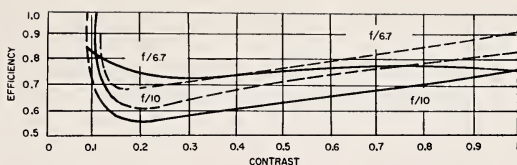
FIGURE 7.10. *Efficiency versus contrast of photographic objectives*
Solid lines, lens with third,order spherical aberration; dotted lines, well corrected lens.

Quality in the Whole Field of an Instrument

The question arises as to whether it is possible to evaluate with one number the quality of an instrument over its entire field.   Two very different cases occur.

1. *Instruments for pure observation.*   These are used to examine minutely, details in objects seen in the field of the instrument.   In this case we are interested in the resolution of details in the center of the field.   It would not be reasonable to include in the measurement the edge of the field when it does not enter into the examination of details.   This instrument will be characterized by the area within which the efficiency is greater than 0.9 times the efficiency in the center.

2. *Instruments to discover objects.*   In this case the periphery of the field has the greatest importance, as is the case for the periphery of the retina, the efficiency in foveal vision having practically no importance.   The efficiency to be introduced is the efficiency that corresponds to a point on the retina situated at a distance from the fovea equal to the field of the instrument, and the whole efficiency of the field will be equal to the mean peripheral efficiency for the whole field.

Instruments With Diffusing Screen (Photographic Emulsion)

Theory

The method that is described here has been completed by Mme. Marquet.[4]   Similar methods have been studied by various other workers.[5]

The principle of the method is the same as for the visual case, but instead of the angular resolving power of the eye, we have here the linear resolving power of the eye-emulsion combination, the optical image of the object being focused on the emulsion with a perfect objective of variable aperture.   If S and T are the resolving powers of the instrument, Ω and $n \sin U$ the linear or numerical apertures of the objective in the object space, $n \sin u$ the numerical aperture in the image space for a perfect instrument, we have

$$S\Omega = 2n \sin U \cdot T = 2n \sin u \cdot t.$$

The total efficiency E, which is similar to the total efficiency in

---

[4] M. Marquet, Sciences et Industries Photographiques, [2]*18*, 129–142 (May 1947).
[5] A. Couder, Cahiers phys., No. 14, pp. 35–48 (May 1943); Sciences et Industries Photographiques,[2] **14**, 170–174 (August 1943).   L. E. Howlett, Can. J. Research (July 1946).   E. W. Selwyn and J. L. Tearle, [B] (Sept. 1946).

visual instruments, and is used in the calculation of the performance, is given by $E_t = t_\infty/t$, where t_∞ is the resolving power with a perfect objective, the numerical aperture of which is very large. t_∞, being independent of the aperture, is called the resolving power of the emulsion.

The absolute efficiency introduces two kinds of losses in the resolution

1. The increase in resolving power of the perfect instrument as it is stopped down, because the diffraction pattern becomes more and more important in comparison with the area of granularity-diffusion in the emulsion. We evaluate it by the aperture efficiency $E_p = t_\infty/t_0$, where t_0 is the resolving power of the emulsion for a perfect instrument with the same aperature as that of the tested instrument.

Let us note that for ordinary emulsions, $t_0 \sim t_\infty$ for small apertures, 2 sin $u = 0.1$, for example. In the case of photographic lenses, the aperture efficiency is rarely important. However, it is always important in the photographic microscope.

2. The decrease in resolving power caused by all the defects in the image produced by the instrument is given by the instrumental efficiency $E_i = t_0/t$. This is the ratio between the resolving power t_0 of the emulsion with a perfect instrument, the aperture of which is the same as the aperture of the tested instrument, and the resolving power t of the emulsion with the tested instrument.

3. We have then, $E_t = E_p \cdot E_i$.

Experimental Procedure

The test objects are illuminated as for visual instruments. We generally use the higher brightnesses only. In order to measure t in the case of the perfect objective, it is necessary to use an instrument sufficiently good to be considered practically perfect. We therefore use astronomic lenses with very short focal lengths, between 20 and 50 mm, or apochromatic microscope objectives, set in a small camera with a high-precision movement for focusing.

A very important point is to know what magnification is to be used when observing the photographic images, this magnification depending upon the conditions of use.

When the photographs are to be seen without magnification, the observation is done with the naked eye at a standard distance of 250 mm.

Frequently we want to determine the best resolving power. It is obtained, for all practical purposes, when the area of granularity-diffusion in the emulsion is seen from the eye within an angle of 4 to 10 minutes. This is obtained for a resolution test when the angular resolution for the eye is between 3.5 and 8 minutes.

Example of an Efficiency Determination

Figure 7.10 gives the variation of efficiency with contrast in the center of the field for an excellent photographic lens operating at apertures of f/6.7 and f/12, and for a lens with third-order spherical and chromatic aberrations.

We might notice two interesting results. A photographic lens, considered to be an excellent one, never reaches an efficiency of 1 for

any contrast, even for $C=1$. This has already been noted by many observers.

The loss in the efficiency is largest for a contrast between 0.1 and 0.2, after which the efficiency increases rapidly, until the perceivable image disappears. This is a different result from that obtained in the case of a visual instrument. This occurs because the emulsion does not reproduce small contrasts very well, and the defects of the instrument have less and less effect on the resolving power.

Conclusion

The method given above, used now for 20 years for visual instruments, and nearly 10 years for photographic instruments, has satisfactorily solved, for us, the problem of determining the quality of an instrument in any point in the field. It avoids the discrepancies that occured very often between the results of tests in the laboratory and the practical use of the instrument. However, it has the serious disadvantage of not separating the eye from the optical image. If it were possible to separate them, then by combining the quality of the optical image, without any receiver, and the special properties of any receiver, it would be possible to predict the quality of the instrument for any condition. This is not yet possible because we do not yet know enough either of the variations of the distribution of illumination with various aberrations taken together, or of the properties of receivers. However, the experimental method maintains its interest because it is rapid, and because it takes into account the totality of the effects produced by instrument defects, some of which cannot be forecast by computation.

Discussion

DR. H. OSTERBERG, American Optical Co., Stamford, Conn.: A question for Dr. Macdonald. If your answer to this question is "Yes," I would like to ask one more. You mentioned the discrimination of squares and circles.

DR. D. E. MACDONALD, Optical Research Laboratory, Boston, Mass.: Yes.

DR. OSTERBERG: In connection with one of your charts?

DR. MACDONALD: I have answered "Yes."

DR. OSTERBERG: Have you studied the proposition as to when an individual would judge a square that has rounded corners to be a square as a function of the distance he stands from this figure?

DR. MACDONALD: I could answer the question "Yes" but I can't give you the information as to what the functional relationship is at this time. I would like to point out that I mentioned in qualifying this—.

DR. OSTERBERG: May I interrupt? I am not going to ask for such a complicated functional relation as you think. What I am interested in is this. What is the ratio of the radius at the corner of the square to the diagonal when this person reports this square with rounded corners to be a square?

DR. MACDONALD: We don't have the information in that form at all. As I tried to point out at the start, this was preliminary data presentation on the basic discrimination or recognition only of this

one type, the simplest type of discrimination we could think of, that is, between a square and a circle. Under different conditions of resolution—this involves the rounding of the corners, but we did not purposely introduce a form factor of this type—we have presented the same information to the analysts at different observing distances to bring in the effect of variable magnification.

As I say, I am not prepared to give any answers on any aspect except the basic square-circle differentiation at this time.

Dr. Osterberg: Thank you.

Chairman: Are there other questions that anybody would like to ask?

Dr. H. R. J. Grosch, General Electric Co., Lockland, Ohio: I would like to point out that these problems of interpretation of photographs, et cetera, remind me very much of the work done in educational and psychological testing because we have essentially a subjective phenomenon and we have experienced photointerpreters who would be willing to tell us this was or was not a good photograph but who would be unable to describe exactly why they thought so.

On the other hand, we have a series of measurements or tests that can be performed under controlled conditions but that do not measure the parameters of the problem partly because we don't know which are important. For instance, we assign arbitrary names such as "contrast" and "resolution" to certain things, but it is by no means certain that these are pure factors just as the terms "intelligence," "quickness," "retentiveness," don't mean too much.

Now, there is a mathematical technique that ranges over into the field of fantasy at times, known as "factor analysis." There are many proponents of this in Washington, hiding in cubby-holes of the Pentagon and similar sin spots, and I would like to suggest both to Howlett and Macdonald that it might be possible to run a standard factor analysis relating—or perhaps I should say correlating—the subjective judgments of their trained observers with measurements that they have made of the objectives, such as motion and stability of the camera.

Dr. Macdonald: I think that is exactly what we are trying to do with this statistical analysis of photographs. We are trying to bring out these subjective factors that are so enmeshed in the personality structure of the photointerpreter. The only way to drag them out seems to be through some statistical form.

We thought perhaps the incentive in a commercial organization for fuller interpretation might result in a more uniform background of training, one that we could trace back more easily than we could in the case of military photointerpreters. So analysis in this type has so far been restricted to the commercial applications of photography.

In terms of the basic units—and this ties in with the definition that we have to make if we are going to define the information content of a picture—we need, perhaps, a basic point where there is a binary decision that the photointerpreter makes. There are several ways of deciding what type of decision he makes. For example, if he is scanning along a line, his decision would be either "Yes, there is an edge," or "No, there is not an edge," as he goes from blur point to blur point.

I am talking about a discreet one-dimensional case and I have mentioned previously that the problem is one of moving to a two-

dimensional source and a continuous case. If the existence of an edge is a logical unit of information, then its evaluation depends on (a) the absolute magnitude of the change, and (b) the gradient.

CHAIRMAN: I believe we had one other question.

MR. A. H. KATZ, Photo Reconnaissance Laboratory, Wright Air Development Center, Dayton, Ohio: This is just an additional comment on the remarks made by Dr. Grosch, to the effect that it is only lately, after 10 years of talking to photointerpreters, that they have even been willing to realize that there is something in the picture other than the scale. This is a waste of good standard deviations.

DR. R. A. WOODSON, Armour Research Foundation of Illinois, Institute of Technology, Chicago, Ill.: I would like to direct a question to two gentlemen, Dr. Howlett and Prof. Arnulf. They both deal with the same question. Should we test a lens in the way it is intended to be used, or should we magnify the image for its inspection? Now, let's consider this first on the basis of the decision of whether a basic lens design is satisfactory, and, second, on whether an instrument made to such a design satisfies the performance requirements.

If you consider that in production manufacture of an optical instrument the operators will have variable visual acuity, you will have disagreements as to the rejection and acceptance of an instrument, whereas if the image is viewed under magnification you eliminate that human factor. I would like comments, please.

CHAIRMAN: Mr. Arnulf, do you wish to comment? Is Dr. Howlett here?

DR. L. E. HOWLETT, National Research Council, Ottawa, Canada: Mr. Chairman, I want to say that it would be acceptable under the circumstances outlined to use whatever degree of magnification was required. Whatever visual criterion is used in the test could be one well correlated with the circumstances of use. Such a procedure would make the test easier, too.

DR. W. WALLIN, U. S. Naval Ordnance Test Station, Inyokern, China Lake, Calif.: Dr. Howlett, it strikes me as somewhat anomalous and bothers me that the philosophy you express seems quite adequate as a functional test, suitable as inspection technique; but let us suppose that we have established to our satisfaction that such a test is inadequate. Then we want to arrive at something more analytic, divide and conquer. There are factors present in the lens system, itself, factors present in the photographic emulsion, let us say, and factors imposed by external circumstances.

If we are to understand the problem we have to analyze them and separate them, and yet once we do that we depart from your basic philosophy, don't we? I would like your comments; or haven't I made my question clear?

DR. HOWLETT: Well, perhaps you have. I will try and answer it. I think there are two distinct things involved there. Perhaps I did not make them sufficiently clear but I did refer to them. Certainly, any philosophy of evaluating quality of an image from the point of view of the user must, in the last analysis be a routine test, but that still has no bearing on any other test. You can do all the analyses that you like but keep it separate until it is in a form where it can be incorporated as an explanation or where it will lead to improvements in the evaluation. You can still evaluate the quality of an image on the photographic emulsion even if you don't know all about the for-

mation of the image, the developers, the various lens elements, et cetera, all of which are extremely important.

Having evaluated in an orderly manner the quality of the image, you can still go on.

DR. WALLIN: You are not saying then that the information we get is not going to help us?

DR. HOWLETT: I did not understand that.

DR. WALLIN: The question was, should we really rely on visual analysis in evaluating a lens that is to be used photographically?

DR. HOWLETT: I think it is quite reasonable to do so if you can show that by so doing your results correlate precisely with the photographic procedures that are going to be used. We have found that there are particular visual points that are always within a reasonable tolerance of where the photographic plate should be. I think that is perfectly acceptable and saves an enormous amount of time.

8. Image Quality as Used by the Government Inspector of Visual Telescopic Instruments

By H. S. Coleman [1]

Introduction

There are a considerable number of points of view from which image-quality evaluations can be considered. It is the purpose of this paper to discuss image-quality evaluation from the point of view of a government inspector whose responsibility is to determine whether or not a given optical system submitted to the Government conforms with its specifications. Because of the widely different procedures that might be followed in inspecting optical systems, it is the purpose of this paper to describe what are perhaps the most popular three inspection processes used in inspection and to show the correlation among the image-quality evaluations that are obtained among these three methods.

For purposes of presenting a description of the above mentioned test equipments, this paper is divided into three parts, corresponding to the three test procedures under consideration. The first of these test procedures involves the use of a device referred to as "the Kinetic Definition Chart Apparatus" [2] (because of its motorized parts), the second involves the use of a "Twyman-type Interferometer," [3] and the third involves the use of a device referred to as a "Dioptometer." [4]

The Use of the Kinetic Definition Chart Apparatus in the Inspection of Visual Telescopic Systems

The following section of this paper describes the Kinetic Definition Chart Apparatus (hereafter abbreviated to K. D. C. apparatus) and presents examples of the type of optical data that can be obtained with this equipment. The basic principles upon which this device depends were first described by Fabry [5] in 1935.

The K. D. C. apparatus can be regarded as an apparatus for measuring the "resolving power" of optical systems in which the spacing between the elements of the test object used and the contrast of the elements of the test object used can be varied continuously.

The Apparatus

The K. D. C. apparatus is essentially a continuously variable resolving-power apparatus in which the target is made to appear (by optical collimation) at any desired range and in which the contrast of

[1] Director of the Scientific Bureau, Bausch & Lomb Optical Co., Rochester, N. Y.
[2] H. S. Coleman and S. W. Harding, J. Opt. Soc. Am. **37**, 263 (1947).
[3] H. S. Coleman, D. G. Clark, and M. F. Coleman, J. Opt. Soc. Am. **37**, 671 (1947).
[4] H. S. Coleman, M F. Coleman, and D. L. Fridge, J. Opt. Soc. Am. **41**, 94 (1951).
[5] C. Fabry, Proc. Phys. Soc. London **48**, 747 (1935).

the target may be varied thus simulating conditions existing in the field. In addition, the test object may be surrounded with an illuminated region designed to simulate natural backgrounds against which targets may be viewed. The apparatus consists of a test object, an optical-reduction unit, a collimating unit, a standard telescope, an auxiliary telescope, an artificial sky, an off-axis fixture, and a mechanism whereby the angular subtense of the elements of the test object may be varied continuously. Each of these parts is discussed separately.

Test object. The basic type of test object used was first introduced by Foucalt [6] in 1858. This object consists of alternate white and gray bands of equal width of the form shown in figure 8.1 In general, it has been found desirable to use test objects in which the markings are oriented in four directions. The relatively coarse numbers shown in figure 8.1 are for purposes of identifying the test object and for focusing. The K. D. C. test objects can be made having any desired spacing and contrast between the white and gray bands. These are calibrated photoelectrically for the number of bands per unit length and for contrast. The test object is illuminated by means of a sphere as shown in figure 8.2.

Optical reduction unit. The apparent size of the test object is reduced optically by means of any one of a set of lenses mounted in a turret shown in figures 8.2 and 8.3. These lenses are referred to as the optical-reduction unit. The apparatus in its present form utilizes microscope objectives having focal lengths of 4, 8, 16, and 32 mm., and Plössel-type eyepieces having focal lengths of 20 and 50 mm., respectively. The optical-reduction unit is required in order to obtain practical limiting distances of resolution. As a result of this feature, the entire length of the K. D. C. apparatus is approximately 6 feet. Early models were as long as 120 feet.

Collimating unit. The collimating unit is used to make the test object appear (optically) to be at any desired range. It consists of a high-grade telescopic objective located such that its focal plane may be placed at any desired distance from the image of the test object formed by the optical-reduction unit.

Standard telescope. The standard telescope consists of highly corrected optical parts and is provided with efficient stray light stops. The objective has a clear aperture of 2.5 inches and an effective focal length of 15.75 inches. Provision is made for varying the aperture of the standard telescope by means of a set of stops so that its entrance pupil can be made equal to that of the optical device under test.

The standard telescope is used to establish the minimum angle of resolution for a "perfect" optical system of a given aperture, to acquaint the observer with the appearance of an image formed by a high-grade optical system, and to adjust the collimating unit for the desired target distance. This adjustment is made by focusing the standard telescope on an outdoor target placed at the desired range and then by adjusting the distance between the collimating lens and the image of the test object formed by the optical-reducing unit so that the test object is at the same focal setting as the outdoor target.

Auxiliary telescope. The auxiliary telescope is an astronomical-type telescope having an entrance pupil of 1.250 inches, which is provided with a set of eyepieces making it possible to vary its magnification from 1 to 30x. The auxiliary telescope is used in series with the instrument under test as shown in figure 8.2.

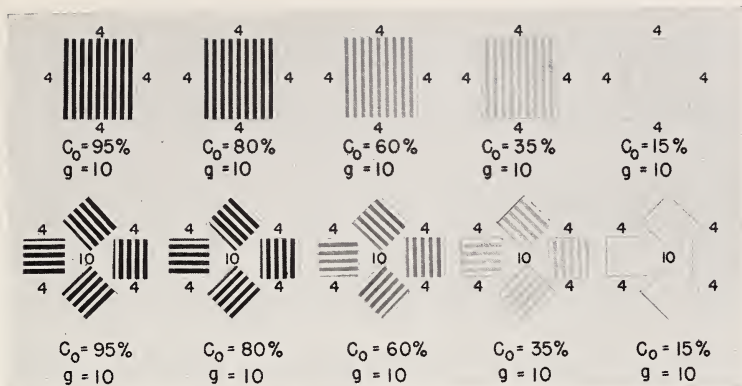[6] L. Foucault, Ann. Observ. Paris, **5,** 197 (1859).

FIGURE 8.1. *Modified Foucault resolution targets.*

$$C_0 = \frac{B-b}{B} = \frac{\text{Reflectivity of white band} - \text{Reflectivity of dark band}}{\text{Reflectivity of white band}};$$

C_0=inherent target contrast; g=number of white bands per inch
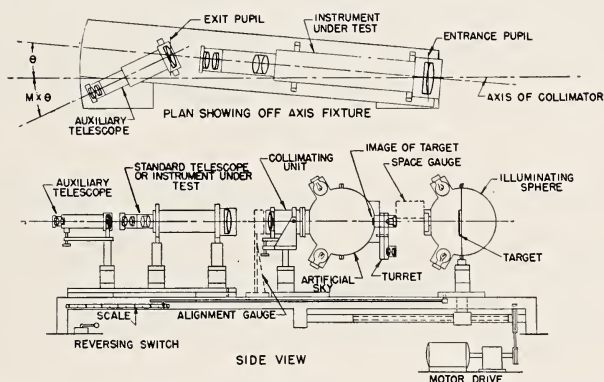

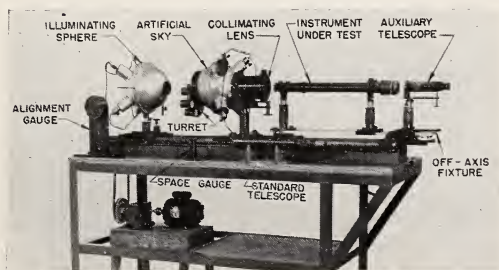FIGURE 8.2. *Schematic diagram of Kinetic Definition Chart apparatus.*

FIGURE 8.3. *The Kinetic Definition Chart apparatus.*

In the quality-control processes, the auxiliary telescope has three functions. The first of these is to make it possible to test the system under consideration at full aperture. This is done by reducing the exit pupil of the combination of the system under test and the auxiliary telescope to less than that of the natural pupil of the eye. Frequently the aberrations in optical systems are greatest at the margin of the aperture and may not be noticeable if the pupil of the eye becomes the effective stop of the system. In addition the pupil size varies from inspector to inspector. The second function is to reduce the portion of the pupil of the eye being used in the measurement to such an extent (1 mm or less) that the eye performs as a "perfect" optical device. This practically eliminates differences in measurements made by various observers. The third function of the auxiliary telescope is to produce a sufficiently large image on the retina of the observer's eye to make an observation "comfortably." In this connection, it should be noted that the auxiliary magnification does not provide a means of exceeding the limit of resolution of the eye because of the inverse relation between the size of the exit pupil at the eye and magnification, and as a result of the limitations imposed by diffraction theory.

Artificial sky. In order to simulate outdoor conditions, the image of the test object formed by the optical-reduction unit is surrounded by an illuminated region. This region is referred to as "the Artificial Sky" and subtends an angle of approximately 30° measured from the center of the collimating objective. The artificial sky is constructed so that it forms an adjustable field stop for the optical-reduction unit. This adjustment is accomplished by means of a series of spherical segments having various apertures and mounted in a turret as shown in figure 8.2.

The artificial sky has two functions. The first is to produce a known brightness level to which the observer's eye is to be adapted. The second function provides a means of taking into account the stray light normally introduced into an optical system by natural backgrounds when the optical instrument under consideration is performing its intended task. Stray light in optical devices reduces the contrast of the image on the retina and, consequently, the range at which objects are visible.

Variable spacing unit. The mechanism for controlling the angular subtense of the target markings consists of a motorized unit by means of which it is possible for the observer to vary continuously the distance between the test object and the optical-reduction unit. This variation changes the size of the image formed by the optical-reduction unit and, consequently, the angular subtense of the target markings. The distance between the optical-reduction unit and the test object is indicated on a scale graduated in tenths of an inch. The optical parts are arranged so that the scale reading is directly proportional to the minification of the image of the test object formed by the optical-reduction unit. In this connection it might be noted that for the range over which the target is moved, the longitudinal displacement of its image formed by the optical-reduction unit is much less than the focal range (depth of focus) of the collimating objective and hence the image viewed is always in focus. This means that neither the instrument under test nor the auxiliary telescope have to be changed in focus when making a K. D. C. measurement.

By virtue of the continuous feature of varying the distance between the test object and the optical-reduction unit, precise resolution measurements can be made since it eliminates the use of the discreet angular intervals of conventional resolution objects.

Gages. In order to facilitate the adjustment of the K. D. C. apparatus two special gages have been devised. These are shown in figure 8.2 and 8.3.

The first of these gages, referred to as the space gage, is used to set the distance between the target and the optical-reduction unit so that the scale reads linearly with minification. This gage is simply a metallic block 5 inches long. Appropriate scale readings have been determined for each of the optical-reduction units when the gage is placed between the face plate of the test-object illuminator and the cell of the optical-reduction unit.

The second gage, referred to as the alinement gage, provides a means of alining the optical axes of the test object, the optical-reduction unit, and the collimating unit. This gage is essentially a combination two-dimensional square and height gage and is used as indicated schematically in figure 8.2.

Off-axis fixture. The K. D. C. apparatus is provided with a mechanical assembly for mounting the optical system under test for making resolution measurements at various field angles. This fixture, referred to as "the off-axis fixture," is designed so that the optical device under test can be alined with the optical axis of the collimating unit and can be rotated about its own entrance pupil. The off-axis fixture also provides a means of rotating the auxiliary telescope about the exit pupil of an optical device under test.

The Procedure

The end product of the K. D. C. test is a quantity referred to as the K. D. C. efficiency. This efficiency is a numerical measure of the quality of an optical system. The axial K. D. C. efficiencies of the better present-day optical systems approximate 100 percent. The K.D.C. efficiency may be used in either of two different forms depending upon whether the measurements are to be used for predicting the performance or for determining the quality of an optical device under consideration.

In predicting the performance of telescopic systems, it is desirable to know how much farther a given object can be seen with the instrument under consideration than with the unaided eye. This point of view gave rise to the definition of the K. D. C. efficiency given by eq 1.

$$\text{K. D. C. efficiency} = \frac{X_i'}{M_i X_e} \times 100, \tag{1}$$

where

$X_i' =$ K. D. C. scale reading at limit of resolution for the eye plus instrument under test,

$X_e =$ scale reading at limit of resolution for the unaided eye,

$M_i =$ the magnification of the instrument under test.

In determining maximum resolution of an optical system, auxiliary magnification is generally used for the reasons indicated above. In

this case, K. D. C. efficiency is defined by eq 2, which is equivalent to the definition presented in eq 1, when less than 1-mm diameter of the eye is used.

$$\text{K. D. C. efficiency} = \frac{X_i'}{S} \times 100, \tag{2}$$

where

$X_i' = $ K. D. C. scale reading at the limit of resolution for the instrument under test using the auxiliary telescope,

$S \ = $ K. D. C. scale reading at the limit of resolution for the standard telescope having the same entrance as the instrument under test and using sufficient auxiliary magnification to eliminate the imperfections that may be introduced by the observer's eye.

It will be noted that when one uses eq 2, he is simply comparing an instrument under test with one known to be of high optical quality and to be practically free from stray light. It should be noted, however, that since the standard telescope has a constant specific coefficient of resolution (defined to be the product of the aperture in inches and the minimum resolvable angle in seconds) of 4.51 inch seconds, the K. D. C. efficiency may be converted into angular units. In making K. D. C. measurements, the limit of resolution is taken to be the scale reading at which the direction of the markings at any one of the four orientations cannot be specified.

The procedure for making K. D. C. measurements for telescopic systems may be divided into two parts. The first of these consists in axial measurements using test objects of various contrasts. The second is the measurement of resolution at various field angles with a test object of fixed contrast. The procedure consists simply in noting the K. D. C. scale readings at the limit of resolution for each target contrast or field angle. The results of such measurements are plotted in the form shown in figures 8.4 and 8.6. As indicated, measurements may be made without the Artificial Sky in operation and at different brightness levels. In order to illustrate the importance of surface defects and cleanliness a thumbprint was placed on the center of the objective of the standard telescope to provide a source of stray light as this instrument is practically free of such defects. In the measurements presented, a uniform surround was used and the target was at the same apparent brightness as the surround and an auxiliary telescope was used. Figure 8.5 is presented to illustrate the relation of K. D. C. efficiency to the scale readings presented in figure 8.4. It is to be noted that the K. D. C. efficiency for any field angle is 100 percent by definition for the standard telescope with clean optics. Figure 8.7 shows an example of the K. D. C. efficiency obtained for a 5x telescope at various field angles without the use of auxiliary magnification.

Possible errors. There are two principal sources of error in making measurements using the K. D. C. apparatus. The first of these is in focusing on the image and the second is the failure to expose the image of the object to the eye long enough. The error in focusing has been reduced by focusing on numerical figures rather than on the resolution markings of the test object. The exposure time required may be as

FIGURE 8.4. *Kinetic Definition Chart scale readings (x) versus target contrast for a standard telescope with and without artificial sky.*
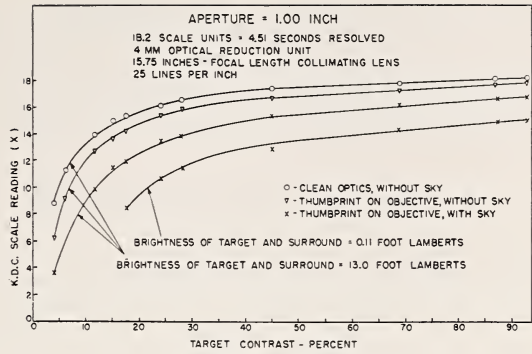
FIGURE 8.5. *Kinetic Definition Chart efficiency versus target contrast for standard telescope with thumbprint on the front of the objective.*
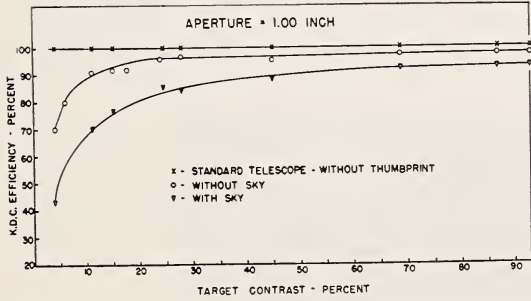
FIGURE 8.6. *Kinetic Definition Chart efficiency versus field angle for 3x telescope with and without artificial sky.*
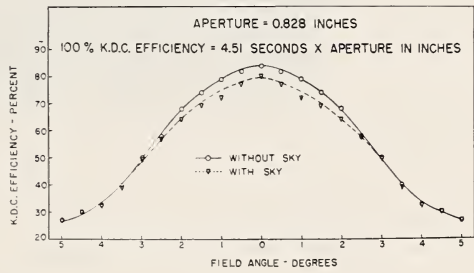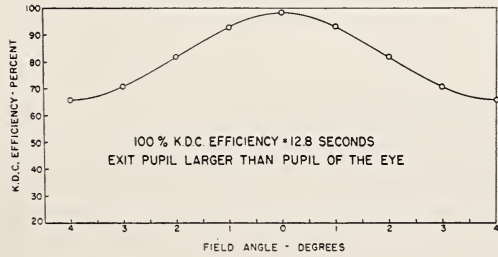
FIGURE 8.7. *Kinetic Definition Chart efficiency versus field angle for 5x telescope without auxiliary magnification.*

101

much as 10 seconds for low-contrast objects or for poor-quality images. This error appears to be related to unsteadiness of the observer's head as the required exposure time can be reduced by using a chin rest. If the above errors are minimized, the probable error in the resolution measurements made using the K. D. C. apparatus does not exceed 2 percent in the case of optical devices of high quality.

The Dioptometer

The dioptometer used to test the instruments mentioned in this paper is of a fairly classical design and therefore is described briefly. It is shown schematically and photographically together with its accessories in figures 8.8 and 8.9. The dioptometer is a simple telescopic system consisting of an objective, an eyepiece, and a reticle, and is calibrated to measure the distance an object appears to be from its objective in units referred to as "the diopter". The diopter is the reciprocal of the apparent target distance in meters. This device is used to focus on the image of a test object, in the shape of a cross, formed by an optical instrument under test. If both the vertical and horizontal parts of the image of the cross can be brought into sharp
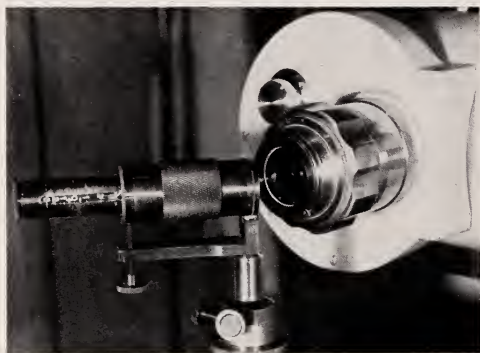

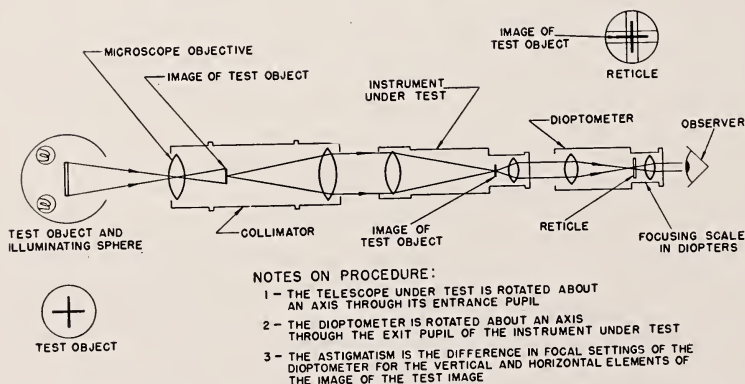FIGURE 8.8. *D–2 dioptometer used to measure the astigmatism of telescopic systems.*

FIGURE 8.9. *Schematic diagram of the apparatus used to measure the astigmatism of telescopic systems by means of the dioptometer.*
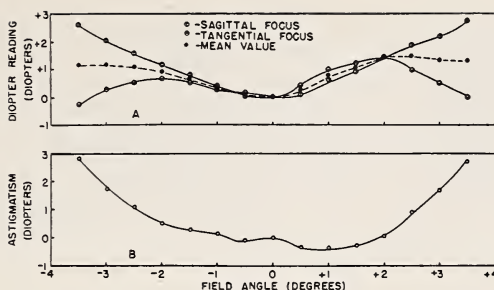
FIGURE 8.10. *Dioptometer readings versus angle of view.*

A. Analysis of a 7- by 50- by 7-degree prism-erecting telescopic system; B. Difference in dioptometer reading between the two foci.

focus when viewed with the dioptometer, the instrument is said to be free from astigmatism. If the two parts of the image of the cross cannot be brought into sharp focus for the same scale reading of the dioptometer, the instrument under consideration is said to have astigmatism. The magnitude of astigmatism is determined by focusing the dioptometer first on the vertical and then on the horizontal parts of the image of the cross and noting the dioptometer scale readings thus obtained. The difference in focus between the vertical and horizontal lines thus obtained is referred to here as astigmatism. Measurements of astigmatism are usually made for various field angles.

Typical dioptometer data are shown in figure 8.10.

The Interferometer

The interferometer used to test the instruments mentioned in this report is shown schematically and photographically in figures 8.11 and 8.12. It resembles the well-known Twyman instrument except in certain mechanical details. The procedure used to evaluate the quality of an optical system, however, is new and is therefore described. Most attempts to use the interferometer for purposes of evaluating the quality of optical systems have been confined to lenses and prisms and depended upon the counting of the number of interference fringes obtained in some specified manner. Such task is almost hopeless in the case of the great variety of complex interference patterns obtained in testing telescopic systems. Because of the difficulty in analyzing the interferometer patterns produced by complex mixture of aberrations, consideration was given to the possibility of evaluating the quality of an optical system on a basis that did not require the specification of the magnitudes of the specific aberrations. Since the main use of the interferometer, from the point of view of this paper, was to evaluate the optical quality of visual telescopic systems, an investigation was made of the type of interferometer patterns that were produced by telescopic systems focused visually. In this investigation a number of instruments were focused on numerals subtending angles well above the limit of resolution and placed optically at infinity. The investigation indicated that the best visual focus was found to give rise to interferometer patterns having one part covered by a single interference area. This region is the area over which the optical-path-difference does not exceed one-quarter of a wavelength (taking
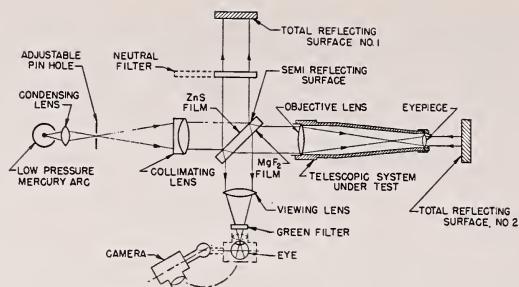
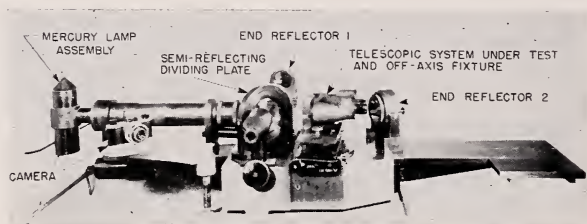FIGURE 8.11. *Schematic diagram of interferometer used to test telescopic systems.*

FIGURE 8.12. *Interferometer used to test telescopic systems.*
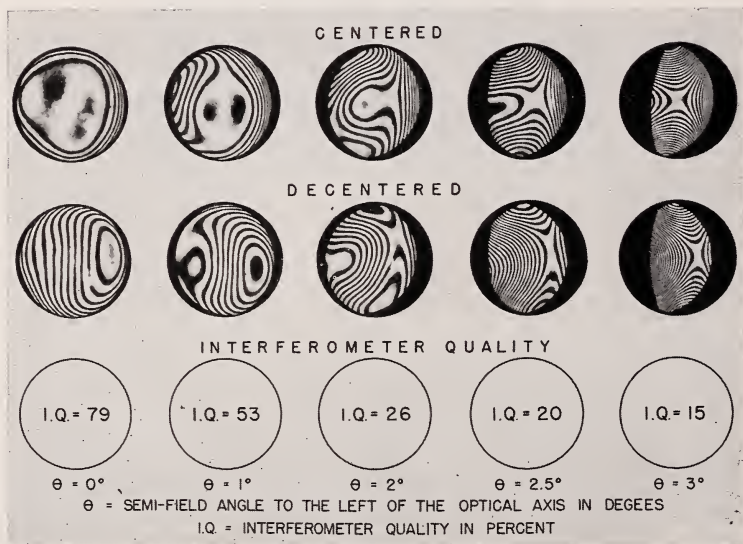

FIGURE 8.13. *The interferometer quality at various field angles of the right barrel of the 7- by 50- by 7.1-degree binocular, Mark 28 Model O serial no. 197549.*

Adjusted for best focus.

into account the double passage of the light through the telescopic system under test). The diameter of this area was found to decrease by any slight change from best visual focus. The diameter of the largest circle that can be inscribed in this area is taken to be a measure of the optical quality. When this diameter is expressed as the percent of the axial entrance pupil of the system under test, the percentage is referred to as the interferometer quality (hereafter interferometer quality may be abbreviated to I. Q.). The diameter of the inscribed circle is referred to as the I. Q. circle, and is defined in terms of the area over which the optical-path-difference, introduced by a system under test, does not exceed one-quarter of a wavelength. For practical purposes, the interferometer quality might be thought of as the percentage of the entrance pupil of the system under test that is free from aberrations.

A sample of the data obtained showing the interferometer quality of an instrument is presented in figure 8.13. Because of the symmetry found for the aberrations across the field of the optical systems studied, data are presented only for the semifield angles of 0, 1, 2, 2.5, and 3 degrees to the left of the optical axis.

Conclusions

Certain general conclusions have been drawn concerning the various experimental procedures mentioned in this paper. These conclusions are stated briefly for each type of apparatus.

The K. D. C. Apparatus. The conclusions concerning the use of the K. D. C. apparatus are summarized as follows: (1) The K. D. C. apparatus is a versatile device that may be used to measure resolving power of a variety of optical devices including telescopes, photographic objectives, microscope objectives, eyepieces, and the human eye; (2) the K. D. C. apparatus provides a means of taking into account the brightness level to which the human eye is adapted, the ill effects introduced by stray light, and the target contrast; (3) the K. D. C. apparatus is sufficiently rapid and impersonal to be used for quality-control purposes for visual telescopic systems being mass produced; (4) the K. D. C. apparatus provides a means of objectively evaluating the design of optical systems; (5) the K. D. C. apparatus may be used to obtain data required for predicting the performance of optical systems.

The Dioptometer. The dioptometer described in this report provides a simple means of evaluating the quality of optical systems in which the predominant aberration is astigmatism.

The Interferometer. The conclusions reached concerning the use of the interferometer mentioned in this report are as follows: (1) The interferometer may be used to evaluate the quality of optical systems of moderate dimensions being mass produced; (2) the procedure described for using the interferometer in this paper is simple and rapid and does not require a technically trained operator; (3) the interferometer is the most precise test fixture for evaluating the quality of optical systems available at the present time; (4) It is evident from figures 8.14, 8.15, 8.16, and 8.17 that a high degree of correlation exists among the data obtained using the K. D. C. apparatus, the interferometer, and the dipotometer for the instruments mentioned in this paper; (5) The high degree of correlation obtained

between the Dioptometer values of astigmatism values and the K. D. C. efficiency, indicates that astigmatism is the principal aberration responsible for the rapid decrease in optical quality across the field for the present design. This may be seen by comparing figure 8.14 with figure 8.17.

FIGURE 8.14. *Average values of the Kinetic Definition Chart efficiency and the astigmatism in diopters for all the instruments tested.*

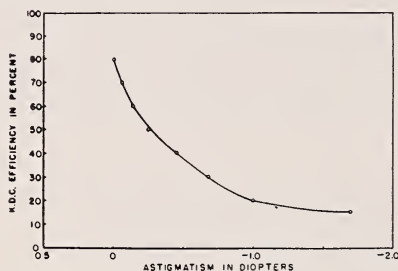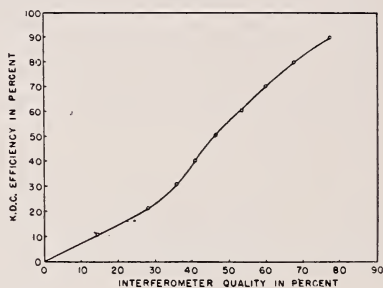FIGURE 8.15. *Average values of the Kinetic Definition Chart efficiency and the interferometer quality for all the instruments tested.*
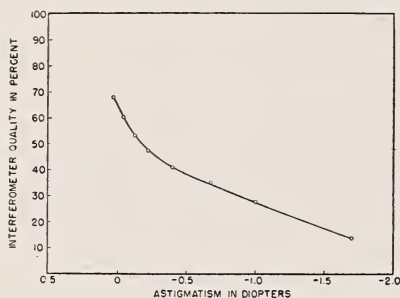

FIGURE 8.16. *Average values of the astigmatism and the interferometer quality for all the instruments tested.*
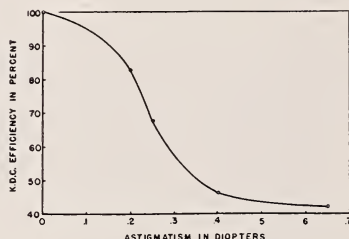

FIGURE 8.17. *Relation between Kinetic Definition Chart efficiency and astigmatism for pure primary astigmatism.*

9. Application of Fresnel Diffraction to Measurements of High Precision

By A. C. S. van Heel [1]

Introduction

On a previous occasion it has been emphasized that a direction can be determined with surprising precision by means of a diffraction pattern of the Fresnel class and especially when settings are made on the color transitions, occurring in the usually very diversely colored patterns obtained when a white light source is used.[2, 3] As one such application, a spherometer for large radii and a flatness tester were described. As this work has been continued and the results have been found to be of desirable accuracy, although the simple method has not been modified essentially, a description of these developments, together with some comments and corrections of the previous publication, appears justified. It will be shown that the scope of the method is exceedingly general and its application to the measurement of even very small aberrations is quite possible. The method is essentially a means of studying the structure of wave fronts, or, when the conception of light rays as normals to these is preferred, the structure of light pencils.

Study of a Flat Surface

The determination of the form of a nearly flat surface may serve as a relatively simple example and will be described in some detail.

In figure 9.1, S is a horizontal slit, some tenths of a millimeter wide, receiving light from a glow lamp or, better, from a carbon arc. At a distance a of about 2 m, a metallic grating G is placed, at TT′ the surface to be tested is placed. The light reflected from this surface is observed at MM′ by means of a traveling microscope of low magnification (10), reading to approximately 1 micron. The distances b and c are approximately 1 m.

In order to remove the observing microscope from the path of the incident light the surface is tilted about the line TT′ by an angle α, so that MM′ is not in the plane of the figure but behind it. When calculating the results this tilt has to be taken into account, a simple procedure for flat surfaces, but requiring a rather complicated correction formula when the surface is convex or concave. Fortunately, the formula can be simplified to

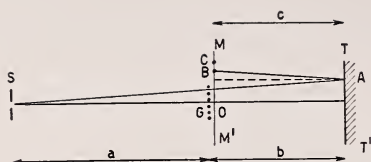$$\delta = g\bar{y} + hR\bar{y} + \ldots,\tag{1}$$
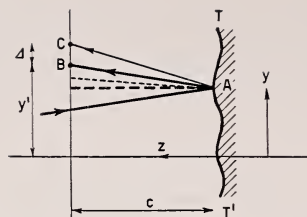
FIGURE 9.1. *Surface tester.*

FIGURE 9.2. *Slope derived from deviation.*

where δ is the correction to the observed distance OB, $\bar{y}$ the distance from SO at which the incident ray intersects the plane of the grating, and R the curvature of the surface; g and h are constants, depending on α, a, b, and c. The value of $\alpha = 2.09° \pm 0.05°$ allows the following terms of (1) to be neglected, as they amount to much less than the tolerances given by the precision of pointing, as long as the radius of the surface is greater than about 1 m.

In the following this correction is always assumed to have been applied, in order to simplify the discussion.

Derivation of the Surface Profile

The diffraction pattern displays its multiple and significant colors MM′ exactly as if this plane were at a distance $b+c$ beyond G and the reflecting surface were not present, provided that it is perfectly plane. Deviations from planeness give rise to changes in direction of the normal. Thus in the case of figure 9.1, if TT′ is not plane the normal at A might have a different direction, and the "correct" position of the intersection with MM′ of the normal to the reflected wave front or reflected "light ray" would be displaced from B to C.

When the deviations from planeness are small the variations of the slope are given with ample accuracy by $1/2\ BC/c$. Referring to figure 9.2 we can write

$$dz/dy = -\tfrac{1}{2}\Delta/c, \qquad (2)$$

where $\Delta = BC$ is the observed deviation from the theoretical point of intersection. The last is known from

$$y_0' = \bar{y}(a+b+c)/a, \qquad (3)$$

where $\bar{y}$ has the meaning given above.

Integrating the curve for dz/dy with respect to y we find the form of the surface $z\ (y)$.

Once the shape of the reflecting surface along a line lying in the plane of figure 9.1 has been established, the coordinates may be transformed (translation and rotation) when this is necessary to compare different series of measurements on the same surface. Again, when the surface has been tilted about an axis normal to the plane of figure 9.1, as may happen when the object has been removed, cleaned, and replaced, a rotation usually must be taken into account together with an accompanying division of c by the cosine of the tilting angle.

108

Straightness of the "Light Rays" and Basis of Measurement

Before proceeding to give detailed results, we must say some words on the "light rays" and the way measurements are performed. As has been stated in a previous publication (see footnotes 2 and 3), the maxima of light in the diffraction pattern correspond more or less to the slits in the grating and the minima to the wires, provided the observing distance $(b+c)$ is small enough to obviate any "forkings" as indicated in figure 4 of the cited article. Still the assumption that the loci of the colored maxima (and of the otherwise colored minima) are straight is even then only approximately true and needs testing. This can be and is done by removing the surface to be tested and measuring the pattern at different distances from G.

Since 1949 the traveling microscope has been improved; the consistency and accuracy of the readings have been considerably bettered, while magnification has been increased from 5 to 10. The small autocar lamp has been replaced by an arc lamp, which permits the color transitions in the pattern to be more clearly and precisely determined. Thus the measurements are more trustworthy.

It appears from careful observations that the observed curvature of these loci [4] is considerably smaller than was then thought and, for a large part, is to be ascribed to observation errors. Instead of the not very consistent curvatures of these loci, we can now state, with more confidence, that the loci are practically straight (with a grating period of about 0.6 mm and at distances from G up to 2 m). Deviations from straightness are never more than 3 microns, even up to the eleventh grating slit above or below the middle slit O.

This fact has emboldened us to speak of "light rays" as a short expression for these loci and facilitates in no mean measure their use. The way in which the curvature begins to play a part when the dimensions are altered and especially how the loci behave when the light is reflected has to be studied separately. For our present purpose we will let this point rest and make use of the fact that the loci are straight within less than $3 \cdot 10^{-3}/2 \cdot 10^3$, or 0.3 second of arc.

Not only the maxima and minima are colored, they are also accompanied by fringes of divers hues and intensities. Each period repeats itself with the same sequence of colors, symmetrical about the maximum as well as about the minimum. Thus there is more than one characteristic to determine their position.

Even with one characteristic for each period we found a mean error of 8.2 microns in five series of determinations on 19 maxima with four pointings on each maximum; the five series gave respectively 6.4, 8.3, 8.1, 9.6, and 8.3 microns.

The distances a, b and c were measured several times in a period of 6 months and could be trusted to about 0.2 mm.

At 19°.0 C their values were: $a=2342.1$ mm, $b=1204.2$ mm, $c=1179.4$ mm. A discussion of errors shows that the uncertainty of these distances has no appreciable influence on the final results. The supports being beams of steel, the values can be corrected for temperature with a dilatation coëfficient of 12.10^{-6}. The object whose surface was to be tested was applied by a double spring to a contact ring. The distances b and c were measured to and from this plane. When

---

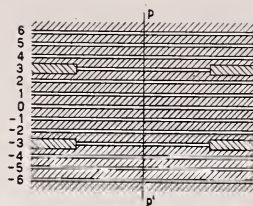[4] A. C. S. van Heel, J. Opt. Soc. Am. **40**, 812 (1950).

FIGURE 9.3. *Grating with partly blocked slits.*

FIGURE 9.4. *General view.*

convex or concave surfaces are studied a correction for the sagitta has to be applied.

During the measurements it is desirable to know with which slit of the grating one is concerned. This was effected by blocking parts of the slits as indicated in figure 9.3. The slit 0 is at O in figure 9.1, the slits 3 and −3 have been blocked except in the central parts. The coloration of the diffraction fringes corresponding to these is distinctly affected, thus they can be identified, but only at their marginal parts. Although the influence of the blocking extends farther towards the center than corresponds to the geometrical shadow of the stops, the neighborhood of PP′, where the measurements are made is unaffected. This was ascertained by a separate series of measurements.

Practical Realization

Figure 9.4 gives a general view of the instrument. It has been set up in a basement, whose temperature is constant for long periods of time within 1° C. Brackets in the wall support two strong steel pipes of 4-m length and 3.2-cm outside diameter. Slit S, grating G, microscope M, and object A are mounted on robust carriages clamped on the pipes. The light from an arc lamp (not shown in the figure) is thrown on to the slit by means of a mirror, so that the arc lamp is at a sufficiently large distance from the apparatus to avoid warming it by radiation or convection.

The microscope has a doublet object glass. The light pencils reflected by the tested surface, which are approximately horizontal, are received in the plane of a cross wire on which the microscope is focused. By means of a plane mirror the light is turned into the vertical direction to facilitate observation.

It may be noted in passing that optical parts are not inserted in the light path between the source and the cross wires (with exception, of course, of the test surface itself), but that beyond the cross hairs no error is introduced by any reasonably good optics. The warning may

be repeated here not to make use of a grating on glass, as every slight inhomogeneity of the glass and every deviation from flatness of the surfaces of it are deleterious to the perfect formation of the diffraction pattern. Thus only metal gratings with air gaps can be used.

Probably the gratings are best made by stretching wire of constant thickness over the grooves of two parallel precision screws. We have applied ourselves to the construction of such a harp, plugging the ends of each wire into holes in a robust frame, on which also the screws were mounted, but as yet either the wires were not sufficiently straight or their thickness, about 0.3 mm, changed by the too strong stress. We also tried metal sieves, whose holes act as slits, the light source being a slit. This entailed the necessity of calibrating each row of holes, this calibration being by means of a precision traveling microscope, provided with a supplementary cylindrical lens. Thus the position of each "slit" was known to within a few microns and had to be taken into account in the calculations. Although precision is not impaired seriously in this way the evaluations are somewhat tedious, and we are still trying to produce a grating with sufficient precision to avoid correction terms.

Flat Surface

A standard flat of natural quartz, cut perpendicular to the axis, diameter 65 mm, thickness at the center 20 mm, was one of the first objects to be studied. The cylindrical slab is slightly prismatic, having a thickness of 18 mm at one side and of 22 mm at the other side. The form of the standard surface was investigated in two perpendicular directions. In position I the virtual line of intersection of front and back surface is vertical; this gives the profile along the diameter a–a in figure 9.5. In position II the intersection line is horizontal, the thickest part of the slab being at the bottom; this gives the profile along the diameter b–b in figure 9.5. The back surface had been dulled with paraffin wax. The front surface has been studied as it came from the polisher.

The diameter of the contact ring, on which the slab was pressed with a force of the order of 1 to 2 kg, was 52 mm.

In figure 9.6 we give the results of the measurements in position II, as these were repeated under different conditions. Position I gives comparable results.

The first series gave the results indicated by open circles. The slab was then removed, cleaned again as it had been before the first series with a piece of cotton wool and acetone, replaced and the second
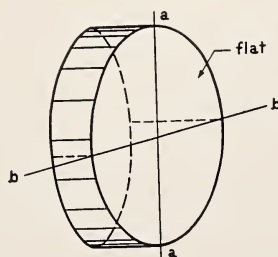

FIGURE 9.5.   *Tested slab.*

FIGURE 9.6.   *Profile of flat; position II.*

series was made, giving the dots in figure 9.6. Although the correspondence was very good, we were not satisfied with the fidelity to the true form of the profile unless it was tested again with a transverse translation of the surface. Only thus could we make certain that the ups and downs of the curve are to be ascribed to the surface and not to some faulty calibration of the grating or other error. The slab, therefore, was removed again, cleaned in the same way and replaced in a slightly higher position, shifted 6.5 mm compared to the former positions. The third series of observations then gave the values indicated by crosses in figure 9.6.

To each of the curves yielded by these series small vertical translations (in the vertical direction of the figure corresponding to the z-direction in figure 9.2) and small rotations have been applied, in order to obtain a satisfying coincidence. Even a small particle of dust on the contact ring gives rise to such variations of the position of the flat.

The curve corresponding to the third series, moreover, has been shifted in a horizontal direction along a distance of 6.5 mm.

We were gratified to find a very good coincidence of the three independent curves. The mean deviation from the mean curve amounts to only $1.2 \cdot 10^{-6}$ mm or 12 A or about $\lambda/475$.

The integration of the dz/dy-curves gives, as is well known in such cases, reasonably good results.

Further, a parabola (thin line in fig. 9.6) has been calculated to account for a mean curvature of this part of the surface (and along the line b-b of fig. 9.5). The radius of curvature is 2.8 km (concave).

It seems to us that the following can be inferred. (1) Irregularies of a reflecting surface can be studied in this way to about 12 A; it is to be understood that only the relative positions are found of *means* over areas of, as yet, unknown extension, because we have not yet ascertained how far the diffracted light from a given slit has an appreciable influence; a preliminary investigation seems to indicate 5 mm² as the order of magnitude of these areas. (2) The mean radius of a reflecting surface can be found, *again without touching the surface and without a reference surface*, over a range of 1 to 2 cm with a precision corresponding to an error in the sagitta of 12 A.

The last conclusion can also be put as follows: In the manner described, deviation from flatness of a reflecting surface is found in an "absolute" way when the surface is 1 to 2 cm in diameter and the radius is smaller than 34 km.

It is not possible without serious variation of the method, to find the thickness of parallel layers. Thus the surfaces a and b in figure 9.7 are not to be distinguished in the manner described.

As has been mentioned, the observations of the same flat in position I give analogous results (See fig. 9.8). Here two series were observed.

FIGURE 9.7. *Parallel layers not indicated.*

FIGURE 9.8. *Results of two series of measurements on a flat surface.*

For the central portion (11 mm) the mean error of the mean values is 11 Å or λ/520. The whole range of 18 mm of the measured profile, however, has a larger mean error (32 Å or λ/180), but inspection of the curves for the first series (open circles) and the second series (dots) of figure 9.8 shows that the general form of the surface is changed between the two series. The mean radius of curvature is 1,4 km for the first series and 2,4 km for the second series. Presumably, the force applied by the springs to press the slab on to the contact ring has been distributed in a somewhat different way in the two cases. This conclusion was strenghtened by an inspection of the construction of the springs, which are not very apt to act evenly when the wedge-like slab has its edge vertical.

A rather crude estimation of the stress made this explanation plausible. In position II the force of the two symmetrically placed springs will be less liable to be disturbed unevenly, even when the slab is shifted upwards.

Curved Surfaces

The study of curved surfaces proceeds along the same lines. Again the observations must be corrected for the tilt α and then can be compared to the theoretical positions, calculated with a preliminary value of the surface curvature.

In figure 9.9 we give the results obtained with the convex and concave parts of a gage of about 4 m radius. Again two positions, I and II, are investigated as in the case of the flat surface.

From the dz/dy-curves the profiles follow by integration and the curves thus obtained are reduced to a mean orientation of the surface. It is to be remembered that here the effects of the *mean* curvature has been left out. The radius for the convex surface is 3987.5 mm, for the concave surface 3984 mm.

On applying the two parts to one another and illuminating with sodium light, figure 9.10 was obtained; for this the surfaces were silvered with a rather thick transparent layer to get sharp interference fringes. The diameter of the gages is 62 mm, but only the central part with diameter 18 mm, where the fringes are nearly straight, has

113
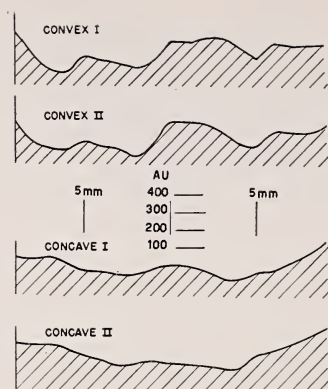
FIGURE 9.9.   *Profiles of convex and concave surfaces.*
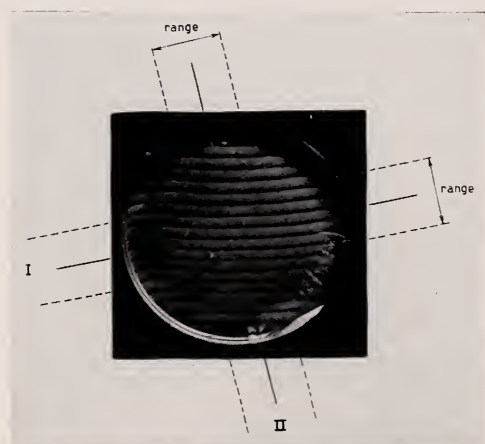
FIGURE 9.10.   *Interference test.*

been tested by the diffraction method.   The directions I and II are indicated in the figure.

The general trend of the profiles can be recognized in figure 9.10.

As an essential difference between the two methods of observation, the described diffraction method and the interference method, we want to emphasize that with the first the mutual influence of the two glasses is obviated.   Indeed, the possibility now arises of investigating the eventual influence of the presence of surface forces on the form of the surfaces.

Further, it is no mean advantage that each surface is studied separately without the aid of an auxiliary surface.   In that sense the diffraction method might be called an absolute one.

The spherometry of surfaces with large radii offers no difficulty with the apparatus as described.   The precision tends to be too great for practical and workshop uses.   For these purposes, however, there is no need to spend too much time on the observation of many lines.   If only the two marked or blocked lines are observed a pre-

cision is easily and speedily attained, which is amply sufficient for current use. In this connection it is valuable that no silvering or other treatment of the surfaces is needed to insure sufficient illumination of the image.

Focussing of Lenses or Mirror Systems and Determination of Aberrations

The preceding account of the study of reflecting surfaces applies largely to the investigation of the performance of lenses. Indeed, the reflected wave front has been the object of study. Only by dividing by 2 c were inferences made about the reflecting surface. Essentially the described method is a tool to look into the structure of plane or moderately curved wave fronts, that is, of parallel or moderately divergent or convergent pencils.

The determination of the position of the focus of a well corrected positive system can be done with extreme precision by placing the slit approximately at that point and studying the convergence or divergence of the emergent pencil by means of the grating, observations being made at say 2 m distance from it.

From the mean curvature of the emergent wave front, the distance of the slit from the focus can be calculated with any desired accuracy if the focal length is approximately known.

Assuming a mean error of the sagitta of the wave front of 25 A, the error in the position of the "mean" focus due to these observational errors amounts to:

> 0.6 micron for a focal length of 100 mm,
> 16 microns for a focal length of 500 mm,
> 60 microns for a focal length of 1000 mm.

If the aperture of the system is large enough and its correction sufficiently good this accuracy can even be bettered by making use of gratings of larger breadth.

Of more importance perhaps is the determination of aberrations. Path differences can be measured with a precision of 25 A. Conversion into other measures of aberration is only a question of elementary theory.

As was mentioned before,[5] the (axial) aberrations of a very good objective lens, f=60 cm, aperture 1/10, were determined with the given precision by means of the diffraction method.

A decided advantage over the diffraction methods working with a grating in the neighborhood of the focus is the great simplicity with which the observations are reduced to data of path differences, while on the other hand the apparatus is much simpler in construction than the lens interferometer. It is not known to the author what accuracy can be attained in practice with the interferometer. Only tests of given well corrected systems by the described method, might give some ideo of how the accuracies of the two methods compare.

The same applies to the elegant and simple method of Gardner and Bennett.[6]

When aberrations of optical systems are the chief object, the ap-

[5] A. C. S. van Heel, J. Opt. Soc. Am. **40**, 815 (1950).
[6] I. C. Gardner and A. H. Bennett, J. Opt. Am. **11**, 441 (1925).

paratus, up to now only used incidentally for this purpose, should be reconstructed in order to facilitate adjustment, especially when the observation is to be made photographically. A vertical mounting then seems to be indicated.

Photographic Method

It was felt during the somewhat lengthy observations that a more rapid recording of the diffraction patterns would be an improvement, as for instance, the influence of temperature variations should be obviated. Therefore, exposures on Kodak Ektachrome and Kodachrome films have been made. It appeared that the times of exposure were reasonably short, well below 1 minute.

Measurement of the films showed that practically the same precision is attained (the grain not being observable at the low magnification used), even when the rendering of color is not perfect. The mean error amounts to 9 microns, which compares, not unfavorably, with the values given on page 111 for direct observations.

The regular and irregular variations of the film during the processing can, with careful handling, be kept within this limit. We should make most of our observations photographically, as measurement can then be done at leisure, but development of the films still forms a drawback. The processing is lengthy and cumbersome and sending to a processing factory of the firm still entails too long a delay. We expect that these drawbacks will be solved in future.

Concluding Remarks

We hope to have it made clear that it is advantageous to use Fresnel diffraction for the study of the form of wave fronts in the field of technical optics. Adaptation to special cases can facilitate the production of the desired data with great accuracy. No very special parts are needed, although a good wire grating helps much to simplify the evaluation of the results.

Precision might still be bettered by making use of more than one color transition at each period of the diffraction pattern. An error in the determination of light path differences of 15A seems to be attainable.

116

10.   Image Structure and Test Data

By James G. Baker [1]

It has been the good fortune of the author over a period of 16 years to have been closely concerned with the design, construction, testing, and application of over 50 uniquely different imageforming optical systems and in fact, with the projects requiring such equipment. Many of these systems have been of substantial size and complexity, in both the visual and photographic domains. Several have employed as many as 35 surfaces in an effort to comply with requirements on field angle, aperture, resolution, and basic task.

The instruments in which these systems are mounted list among them large astronomical telescopes, visual telescopes, periscopes, theodolites, spectrographs, navigational sights, and to a very considerable extent, photographic objectives. These latter objectives have had as many as 11 elements and for astronomical purposes have focal lengths up to 320 inches. Apertures have been as large as 33 inches, and formats as large as 27 by 27 inches.

Because of this close association between calculation and the building of equipment, supplemented by observing experience in the use of large astronomical instruments, the author has been able to accumulate a fund of visual testing experience difficult to transcribe on paper. In the case of large aerial lenses the writer has been on numerous high-altitude flights for testing purposes, and has undertaken cold- and pressure-chamber tests of the same lenses. The purpose of all of this procedure was to leave as small a gap as possible between various phases of development of the optical equipment, from conception to final application.

Because of the necessary expenditure of many thousands of hours of computations in the design of these systems, it has been of the utmost importance to the author to form a reasonably consistent picture in his own mind of the many factors governing the performance of the systems for their intended purposes. The follow-through to the final application of any given system, especially where photography is involved, had to be pursued to furnish a safe basis for further work. In this way aerial flight testing as early as 1942 provided the author with first-hand experience on the conditions of aerial photography in the field. The accumulated results over a period of 5 years during World War II have led to what the author hopes are improved compromises in the instruments of the post-war period in which he has been involved.

Unsatisfactory testing conditions encountered during the manufacture of several of the large astronomical Schmidt telescopes in 1938–40 caused the author to set up testing equipment at the Harvard

College Observatory as soon as circumstances permitted. In 1941 a "testing tunnel" 22 feet long and 4 feet square was constructed. An existing 24-inch aperture f/5 astronomical mirror was mounted on a concrete pier for a collimator. This testing tunnel with its collimator was used for visual examination of a number of instruments built during the war at Harvard under the NDRC program. Photographic testing and photomicrographic enlargements were employed at infrequent intervals to supplement the extensive visual work. Photographs of star trails and planets were occasionally made.

No systematic laboratory testing by photographic means was employed at Harvard but the writer maintained close contact with the copious testing accomplished under OSRD contracts at the Massachusetts Institute of Technology, Mount Wilson Observatory, the National Bureau of Standards, the Eastman Kodak Company, and at Wright Field. Some of the Harvard lenses were tested independently at several of these places, and abundant data are available.

If one adds the extensive testing carried on in England and Canada during the war, the volume of data becomes very great indeed. The material is far from being obsolete. Any one interested in the detailed performance of a variety of optical systems would do well to have these data at hand. It would be very difficult even to reproduce the data. The author has drawn liberally on this information in the postwar years.

During the fall of 1945 a number of reports were prepared by the staff of the NDRC project at Harvard on the optical developments carried on there during the period from 1940 to 1945. In the first 3 months of 1946 a further opportunity was presented to the writer in reviewing the testing work and optical developments throughout the nation, and summarizing these as a coauthor in the NDRC volume on instruments. This period of review led to the presentation of a conception of the problems of image quality, as touching on resolution and contrast, for purposes of optical design.

It was pointed out that the lens design could be adjusted in the final balancing of aberrations to go with the level of resolution to be expected as an average in the air, and that by putting as much light energy as possible within a circle of confusion determined by the adopted level of resolution, the designer could improve the *microscopic contrast* at that level. In fact, these considerations were made use of in 1941 in the design of a 40-inch f/5 telephoto. It is very probable that the Zeiss designers were doing more or less the same thing before the turn of the century, inasmuch as lens bench practice almost inevitably requires it.

Users of lens systems often tend to look at the problem from another side, where a different focus may be found to exist for maximum resolution and for maximum contrast. However, it should be observed that the user of a finished system has far fewer variables at his disposal than the designer, and in fact is faced with simply the question of what focus to use. It is incumbent on the designer from the start of his work to control the image in matters of contrast and resolution. Consequently, it is vital that the designer understand as much as possible the details of image structure and test data for the guidance of his efforts and the success of his output. Where the design results can be supplemented by final laboratory adjustments of the finished lens, so much the better, and the old-time optical manufacturers made many test models for the purpose.

Visual Testing

It is well known that in the presence of image aberrations there may be a very distinct difference between the visual and photographic performance. For convenience, however, it is often necessary for the observer to omit the photographic process and to judge from the visual image insofar as he can what the probable photographic performance will be. In systems having very little aberration the visual evaluation can be applied with confidence. The observer looks for a sharply defined image of a source pinhole, even though the pinhole image may be embedded slightly in unfocused light. Many observers prefer to use a brilliant star point, but the writer has found that it is difficult to judge the distribution of light in the image, or the relation between the intense central point surrounded by diffraction rings and the probable photographic performance.

In a few cases where absolute perfection is required in the image within the possibilities of wave optics, a detailed analysis of the structure of the diffraction rings may prove necessary. Photographic systems do not require this kind of analysis directly. One is more interested in whether a bright but not dazzling pinhole, whose image in the focal plane measures perhaps 0.020 mm, presents a "clean" edge even in the presence of color aberrations or monochromatic aberrations. If the color correction by design differs appreciably from the visual range, one is obliged to make use of a suitable filter in which the response of eye and photographic emulsion receive proper weighting. Obviously, this process becomes more and more unreliable in outlying regions of the spectrum. If visual testing is to be employed at all under these circumstances, one must check the performance in the nearest accessible color region for comparison with the calculated design performance in the same region. Obviously, the designer of the system ordinarily prefers to do this kind of checking himself, owing to his familiarity with the aberrations and their color gradients. The success of the venture will depend greatly on the fund of experience available to the designer at the lens bench.

In·1943 a demonstration instrument was made at Harvard to show the difference between the performance of a two-element achromatic doublet, designed according to purely geometrical optics for smallest axial circle of confusion, and of a similar two-element achromatic doublet designed according to strictly physical optics. Two f/4 objectives were constructed from the same melts of optical glass and to the same test plates. The curves and thicknesses were all within critical tolerances. The only difference between the pair of doublets lay in the rear surface radius. The one differed by a calculated number of fringes from the other to span the gap between geometrical and physical correction. On the test bench the performance of the "physical optics" instrument proved to be in close agreement with the usual diffraction image structure and the image was sharply defined. The "geometrical optics" lens showed quite marked spherical aberration in the image.

Again in 1943 a "spherical-aberration analyzer" was constructed to be used at the testing tunnel for studying the image structure corresponding to varying amounts of third- and fifth-order spherical aberrations and stop diameters adjusted into the system. The two-element system was mounted in micrometer form so that it was possible to

obtain the exact design figures corresponding to any given situation of the spherical-aberration analytical terms and focal ratio. The visual information obtained at the testing tunnel was then applied in the further design work of the period. No photographic testing was accomplished with this instrument, but such testing would even now be very worthwhile.

Another brief but interesting experiment was performed in 1943 along the same lines for the purpose of information on balancing of aberrations for contrast and resolution. Even in 1941 the writer had adopted a rule in design of getting as much light from the aperture into a 20-micron circle as possible and the remaining light as far away as possible, for the purpose of maintaining contrast at a fairly critical resolution level. This rule might be recast into modern language in much more detail. The experiment performed in 1943 tended to confirm this rule adequately well. In this experiment, a well made positive simple lens with spherical surfaces was mounted as a telescope objective and used with a microscope as eyepiece. A large amount of under-corrected spherical aberration was necessarily present.

A boy was stationed about 40 feet away in a dark room with his face illuminated by a sodium-vapor lamp. His face was observed through the telescopic system with the 20X microscope mounted on a micrometer slide. Various focal settings were averaged for the reading at which the boy's face presented the most pleasing compromise between contrast and resolution.

Thereafter, a small pinhole was placed over the sodium light at the same distance and used as a new object. It was observed that as the microscope was moved from decidedly outside focus, the image of the pinhole first formed at a rather definite microscope reading. Here, the pinhole image appeared for the first time to present an edge or smooth outline, but was surrounded by a very large single flare of undercorrected primary spherical aberration.

As the microscope was moved closer to the lens in small increments, the pinhole image measuring about 0.020 mm in diameter became brighter and more sharply defined. This fact means that a larger portion of the lens aperture provided light within the Rayleigh limit with a corresponding increase in the brightness of the pinhole image and sharpness of edge. Finally, just beyond the setting for maximum sharpness of the pinhole image, a slight secondary flare began to grow from the edge of the pinhole. At this setting the Rayleigh limit was so far exceeded by an intermediate zone of the aperture as to cause the resulting caustic to add the furry edge to the pinhole image. The pinhole with the surrounding small secondary flare still lay embedded in a much greater flare of spherical aberration from the outer zone of the lens. *The micrometer reading previously ascertained for the best over-all rendition of the boy's face agreed closely with the focal setting on the pinhole showing the slight secondary caustic flare.* With the pinhole image having a diameter of about 0.020 mm, one might judge the diameter of pinhole image plus flare as about 0.030 mm. Farther inside or outside of this focal setting, one observed that the image of the boy's face grew worse, though on the outside at least the edge of the pinhole image itself was sharper.

All of this can be interpreted to mean that for a particular type of object, such as the face of a boy, best pictorial results are given by a certain degree of contrast and resolution. If the resolution were to

be emphasized, the contrast suffers. If a larger blur circle is used to bring a higher percentage of the light into the expanding circle, the resolution may fall as the contrast increases. *The picture "quality" appears then to be a proper balance between contrast and resolution for a given type of object and scale of object in the focal plane, both depending on the quality of the optical instrument.* A fairly high contrast is required for large and small detail to obtain a pleasing pictorial effect, but there must be enough resolution too for defining what is sought in the picture.

The visual testing of the many different optical systems developed at Harvard consisted mostly of examining the image of a pinhole formed from the 24-inch mirror collimator and of such a size at the focus of the collimator to give a 0.020 mm image in the focal plane of the lens under test. The microscope power usually ranged from 50X to 100X with no attempt made to increase the magnification much beyond the merit of the image. Image characteristics were judged from the appearance of the pinhole image at varying off-axis angles and with different filters. Quite frequently, three-line test patterns were examined for guidance, but most judgments as to the residual defects of the lens performance were derived from pinhole images. Differential corrections to airspaces and occasionally to radii were determined from the pinhole images, via design calculations, and applied to the lens with subsequent rechecking. By this process the best compromises were sought and achieved in the final lens.

It has been the author's experience and perhaps bias that visual testing can be much more exhaustive of a system's ultimate performance than photographic testing in the focal plane. No doubt, photoelectric or photomicrographic methods can be substituted for visual testing in order for one to obtain quantitative data. However, in the focal plane itself use of the turbid photographic emulsion spoils the optical image to such a degree that the subtle aspects of the image structure are lost. Photographic testing in the focal plane for photographic objectives is, nevertheless, highly desirable on the basis that acceptance tests on performance ought to be based on what the instrument is supposed to do. If a photographic objective provides nearly perfect visual images, it may be over-designed and too complicated. If an objective is afflicted by many hybrid aberrations, the observer may find it extremely difficult from a visual examination to judge the photographic performance. The average lens of good quality lies between these two extremes.

Visual testing on a lens bench or in a testing tunnel is often necessary for the designer to determine whether his lens is built to his specifications and whether further improvement can be achieved. Ideally, a design ought to go through a prototype stage and at least one differential correction if necessary before the final optimum performance can be determined. With the introduction of modern high-speed computing and with a better quantitative knowledge of image structure, designers may be able before long to eliminate need for a differential correction on any pronounced scale. As it is, an individual with considerable design and testing experience can minimize the gap between design and performance by judgment alone.

The author feels that it is relatively unproductive for a new type of optical system to receive exhaustive testing before the designer himself has done all he wishes on improving the system. It is easy

to make rather small adjustments that will change the apparent structure of the image very appreciably. The designer can look at the image and read into it how it would look if he made certain alterations in the design. The nondesigner can only accept it for what it is, and may spend too long a time reaching conclusions that may be invalid for the next lens off the production line.

Close attention to detail near the end of fabrication of a lens system may result in a gain in performance by a factor of two or more. Such an adjustment may, for example, convert an effective image 0.020 mm in diameter into one 0.010 mm in diameter without any loss of contrast. There may be compromises possible that will bring about an average high-level performance over the field instead of one excessively high on axis and poor near the edge of the field.

The balancing for maximum contrast at a desired resolution level is a case in point. The designer may be tempted to demonstrate a high visual or even photographic resolution, whereas better pictures might be obtained at a more conservative resolution level by improving the microscopic contrast.

In aerial photography pictures having a resolution of about 10 lines/mm, with good contrast at 5 lines/mm or coarser, appear to be of excellent quality to the uncritical observer. If such a performance is taken as the norm, then one overlooks a whole realm of complex considerations in image structure at a level of 30 lines/mm or finer. Visual study of the subtleties of image structure beyond 30 lines/mm becomes then relatively unimportant. For quality work, however, the question of whether a system has received adequate final differential correction and adjustment is still vital. Otherwise, an elaborate and expensive system inadequately made may perform no better on the average than a simpler system carefully adjusted.

Image Structure

An optical image formed well off-axis in a large fast lens is a hybrid structure resulting from the contributions of many types of image errors or aberrations in varying degrees according to the design. Figure 10.1 gives a schematic array of the many conditions at work, each of which in turn requires a number of power series terms for analytical representation. Moreover, the fifth order indicated is still quite insufficient to give an accurate description of the image errors that occur in such a lens as a 12-inch f/2.5 for 9 by 9 format, and one must draw either on still higher order terms with slow convergence or abandon series methods altogether. In such cases a useful procedure is to make use of the first, third and fifth order for analytical control but to analyze image structure by exact ray-tracing.

In some of the systems designed at Harvard all terms of the first, third, and fifth orders given in figure 10.1 have been satisfied with the exception of a residual oblique spherical aberration and secondary spectrum. In others, the latter two aberrations have been controlled as well. An aberration is spoken of as being satisfied when its contribution to image structure taken together with any other balancing residuals falls within a tolerance based on performance requirements. The coefficient of such an error is necessarily small or close to zero if a fast lens is involved, a fact that often implies a zero in the solution, and some control of the aberration by the designer.

Magnification (lateral), E. F. L. (One Condition)

Gaussian Focus (One Condition)

Primary Spectrum in Lateral Color, or Chromatic Difference of Magnification. (One Condition)

Secondary Spectrum in Lateral Color, or Secondary Variation of Magnification with Color. (One Condition)

Primary Distortion, or Third Order Distortion. (One Condition)

Chromatic Variation of Distortion, or Variation of Distortion with Color. (One Condition)

Fifth Order Distortion or Secondary Distortion. (One Condition)

Primary Coma, or Third Order Coma, or Zonal Difference of Magnification. (One Condition)

Chromatic Variation of Primary Coma, or Variation of Third Order Coma with Color. (One Condition)

Fifth Order Oblique Coma, or Variation of Third Order Coma with Angle. (Two Conditions)

Fifth Order Sine Theorem Coma, or Secondary Coma. (One Condition)

Primary Spherical Aberration, or Third Order Spherical Aberration. (One Condition)

Fifth Order Spherical Aberration, or Secondary Spherical Aberration. (One Condition)

Oblique Spherical Aberration, or Variation of Spherical Aberration with Angle. (Two Conditions)

Chromatic Spherical Aberration, or Variation of Third Order Spherical Aberration with Color. (One Condition)

Primary Astigmatism in the Tangential Direction, or Third Order Tangential Astigmatism. (One Condition)

Fifth Order Tangential Astigmatism, or Secondary Tangential Astigmatism. (One Condition)

Variation of Third Order Tangential Astigmatism with Color. (One Condition)

Primary Astigmatism in the Radial Direction, or Third Order Radial Astigmatism. (One Condition)

Fifth Order Radial Astigmatism, or Secondary Radial Astigmatism. (One Condition)

Variation of Third Order Radial Astigmatism with Color. (One Condition)

Primary Spectrum in Logitudinal Color or Chromatic Aberration. (One Condition)

Secondary Spectrum in Longitudinal Color (One Condition)

Figure 10.1. *A schematic array of the 25 conditions affecting lens performance through the fifth order of approximation.*

Fifth-order field curvature and astigmatism can often be controlled in spite of the complexity of the effort. The asymmetrical terms of the fifth order come out quite often to be small or negligible. In a few cases it has proved possible to eliminate the oblique spherical aberration of the fifth order and leave residuals of the seventh or higher orders. In all cases the residual higher order aberrations have been balanced as well as judgment allows, taking diffraction image formation and vignetting into account, and weighted according to the requirements of the problem.

Figure 10.2 shows a typical array of power series terms for the third, fifth and a portion of the seventh orders, evaluated for the case of a 40-inch f/5 telephoto lens for 9 by 9 format. The coefficients are in millimeters when the normalized field-angle θ is equal to 1 in the

Δy	Δz
Third Order	
Distortion $\quad$ 0.186 θ^3	
Astigmatism $\quad$ $-0.001\ \theta^2 y$	Astigmatism $\quad$ $-0.000\ \theta^2 z$
Coma $\quad$ $-0.010\ \theta y^2$ $\quad$ $-0.003\ \theta z^2$	Coma $\quad$ $-0.007\ \theta yz$
Spherical $\quad$ $-0.061\ y^3$ Aberration $\quad$ $-0.061\ yz^2$	Spherical $\quad$ $-0.061\ z^3$ Aberration $\quad$ $-0.061\ y^2 z$
Fifth Order	
Distortion $\quad (-0.136 + 0.00079\ c_N)\ \theta^5$	
Astigmatism $\quad (\ 0.334 + 0.00187\ c_N)\ \theta^4 y$	Astigmatism $\quad (\ 0.111 + 0.00037\ c_N)\ \theta^4 z$
Coma $\quad (\ 0.495 + 0.00178\ c_N)\ \theta^3 y^2$ $\quad (\ 0.166 + 0.00036\ c_N)\ \theta^3 z^2$	Coma $\quad (\ 0.332 + 0.00071\ c_N)\ \theta^3 yz$
Oblique $\quad (\ 0.084 + 0.00084\ c_N)\ \theta^2 y^3$ Spherical $\quad (\ 0.194 + 0.00050\ c_N)\ \theta^2 yz^2$	Oblique $\quad (\ 0.194 + 0.00050\ c_N)\ \theta^2 y^2 z$ Spherical $\quad (\ 0.304 + 0.00017\ c_N)\ \theta^2 z^3$
Sine Coma $\quad (\ 0.004 + 0.00020\ c_N)\ \theta y^4$ $\quad (\ 0.005 + 0.00024\ c_N)\ \theta y^2 z^2$ $\quad (\ 0.001 + 0.00004\ c_N)\ \theta z^4$	Sine Coma $\quad (\ 0.003 + 0.00016\ c_N)\ \theta yz^3$ $\quad (\ 0.003 + 0.00016\ c_N)\ \theta y^3 z$
Fifth Order $\quad (\ 0.079 + 0.00002\ c_N)\ y^5$ Spherical $\quad (\ 0.158 + 0.00004\ c_N)\ y^3 z^2$ $\quad (\ 0.079 + 0.00002\ c_N)\ yz^4$	Fifth Order $\quad (\ 0.079 + 0.00002\ c_N)\ z^5$ Spherical $\quad (\ 0.158 + 0.00004\ c_N)\ y^2 z^3$ $\quad (\ 0.079 + 0.00002\ c_N)\ y^4 z$
Seventh Order	
Distortion $\quad$ $-0.023\ \theta^7$	
Astigmatism $\quad$ $0.070\ \theta^6 y$	Astigmatism $\quad$ $0.022\ \theta^6 z$
Coma $\quad$ $0.054\ \theta y^6$	
Spherical $\quad$ $-0.015\ y^7$ Aberration	

FIGURE 10.2. *A typical array of coefficients of third, fifth and seventh orders of approximation.*

corner of the format, and where the maximum values of y or z are equal to 1 in the entrance pupil. The quantities Δy and Δz refer to the intercept of a ray, defined by θ, y and z in the entrance pupil, in the focal plane in millimeters relative to the Gaussian point of zero distortion. The quantity c_N refers to an aspheric coefficient of the last surface of the system. The series was evaluated before reintroducing first- and third-order residuals for balancing of aberrations. Figure 10.3 shows the image form calculated from the series of figure 10.2, but with balancing included. The numerical quantities are in millimeters and indicate that the effective image is of the order of 0.020 mm in diameter. The image structure in figure 10.3 is purely geometrical optics, and by physical optics the image is very nearly the usual spurious disk surrounded by diffraction rings with very little extra light in the rings. The vertical elongation of the image in the skew direction was indeed observed in the lenses built, but, neverthe-
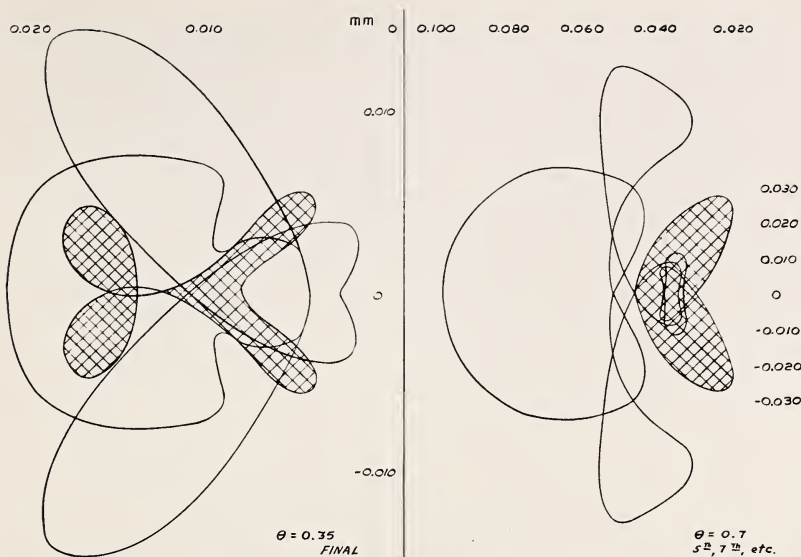
FIGURE 10.3. *Enlarged geometrical image calculated from series for a 40-inch f/5 telephoto lens.*

less, resolution on Super-XX emulsion at high contrast reached a level of 45 lines/mm.

In the meridional plane, owing to maximum deviations of the rays at the rim of any given surface, the aberrational elongations in the image can easily dominate over the skew aberrations. Thus, comatic flares and oblique spherical aberration usually cause a faster deterioration of the image in the meridional plane than do the skew terms. However, owing to the greater ease of designing in the meridional plane, one often finds near the end of his work that the meridional elongations in the image have been brought nearly to zero, but that the image squeezes out still in the vertical direction, above and below the meridional plane. Figure 10.3 indicates that this residual defect occurred in the case of the 40-inch f/5 lens, but other examples at hand show the effect even more strongly. With the advent of faster computing methods by automatic machine, it is probable that the skew direction will come in for more evaluation during the course of a design with improved chances for better image structure in all azimuths.

Much more complicated off-axis images have been encountered than can be shown here. Variations with color have to be considered, and if present in appreciable amount, the effect on the location of focus and image structure must be dealt with during the course of the design. The final balancing of aberrations must also involve contrast-resolution factors, aperture, field, and color errors, weighted according to requirements. It is clear that a designer must be as familiar as possible with the behavior of systems of known corrections in order to determine in advance just what he must do for optimum results. It is probable that German designers of 50 years ago had a very good idea of image structure and balancing, but kept such information only as part of their experience or as company data. Nowadays we may be plowing up much old ground, but this time the growing
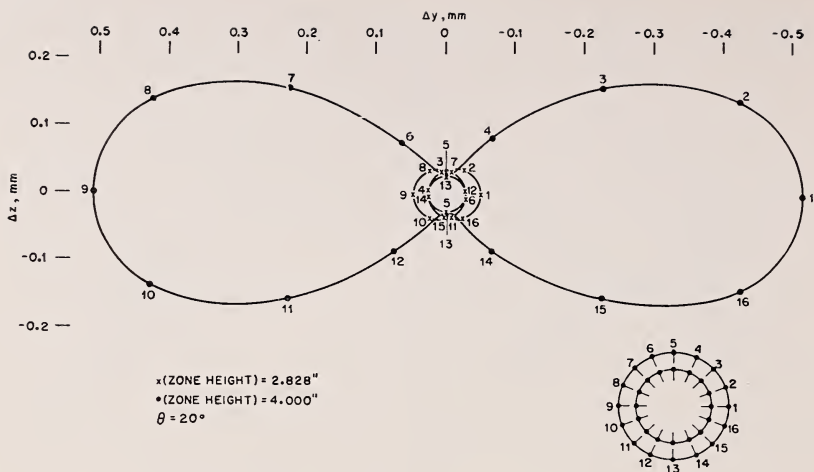
125

Figure 10.4.   *Image errors of an f/1 Schmidt camera.*

x(ZONE HEIGHT) = 2.828"
•(ZONE HEIGHT) = 4.000"
θ = 20°

knowledge of physical optics, photographic effects, psychological factors and the like are undoubtedly going to lead to scientific results available to everyone.

Figure 10.4 shows the oblique errors of a standard Schmidt system as calculated from series and verified by selected rays. The aberration is almost entirely oblique spherical aberration caused by foreshortening of the oblique beam and by the increase in the refractive strength of the plate for inclined rays. Figure 10.4 shows that the unvignetted off-axis image errors of an 8-inch f/1 Schmidt camera are by no means inappreciable. Star photographs made with a 12-inch f/1 Schmidt demonstrate the radial elongation of the images very clearly.

During the summer of 1940 the author made a prolonged attempt to improve the Schmidt by introducing more correcting plates. Figure 10.5 shows an array of terms for the various aberrations of the third and fifth orders for the generalized system of the same figure. Algebraically, the equations hold for a solid glass type of Schmidt system with air-lens aspheric surfaces that may have finite bendings. Aspheric terms on the mirror are also included. One needs only to substitute proper values for the indices of refraction in order to obtain valid equations for a mirror in air, combined with three thin correcting plates that may have finite bendings.

The equations in figure 10.5 show many relationships of interest with respect to the fifth order, and demonstrate how very lengthy an explicit solution through the fifth order for even a Cooke triplet would be. The Schmidt system is about as tractable as any that can be handled in this way. No attempt has been made to find a solution to the analytical expressions, for in order to improve on the ordinary Schmidt, one must have all of the expressions zero simultaneously. Physical considerations can circumvent the use of such complicated expressions, and indeed in 1943 the author came across a rather simple solution to the problem in the form of the "double shell" Schmidt system with a correcting plate of higher order asphericity.

Such a system was developed for television purposes in December, 1943 for the Perkin-Elmer Corp. Shell systems were designed independently elsewhere, as information coming to hand at the end of the war indicated, with some as early as 1940. In the spring of 1946 the double-shell system was further elaborated at Harvard into the Super-Schmidt Meteor Cameras to include an achromatized correcting plate and arranged to have the light coming through the second shell twice with a favorable effect on the image quality.

These "Super-Schmidt Meteor Cameras" now in successful operation cover a field of 52 degrees on specially curved film, have a speed of f/0.65, a color range from 3700-5000 and a clear aperture of 12.2 inches for an 8-inch focal length. The effective star image appears to be of the order of 0.025 mm in diameter over most of the field. The extremely low f/number causes focusing to be extremely critical to the extent that a focal error of 0.010 mm already shows to disadvantage. The secondary spectrum of this system is negligible by ordinary standards, but the focal tolerance here is so critical that the image is enlarged appreciably by the secondary spectrum of the correcting plate and shell combination. There appears to be no easy solution to the defect caused by the secondary spectrum, inasmuch as the requisite crystalline materials or special glasses are either not available or would have achromatizing curves too strong to be useful.

The observed performance of the Super-Schmidt Meteor Cameras is all the more striking, because the central 7.4 inches diameter of the 12.2-inch aperture is blocked out by the film and film-holder. The sharp images are thus formed from an annular entrance pupil, a fact that causes the focal tolerance to be all the more critical. This outer annulus is the region of the worst aberrations, and the core of the image normally formed by the inner zones is missing. Nevertheless, the color error remains the limiting defect in the center of the field, and only slight comatic wedges defined by the vignetting characteristics appear in the outer field.

Figure 10.6 shows a cross-sectional cut-away of a 36-inch f/8 apochromatic lens for a 9 by 9 format, designed in 1943. The apochromatic correction was achieved through a combination of a fluorite element with light crown elements elsewhere. This fluorite element was cemented between two glass elements in the central negative component. The design of this apochromatic system was particularly interesting in view of the first-time use of a large element of fluorite in a lens system. The standards of the design had to be set at a high level of performance. There is no gain in eliminating secondary spectrum if the monochromatic aberrations are made larger thereby. The true apochromat is corrected to have a common focus for three colors and is also corrected for spherical aberration and coma for two colors. This aerial lens also had to have flat field, freedom from astigmatism, and distortion, and had to have an off-axis image sensibly free of oblique spherical aberration.

Figure 10.7 shows the residual meridional aberrations. The inner part of the field shows some outward coma, which at ¾ field has practically disappeared with the up-swing at the right vignetted away. The very corner image comes from the portion of the bottom curve between the vertical lines, the rest being vignetted away, and is evidently astigmatic. Figure 10.8 shows the effective field curves. The tangential astigmatism causes trouble only in the very corner.

Figure 10.5.   *Third- and fifth-order conditions for a multiple Schmidt camera.*

Series Expansion for Solid Glass Schmidt with 3 Aspheric Correcting Air Lenses of Finite Bendings and Aspheric Mirror.

$$x_i = a_i(y^2 + z^2) + b_i(y^2 + z^2)^2 + c_i(y^2 + z^2)^3 + \cdots$$

(Stop at d_0)

$$d_0 = d_1 = d_3 = d_5 = 0$$
$$d_2 = d_4 = d_6 = -d_7 = \frac{1}{2}$$
$$a_1 = a_2 \quad a_3 = a_4 \quad a_5 = a_6$$
$$B_1 = -2\frac{1}{n}(n-1)(b_2 - b_1)$$
$$B_3 = -4\frac{1}{n}(n-1)(b_4 - b_3) - 1$$
$$B_5 = -2\frac{1}{n}(n-1)(b_6 - b_5)$$
$$C_1 = -6\frac{1}{n}(n-1)(c_2 - c_1)$$
$$C_3 = -8\frac{1}{3}\frac{1}{n}(n-1)(c_4 - c_3) - 1$$
$$C_5 = -6\frac{1}{n}(n-1)(c_6 - c_5) - 1$$
$$+ \frac{3}{64}(n_3 - 1)a_3^2(1 + B_3)$$
$$+ 3(n_5 - 1)B_5 a_5^2$$
$$+ \left(\frac{33}{16} + \frac{1}{2}n_5\right)B_5 - \frac{27}{16}S_7 - \frac{1}{8}S_8$$
$$- 27 B_5 S_7$$
$$+ \frac{9}{256}C_3$$

$$y_8 =$$

$$[\mu_0] = \frac{1}{2n}$$

$$[y_0] = 0$$

$$[\mu_0^3] = \frac{1}{n^3}\left[\frac{1}{16}B_3 + B_5 - \frac{27}{2}S_7\right]$$

$$[\mu_0^2 y_0] = \frac{1}{n^2}\left[\frac{3}{8}B_3 + 3B_5 - 27 S_7\right]$$

$$[\nu_0^2 y_0] = \frac{1}{n^2}\left[\frac{1}{8}B_3 + B_5 - 9 S_7\right]$$

$$[\mu_0 y_0^2] = \frac{1}{n}\left[\frac{3}{4}B_3 + 3B_5 - 18 S_7\right]$$

$$[y_0^3] = \left[B_1 + \frac{1}{2}B_3 + B_5 - 4 S_7\right]$$

$$[\mu_0^5] = \frac{1}{n^5}\left[\frac{1}{8} + \frac{1}{32}n_3\right.$$
$$+ \frac{5}{64}(n_3 + 2)a_3(1 + B_3)$$
$$+ \frac{5}{2}(n_5 + 2)B_5 a_5$$
$$+ \left(\frac{3}{32} + \frac{1}{32}n_3\right)B_3$$
$$+ \frac{3}{16}B_3 B_5 - \frac{27}{8}B_3 S_7$$
$$\left. + \frac{1}{2}C_5 - \frac{243}{32}T_7\right]$$

$$[\mu_0^4 y_0] = \frac{1}{n^4}\Bigg[+ \frac{3}{4} + \frac{3}{16}n$$

$$+ \frac{5}{8}(n+2)\,a_3\,(1+B_3) \qquad + \frac{15}{32}(n-1)a_3^{\,2}\,(1+B_3)$$

$$+ 10(n+2)\,B_5a_5 \qquad + 15(n-1)\,B_5a_5^{\,2}$$

$$+ \left(\frac{3}{8}+\frac{3}{16}n\right)B_3 \qquad + \left(\frac{57}{8}+\frac{3}{2}n\right)B_5 \quad - \frac{63}{2}S_7 \; - \frac{1}{8}S_8$$

$$+ \frac{3}{2}B_3B_5 \quad - \frac{99}{4}B_3\,S_7 \qquad - 117\,B_5\,S_7 \qquad + \frac{45}{128}C_3$$

$$+ \frac{5}{2}C_5 \quad - \frac{405}{16}T_7\Bigg]$$

$$[v_0^4 y_0] = \frac{1}{n^4}\Bigg[+ \frac{1}{8} + \frac{1}{16}n$$

$$+ \frac{1}{8}(n+2)\,a_3\,(1+B_3) \qquad + \frac{3}{32}(n-1)\,a_3^{\,2}\,(1+B_3)$$

$$+ 2(n+2)\,B_5a_5 \qquad + 3(n-1)\,B_5a_5^{\,2}$$

$$+ \left(\frac{1}{8}+\frac{1}{16}n\right)B_3 \qquad + \left(\frac{15}{8}+\frac{1}{2}n\right)B_5 \quad - \frac{21}{2}S_7 \; - \frac{1}{8}S_8$$

$$+ \frac{1}{4}B_3B_5 \quad - \frac{15}{4}B_3\,S_7 \qquad - 21\,B_5\,S_7 \qquad + \frac{9}{128}C_3$$

$$+ \frac{1}{2}C_5 \quad - \frac{81}{16}T_7\Bigg] .$$

$$[\mu_0^3 y_0^2] = \frac{1}{n^3}\Big[\; \frac{3}{2} + \frac{3}{8}n$$

$$+ \frac{15}{8}(n+2)(1+B_3) \quad + \frac{15}{8}(n-1)\,a_3^{\,2}(1+B_3) \quad + 15\,(n+2)B_5a_5$$

$$+ 30\,(n-1)\,B_5a_5^{\,2} \quad + \left(\frac{3}{8}+\frac{3}{8}n\right)B_3 \quad + \left(\frac{213}{16}+\frac{3}{2}n\right)B_5 \quad - \frac{273}{2}\,S_7$$

$$+ \frac{75}{16}B_3B_5 \;-\; 69\,B_3\,S_7 \quad + \frac{45}{32}\,C_3 \quad + 5\,C_5 \quad - \frac{135}{4}\,T_7\;\Big]$$

$$[\mu_0^3 z_0^2] = \frac{1}{n^3}\Big[\; \frac{1}{4} + \frac{1}{4} + \frac{1}{8}n$$

$$+ \frac{3}{8}(n+2)\,a_3(1+B_3) \quad + \frac{3}{8}(n-1)\,a_3^{\,2}(1+B_3) \quad + 3\,(n+2)\,B_5a_5$$

$$+ 6\,(n-1)\,B_5a_5^{\,2} \quad + \left(\frac{1}{8}+\frac{1}{8}n\right)B_3 \quad + \left(\frac{47}{16}+\frac{1}{2}n\right)B_5 \quad - \frac{61}{2}\,S_7$$

$$+ \frac{17}{16}B_3B_5 \;-\; 17\,B_3\,S_7 \quad + \frac{9}{32}\,C_3 \quad + C_5 \quad - \frac{27}{4}\,T_7\;\Big]$$

$$[\mu_0^2 y_0^3] = \frac{1}{n^2}\Big[\; 1 + 1 + \frac{1}{4}n$$

$$+ \frac{5}{2}(n+2)\,a_3(1+B_3) \quad + \frac{15}{4}(n-1)\,a_3^{\,2}(1+B_3) \quad + 10\,(n+2)\,B_5a_5$$

$$+ 30\,(n-1)\,B_5a_5^{\,2} + \frac{1}{2}n_1 B_1 + \quad \frac{1}{4}n\,B_3 \quad + \left(\frac{33}{2}+\frac{1}{2}n\right)B_5 \quad - 196\,S_7$$

$$+ \frac{3}{8}B_1B_3 + 6\,B_1B_5 \;-\; 81\,B_1\,S_7$$

$$+ \frac{57}{8}B_3B_5 \;-\; 90\,B_3\,S_7 \quad + \frac{45}{16}\,C_3 \quad + 5\,C_5 \quad - \frac{45}{2}\,T_7\;\Big]$$

$$[v_o{}^2 y_o{}^3] = \frac{1}{n^2}\left[+\frac{1}{4}n \quad +\frac{1}{2}(n+2)a_3(1+B_3) \quad +\frac{3}{4}(n-1)a_3{}^2(1+B_3) \quad +2(n+2)B_5a_5 \right.$$

$$+6(n-1)B_5a_5{}^2 + \tfrac{1}{2}n_1B_1 + \quad \tfrac{1}{4}nB_3 \quad +(\tfrac{7}{2}+\tfrac{1}{2}n)B_5 \quad -44S_7$$

$$+\tfrac{1}{8}B_1B_3 + 2B_1B_5 \; - 27B_1S_7$$

$$\left. +\tfrac{11}{8}B_3B_5 \quad -18B_3S_7 \quad -33B_5S_7 \quad +\tfrac{9}{16}C_3 \quad +C_5 \quad -\tfrac{9}{2}T_7 \right]$$

$$[\mu_o y_o{}^4] = \frac{1}{n}\left[+\frac{5}{2}(n+2)B_1a_1 \quad +\frac{5}{4}(n+2)a_3(1+B_3) \quad +\frac{15}{4}(n-1)a_3{}^2(1+B_3) \quad +\frac{5}{2}(n+2)B_5a_5 \right.$$

$$+15(n-1)B_5a_5{}^2 \qquad\qquad +\tfrac{45}{4}B_5 \quad -115S_7$$

$$+\tfrac{3}{2}B_1B_3 + 12B_1B_5 \; - 108B_1S_7$$

$$\left. +\tfrac{21}{4}B_3B_5 \quad -54B_3S_7 \quad -72B_5S_7 \quad +\tfrac{45}{16}C_3 \quad +\tfrac{5}{2}C_5 \quad -\tfrac{15}{2}T_7 \right]$$

$$[y_o{}^5] = \left[+3(n-1)B_1a_1{}^2 \qquad\qquad +\frac{3}{2}(n-1)a_3{}^2(1+B_3) \right.$$

$$+3(n-1)B_5a_5{}^2 \qquad\qquad +3B_5 \quad -24S_7$$

$$+\tfrac{3}{2}B_1B_3 + 6B_1B_5 \; - 36B_1S_7$$

$$\left. +\tfrac{3}{2}B_3B_5 \quad -12B_3S_7 \quad -12B_5S_7 + \tfrac{1}{2}C_1 \; +\tfrac{9}{8}C_3 \quad +\tfrac{1}{2}C_5 \quad - T_7 \right]$$

131

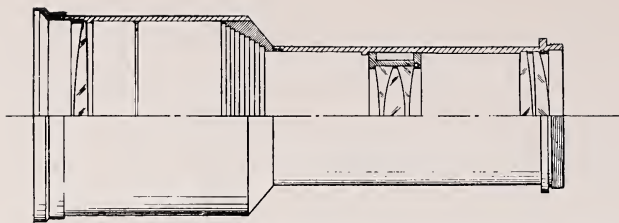FIGURE 10.6.   A cut-away view of a 36-inch f/8 apochromatic lens in which a single small fluorite element is used.
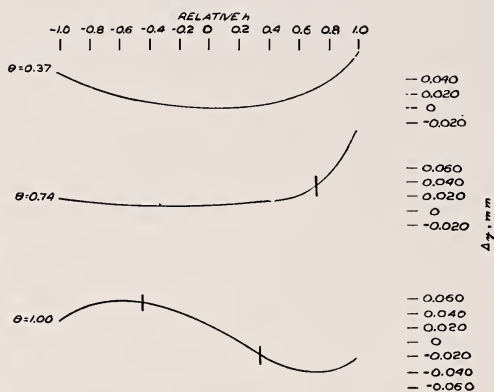

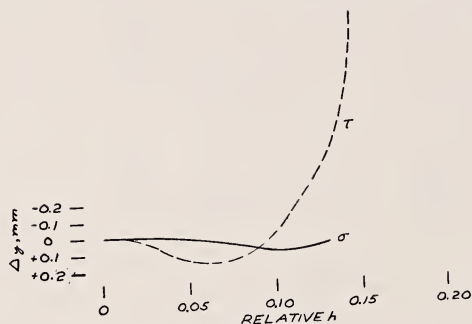FIGURE 10.7.   The aberrations of the 36-inch f/8 apochromat in the meridional plane.

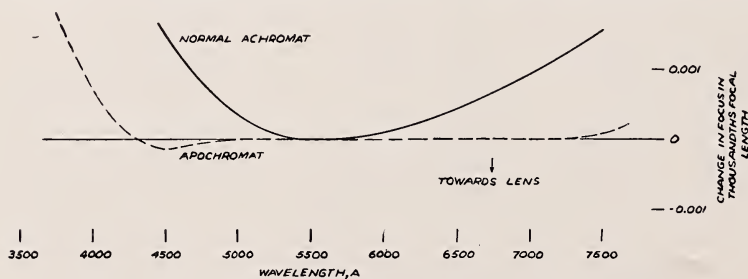FIGURE 10.8.   The radial and tangential image surfaces of the 36-inch f/8 apochromat.

FIGURE 10.9.   The color curve of the 36-inch f/8 apochromat.

The performance of this system turned out to be quite satisfactory even for astronomical purposes over a 9-inch diameter circle of the field. The apochromatic correction was fully verified by independent tests at Mount Wilson, and at Eastman Kodak Co. Figure 10.9 shows the color curve of the system as evaluated on the optical bench at Mount Wilson. Diffraction came to the aid of the residual image defects shown in figure 10.7 and figure 10.8, so that the star images remained sharp even to photomicrographic magnifications up to 80X.

Figure 10.10 shows the color curve of a 100-inch f/10 astronomical lens of four elements, designed for optimum performance in red light. It can be seen that the color curve, though normal for an achromat, is very pronounced in the yellow, green and blue, owing to the location of the minimum in red light. This lens was actually built and tested thoroughly. The visual image without filter showed a striking color flare around a red "core", enhanced by the sensitivity of the eye to green light. However, the photographic image with yellow filter showed almost no effect of the color curve on Super-XX film. While color pictures taken with such a lens would no doubt be deficient, a lens of this kind seems entirely adequate for black and white pictures on panchromatic emulsion with yellow or red filter. The tests in the laboratory proved that the sharp core of the image in orange and red light accomplished more than the large flare in yellow and green light detracted. Star images photographed in red light appear fully round and sharp over a 14 by 14 photographic plate.

Many of the systems designed by the author have received similar tests. The astronomical systems are in use and much data are at hand on their performance. However, this report need not be burdened with more information of this kind, information that might better appear in the form of a manual or compendium of results. It is necessary here only to show that the large variety of optical images designed, produced and studied by the writer have built up a certain fairly consistent set of experience that make image interpretation reliable and further design work more precise.

As a general rule, the lower order image errors must be satisfied properly in order to draw on the liberal tolerances of the higher order errors. In this sense, one must expand the image errors in terms of the aperture, rather than of the field. Thus, fifth-order astigmatism must be regarded as a first-order error in the aperture. The aberration is just as important in the outer field as the simple focal error is on axis. At each point in the field, then, the image quality must be
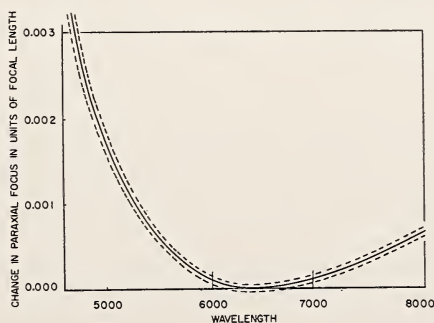

FIGURE 10.10. *Color curve of a 100-inch f/10 Ross lens for red light.*

studied in powers of the aperture. The focal errors such as curvature of field and astigmatism must be satisfied first. Next, the variations of these errors with color must be satisfied, together with lateral color, which is of first-order importance. Next come coma-like errors, varying as the square of the aperture. Thereafter come errors of spherical aberration, easy to control on axis but very difficult to control at far off-axis angles.

Lens-Film Performance

Figure 10.11 is reproduced through the courtesy of the Eastman Kodak Company. The several characteristics of Super-XX and Pana-tomic-X roll films are shown all on one type of graph. The target contrast is given as the log of the ratio of high light to low light intensity. It will be seen that at a contrast level of 2:1, the resolving power is still as good as 25 lines/mm. Figure 10.11 also shows how the resolving power determined from conventional three-line patterns is a function of exposure. The peak resolving power occurs at a density of about 1.7 for Super-XX and about 1.4 for Panatomic-X for high target contrast. It is significant that the maximum resolving power
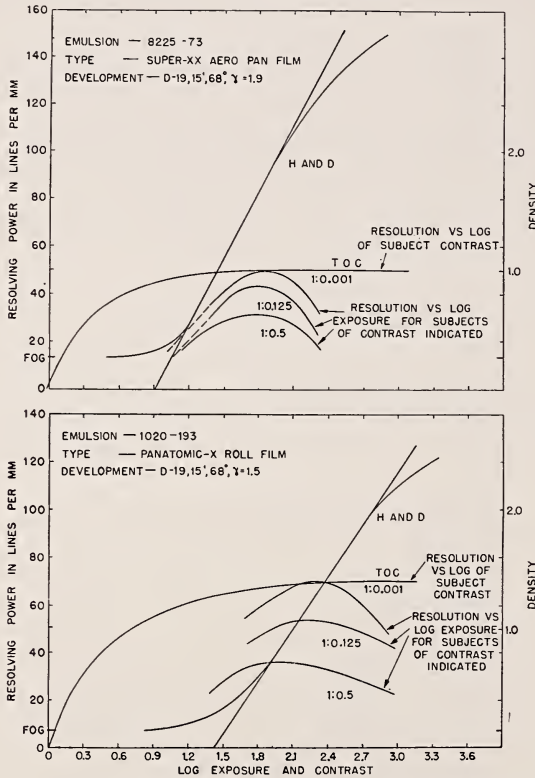

FIGURE 10.11. *Photographic properties of Super-XX and Panatomic-X aerial film.*
(Courtesy of the Eastman Kodak Company).

for lower contrast shifts to lower density levels for both emulsions, an effect more pronounced for Pan-X than for Super-XX.

In view of the nature of the curves in figure 10.11, combined with experimental microphotometer tracings made on test patterns, the author was led to propose late in 1945 a type of testing involving the concepts of *macroscopic* and *microscopic* contrasts. The term macroscopic contrast refers to the contrast of large areas on the emulsion. The term microscopic contrast refers to the measured contrast at a given line number in a sequence of decreasing line widths and separations. Microscopic contrast is therefore a function of line number and goes into macroscopic contrast in the limit for large areas. Microscopic contrast goes to zero at the limit of resolution of the photograph, if measured in terms of density difference.

It is evident that the visibility of resolving-power lines depends, to a large extent, on density difference between lines and spaces on the test emulsion. As far as the emulsion is concerned, it does not matter whether the reduction in density difference in the finer lines is caused by aberrations of the lens system, by general fogging, or by turbidity within the emulsion. The final microscopic contrast of a given lens-film combination will depend functionally on the intrinsic characteristics of the emulsion, the target contrast, the lens quality, the exposure, development, and means for measuring the contrast.

For each value of target contrast one can plot a curve of microscopic contrast, as measured in terms of density difference, against line number from macroscopic contrast down to zero, in any given test of lens-film under standardized conditions. The resultant curves are then to be related to performance characteristics that can be considered of direct importance to the ability of the combination to reveal detail and contrast.

Such a family of curves can be plotted for a "perfect lens-film combination", and then for the test lens-film combination, all other conditions being held constant. A good test lens will have the microscopic contrast as close as possible to the ideal limiting curve all the way down the progression of line number. One can interpret from such curves the performance of the lens under varying conditions of haze, and target contrast.

While this procedure was described in 1945, no one has carried it out systematically so far as the author is aware. However, the new resolution-contrast test pattern of the Bureau of Standards accomplishes essentially the same thing, except for differences produced by the use of long lines. The differences may be marked where the aberrations are large. The Bureau of Standards results may indicate higher resolution values over tests conducted with the standard three-line patterns.

If we view the limiting resolution from the standpoint of the emulsion, we can relate the target contrast needed for the imperfect lens-film to attain the stated resolution, to the target contrast needed for the perfect lens-film to attain the same resolution. In this way the loss of picture quality can be expressed in terms of the *microscopic equivalent target contrast*. Thus, at a given resolution limit the inferior lens will require more contrast in the target for resolution than will the perfect lens, and the equivalent target contrast will be a measure of the loss of quality.

TABLE 10.1.

Macroscopic ground target	Super-XX	Pan-X	Ground target contrast	Haze	Microscopic equivalent target contrast
High contrast					
	Lines/mm	*Lines/mm*		%	
Perfect lens_____ {	50	70	1/0.001	None	1/0.001
	34	43	1/0.001	50	1/0.33
Imperfect lens_____ {	26	32	1/0.001	None	1/0.50
	23	27	1/0.001	50	1/0.60
Low contrast					
Perfect lens_____ {	26	32	1/0.50	None	1/0.50
	20	23	1/0.50	50	1/0.67
Imperfect lens_____ {	16	17	1/0.50	None	1/0.74
	14	14	1/0.50	50	1/0.80

Table 10.1 gives a tabulation of contrast values worked out from typical results of aerial photography. It is evident that the poor lens gives lowered microscopic contrast, which is the equivalent of added haze for the perfect lens. Increased gamma of development may restore the contrast in the coarser patterns but cannot affect the finer patterns appreciably. Excessive macroscopic contrast produced by prolonged development for the purpose of increasing the microscopic contrast will then distort the tonal values of the photograph.

It is also evident from table 10.1 that the perfect lens has its performance lowered more rapidly by the addition of haze than does the poor lens, though at all times the perfect lens stays systematically superior to the poor lens. Another way of stating the case is to say that the law of diminishing returns sets in when the image quality is perfected beyond a certain point, and that one pays dearly for a slight increase in contrast and resolution. Much aerial photography occurs at low contrast where the superiority of the perfect lens is less clearly defined. All of such troubles are caused essentially by the diffusing nature of the emulsion.

Table 10.1 proves that an increase in focal length is more important than an increase in lens performance, where we can assume a certain reasonable quality to every professional lens system. If haze is so bad that even a perfect lens-film will yield only 16 lines/mm, say, the imperfect lens-film of average performance may still resolve 12 lines/mm. If the focal length of the imperfect lens is increased by 33 percent, the resulting performance referred to the target will be approximately identical to that of the smaller perfect lens, inasmuch as the haze factor is independent of focal length. In aerial photography at a given altitude under such bad haze conditions, either one can use a greater focal length to achieve ground resolution or else fly lower with the same camera, all to achieve a larger scale that in the presence of haze may permit the desired object to be resolved.

It is well known that comparative photographs taken with large and small lenses have about the same macroscopic contrast in the presence of haze, but have ground resolution approximately proportional to focal length. There are many other factors, such as innate lens quality, the mounting, shutter, filter, used, etc. However, if

136

all these things are held equal, and if the laboratory performance in lines/mm remains independent of focal length as is largely true of modern aerial lenses for the USAF, then the focal length is the most important factor in achieving ground resolution. Desirably, one should have as perfect a lens as prudence and economy of construction can supply, but the focal length is still the most important factor.

Where attention is given on a high technical level to every detail of the aerial lens, photography and flight conditions, one can readily achieve an average of 25 lines/mm in the air with a top quality lens. The imperfect lens may still produce an average of 18 lines/mm, if all other factors are optimized. These figures are obtainable only in the absence of haze and for high target contrasts on the ground.

It has proved possible under excellent conditions to reach a resolving power in the air of 42 lines/mm in a direct test run over ground targets, a figure that the author believes to be the peak so far recorded anywhere. This one test run proves that on excellent days the pure gaseous atmosphere remaining contributes only a slight amount of haze if a red filter is used. However, over much of the world, the atmosphere on the average is very hazy, whereby the differences between good and bad lenses in practice are much obscured.

Focal Problems in Aerial Photography

Possibly the most direct cause of inferior aerial photographs is simply an inadequate focal setting. A "poor" lens may have a greater depth of focus than a "perfect" lens. Therefore, if owing to errors of usage, the lens is out of focus, it can easily happen that the poor lens will return a better picture than the perfect lens. The perfect lens has a higher peak resolution but may have a shorter base to the curve of resolution against focal setting. In the limit with a perfect lens the depth of focus curve can be calculated from known emulsion properties and the nature of the target. The agreement is exact. With an imperfect lens one finds it difficult to calculate the curve of resolution against focal setting, owing to the influence of color aberrations, zones of the lens, etc.

The problem of focusing a lens system depends greatly on the nature of the images produced and on the light source, filter and emulsion. Usually, for an imperfect lens there will be a focal position of best resolution and another of best contrast, though the latter may change with the target contrast. For the perfect lens the optimum focal settings for best resolution and best microscopic contrast coincide. At moderate levels of resolution such as 10 lines/mm, the limitation being produced by vibration, image motion and/or by haze, *focusing for best microscopic contrast* in the laboratory may prove to be the best answer. If vibration and image motion are eliminated, and if haze is moderate, at a level of 20 lines/mm, *focusing for best resolution* probably will yield optimum results. An experienced photographer in the field might consider the situation and set his focus accordingly. It is unlikely that routine observers will have sufficient training, and hence fixed focus cameras are still necessary.

Many workers in aerial photography are not acquainted with the fact that the focal position of a lens is not necessarily stationary. A lens may be said to be *factory-focused*. This statement implies that there exists a focus and that the only problem is to find its position.

Once such is accomplished, the aerial pictures taken thereafter are always supposed to be in focus.

The smaller aerial lenses have such slight changes of focus within the large tolerance permitted by the observed depth of focus curves as to cause little or no focus troubles in practice. Large aerial lenses, however, exhibit a noticeable shift in focus caused by: (1) Temperature changes between equilibrium conditions; (2) thermal gradients or transients; (3) air density effect; (4) ground distance, according to vertical or oblique; (5) simple flexure of the camera and its component parts; (6) change in focal setting with change of filter and/or emulsion; (7) change in focus caused by stopping down, if lens imperfect. We are concerned here with only the first four causes.

Figure 10.12 shows the results of observations made in a cold-chamber test of a 40-inch lens. The total range of focal changes amounts to nearly 1 mm, owing to thermal gradients and to changes between equilibrium temperatures. The depth of focus at the level of 40 lines/mm is of the order of 0.15 mm. A focal error of the order of 0.5 mm can already cause a loss of resolution to perhaps 20 lines/mm or so.

Figure 10.13 shows the combined effects of ground distance for vertical photographs and focal changes caused by loss of air from the lens at the various altitudes. Here again appreciable errors are
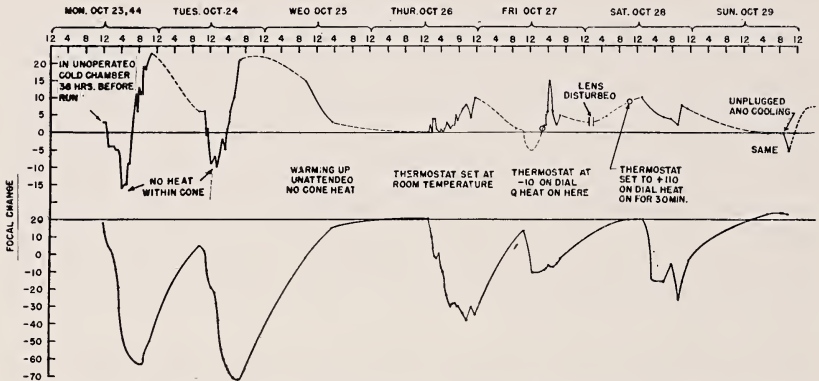

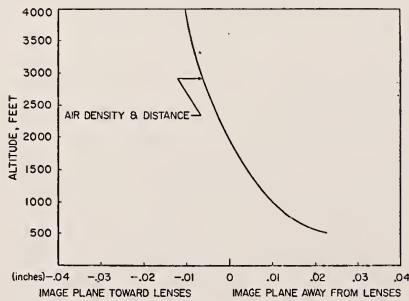FIGURE 10.12.   *Focal changes of a 40-inch f/5 telephoto lens as a function of temperature and temperature gradients.*

FIGURE 10.13.   *The air-density effect and ground-distance image shift for a 40-inch f/5 telephoto lens.*

encountered for both effects. At very low altitudes, of course, the ground distance effect becomes very large.

Table 10.2 contains calculated shifts for the case of a 24-inch and of a 144-inch focal length. It is clear that the focal errors are so large as to represent a serious problem to the aerial photographer. Thermal-gradient changes are even more serious than equilibrium changes, as shown by figure 10.12.

TABLE 10.2. *Change in equilibrium temperature*

A change of −80° C causes the image in the case of a 24-inch focal length to move 0.90 mm beyond the film plane away from the lens. The corresponding shift in the case of a 144-inch focal length is 6.30 mm.

Air-density effect		
	Focal length	
Altitude	24-inch image shift	144-inch image shift
ft	*mm*	*mm*
Infinity	−0.65	−3.44
40,000	− .50	−2.60
20,000	− .32	−1.71
10,000	− .19	−1.00
5,000	− .11	−0.54
2,000	− .05	− .24
Ground	.00	.00
Ground-distance effect		
Infinity	0.00	0.00
40,000	.03	1.10
20,000	.06	2.20
10,000	.12	4.39
5,000	.24	8.78
2,000	.61	21.95

Ideally, large lenses ought to be focused in the air immediately before a picture run is made. Short of this, the next best procedure is to prepare tables of focal changes from laboratory cold- and pressure-chamber observations. There is little justification in one's guessing at the focal setting for a large lens unless such tables have been prepared. A trained observer should also take thermal gradients into account. If the focal problem is carelessly handled, pictures as poor as 3 lines/mm might result and one may as well make use of a lens of shorter focal length.

Several of the lenses discussed above have been tested thoroughly at the Eastman Kodak Co. through the work of L. A. Jones, R. N. Wolfe, and associates. A very few of the voluminous and careful test results are reproduced here by permission of the Research Laboratories, for which acknowledgment is hereby made.

Figure 10.14 and figure 10.15 show the results of resolving power against focal setting for a 6-inch f/3.5 spherically-concentric lens. Because of the spherical symmetry no single optical axis exists. The very slight deterioration of the tangential resolving power in figure 10.14 is brought about by the fore-shortening of the aperture in the oblique beam with the consequent loss of diffraction resolving power. This lens was corrected for use with a red filter, and the spherical correction at f/3.5 was just slightly beyond one Rayleigh limit. The curves at high contrast show peak resolutions as much as 70 lines/mm
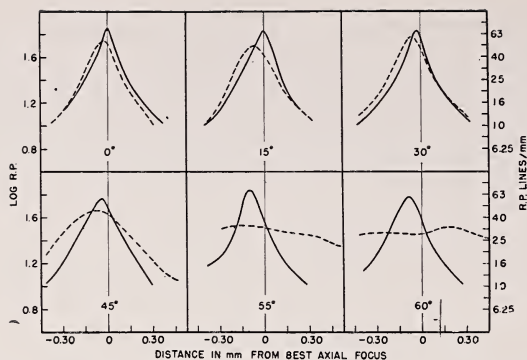
FIGURE 10.14. *Resolving power versus focal setting for a 6-inch f/3 wide-angle lens with red filter, Super-XX Aero Pan, and high-contrast 3-line test object.*
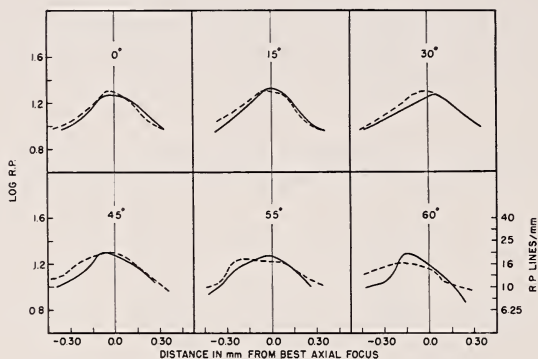---, radial; ——, tangential.

FIGURE 10.15. *Resolving power versus focal setting for a 6-inch f/3 wide-angle lens with red filter, Super-XX Aero Pan, and low-contrast Cobb test object.*
---, radial; ——, tangential.

on Super-XX. The apparent large depth of focus is brought about by the broadening of the bright lines of the test pattern at the expense of the dark spaces between the lines. That is, at a position 0.3 mm outside of focus, the dark spaces were only 0.020 mm wide on the test negative but still one could see three separate lines. This effect increases the depth of focus of the three-line pattern over and beyond what one would normally expect from the size of an out-of-focus star image.

Figure 10.15 shows the peak resolution to be expected with the Cobb 2-line test chart at low contrast (log contrast equals 0.17). This f/3.5 lens is nearly as good a lens as might be required to test the film itself. Hence, the loss of resolution down to about 20 lines/mm is an emulsion property. It is clear again that good image correction can be swallowed up at low contrast by the turbidity of the emulsion.

Testing at low contrast provides only a compressed scale for distinguishing between good and poor lenses. Table 10.1 is typical of what happens. Figure 10.15 as compared to figure 10.14 points it out again. If testing at low contrast is to be accomplished without additional testing at high contrast, the observer must content himself

140

with considering that no lens is very good and that every *designed* lens is fair.

The author believes that the high peak resolution observed may be brought about by the f/3.5 speed rating of the spherical lens, along with a near absence of spherical aberration and color. The 40-inch f/5 telephoto lens in the controlled models at least was just as well corrected for spherical aberration but in red light gave somewhat lower peak resolutions. The probable slight dependency on f/number ought to be kept in mind as more data become available. The dependence is shown to some extent by the many Eastman tests.

Figure 10.16 shows wedge photographs made with the 40-inch f/5 telephoto. The figure is more or less self-explanatory. The ordinates are logarithmic and hence the length of the sharp peak indicates that most of the light is where it belongs. For comparison purposes figure 10.17 shows similar wedge patterns for the standard 24-inch f/6 aerial lens, which is of Tessar construction, and about at the limit of what a Tessar can be expected to accomplish. The sharp peaks of figure 10.16 are intended to produce high resolving powers with good microscopic contrast at the level of 30 lines/mm. The shallow
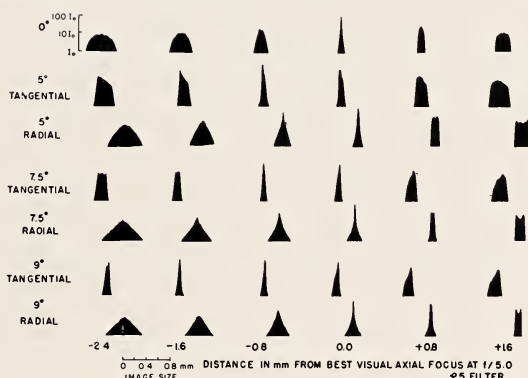

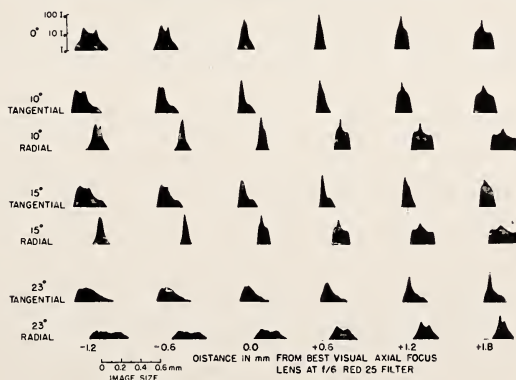FIGURE 10.16. *Wedge tests on images from a 40-inch f/5 telephoto lens.*

FIGURE 10.17. *Wedge tests on images from a 24-inch f/6 Tessar lens.*

peaks of figure 10.17 in cases of heavy exposure produce "muddy" photographs so far as the microscopic contrast is concerned, at levels around 20 lines/mm. Figures 10.18 and 10.19 show the resolution curves for this same f/6 standard Tessar for high and low contrast.

Figure 10.20 shows the wedge photographs for a 100-inch astronomical Ross lens. The peaks are sharp, even though there is very
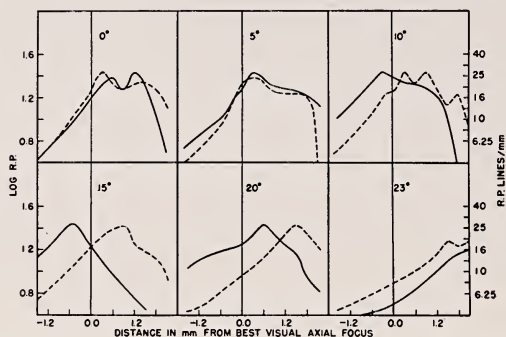

FIGURE 10.18. *Resolving power versus focal setting for a 24-inch f/6 Tessar lens, with Super-XX Aero Pan, tungsten, no. 12 filter, and high-contrast 3-line test object.*
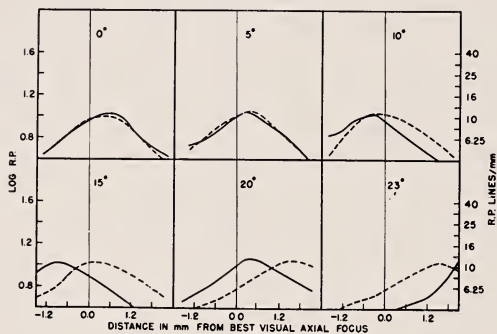
- - -, radial; ——, tangential.

FIGURE 10.19. *Resolving power versus focal setting for a 24-inch f/6 Tessar lens, with Super-XX Aero-Pan, tungsten, no. 12 filter, and low-contrast Cobb test object.*
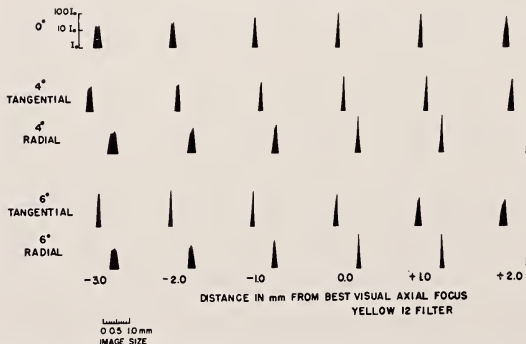
- - -, radial; ——, tangential.

FIGURE 10.20. *Wedge tests on a 100-inch f/10 Ross lens.*

142

considerable secondary spectrum. Figures 10.20 and 10.22 show the observed resolving powers at high and low contrasts. It should be noted that the optical glass of this lens was deficient. The observed axial visual image in red light showed no "clean" edge, but rather a hazy patch of small size. The loss of contrast in the edge causes a loss of microscopic contrast for the finer lines. This loss is barely discernible in the wedge photographs of figure 10.21, where the approach to the peak is slightly broader than would have been the case if the lens had been fully corrected. Another lens of the same design made since shows resolving powers as high as 55 lines/mm at high contrast.

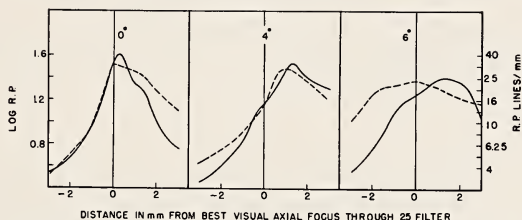Figures 10.23 and 10.24 show the resolving power results for the

FIGURE 10.21. *Resolving power versus focal setting for a 100-inch f/10 Ross lens with Super-XX Aero Pan, tungsten, no. 12 filter, and high-contrast 3-line test object.*

_ _ _, radial; —, tangential.

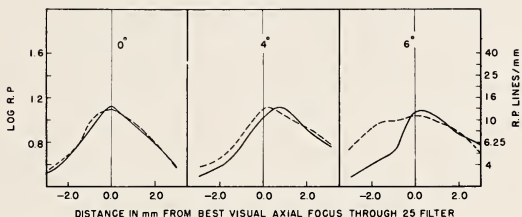FIGURE 10.22. *Resolving power versus focal setting for a 100-inch f/10 Ross lens with Super-XX Aero Pan, tungsten, no. 12 filter, and low-contrast Cobb test object.*
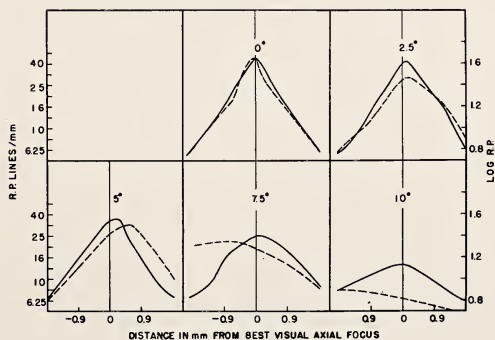
_ _ _, radial; —, tangential.

FIGURE 10.23. *Resolving power versus focal setting for a 36-inch f/8 apochromat with Super-XX Aero Pan, "daylight", no. 12 filter, and high-contrast 3-line test object.*
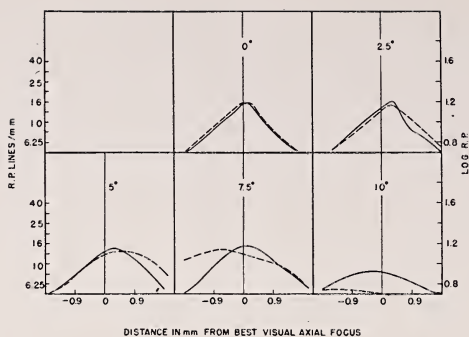
_ _ _ _, radial; —, tangential.

143

FIGURE 10.24. *Resolving power versus focal setting for a 36-inch f/8 apochromat with Super-XX Aero Pan, "daylight" no. 12 filter, and low-contrast Cobb test object.*

Object distance=28.8 focal length; — — — —, radial; ——, tangential.

36-inch f/8 apochromat discussed above. The effect of the apochromatism is not evident in the resolving powers as such but would show up more in the level of microscopic contrast around 30 lines/mm. Such lenses as the 36-inch apochromat would take aerial pictures of good microscopic contrast around the average expected resolving power in the air of 25 lines/mm., where other factors have been minimized.

The Calculated Photographic Image

With the coming use of electronic calculating equipment in optical design, consideration must be given as to whether the photographic image can be calculated. The cost of a large lens is so considerable as to make it desirable to go as far as possible on paper during the design stages. To a considerable extent the views described above have been drawn on repeatedly by the author in designing photographic systems and in predicting by judgment alone what might resonably be expected of the lens in photographic test. However, if it should prove possible to calculate the photographic image with accuracy, a step in quantitative analysis will have been achieved.

The author with the valued assistance of W. Randolph Angell, Jr. has made a preliminary attempt to calculate the photographic image in a special case. A simple plano-convex lens of barium crown glass with its normal complement of primary color was set up in the laboratory. Infrared photographic resolving power focusing runs were made at a 10-fold reduction between collimator and lens. A minus-blue filter was used, but otherwise the standard aerial infrared film was exposed to the spectral colors from a 3,000° tungsten source. Figure 10.25 reproduces a portion of the photographs taken. The second pattern seems resolved and corresponds to a resolution of 16.6 lines/mm When one considers that the primary spectrum from 5000 to 9000 angstroms is altogether uncorrected, that the lens has some spherical aberration, and otherwise is f/4, the observing of 16.6 lines/mm indicates to some extent how inadequate a lens can be and still give fair photographs. Of course, the microscopic contrast is low and even at a level of 5 lines/mm the photograph resulting would appear "muddy"
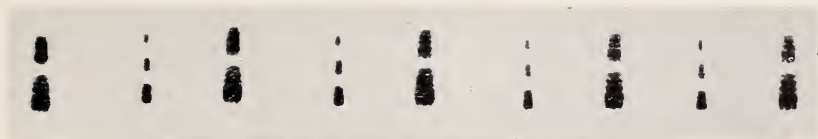
144

FIGURE 10.25. *Enlarged view of test images produced by a simple f/4 lens of 10-inch focal length.*
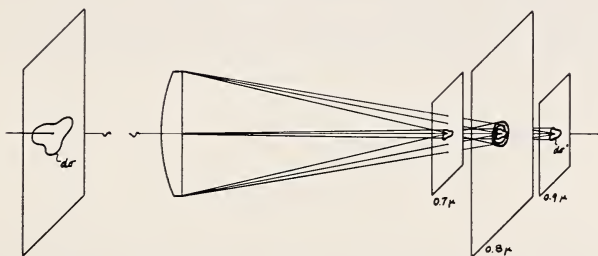
Pattern 2 is 16.6 lines/mm.

FIGURE 10.26. *Schematic view of the problem of calculating the photographic image in the presence of large chromatic aberration.*
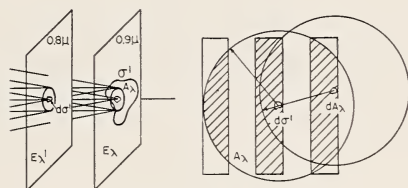

FIGURE 10.27. *Enlarged view of the image region in the problem of calculating the photographic image in the presence of large chromatic aberration.*

and unsatisfactory. Some of the out-of-focus colors have blur circles considerably larger than the entire test pattern.

Figure 10.26 shows the initial circumstances of the attempt to calculate the image. A mean focal plane was chosen to lie at 0.8 micron, though the results of figure 10.25 are not necessarily found optimum at 0.8. A small object element of area providing the illumination images into a similar area, $d\sigma'$. The illumination on the mean focal plane at 0.8 micron consists of the in-focus 0.8-micron light plus the inside and outside focus neighboring colors.

The illumination of any element $d\sigma'$ in its own focal plane is given by

$$E_\lambda d\lambda d\sigma' = k\pi \sin^2 \theta' B_\lambda d\lambda d\sigma,$$

which is the standard formula. We need it here only to show that the distribution of the illumination E_λ with wavelength is proportional to the similar emission from the object. The *dilution factor* k need not be evaluated, and we can work directly with E_λ, instead of B_λ.

Figure 10.27 shows an enlarged view of the space around the image. element. E_λ' (not a derivative) is taken as the illumination function on the chosen focal plane at 0.8 micron, resulting from the illumination E_λ on the displaced focal plane at another wavelength, λ. If the ele-

145

ment of area $d\sigma'$ is taken to be extremely small compared to the color aberration, then the illumination in the focal plane at 0.8 micron will be

$$E'_\lambda d\lambda d\sigma' = \iint_{A_\lambda} \frac{E_\lambda d\lambda d\sigma'}{(A_\lambda)_{\text{max.}}} \, dA_\lambda,$$

where $(A_\lambda)_{\text{max.}}$ is the total area of the out-of-focus blur circle striking the 0.8-micron focal plane, and where the integration is taken over the entire illuminated area of the test pattern, whatever its shape, within the radius of the blur circle of area $(A_\lambda)_{\text{max.}}$.

This integration can be performed at each wavelength, and in fact

$$E'_\lambda d\lambda d\sigma' = \left[\frac{\iint_{A_\lambda} dA_\lambda}{(A_\lambda)_{\text{max.}}} \right] E_\lambda d\lambda d\sigma'.$$

If we multiply by the dependence of filter transmission on wavelength, F_λ, and by the sensitivity of the emulsion relative to wavelength, S_λ, we have

$$F_\lambda S_\lambda E'_\lambda d\lambda d\sigma' = \left[\frac{\iint_{A_\lambda} dA_\lambda}{(A_\lambda)_{\text{max.}}} \right] (F_\lambda S_\lambda E_\lambda) \, d\lambda d\sigma'.$$

To convert to the characteristic curve, we have simply

$$\log \mathscr{E} = \log \int_0^\infty \left[\frac{\iint_{A_\lambda} dA_\lambda}{(A_\lambda)_{\text{max.}}} \right] (F_\lambda S_\lambda E_\lambda) \, d\lambda + \log \kappa,$$

where κ is a constant. Where the blur circle is completely illuminated for a macroscopic area, σ', then the photographic density can be set at any desired value by an adjustment to the abscissa scale. In particular, for infrared aerial film we can set this maximum density that would be obtained by photographing a uniformly illuminated surface at 1.6 on the standard characteristic curve for this emulsion. For any other pattern to be calculated, we have only to evaluate

$$\log \int_0^\infty \left[\frac{\iint_{A_\lambda} dA_\lambda}{(A_\lambda)_{\text{max.}}} \right] (F_\lambda S_\lambda E_\lambda) \, d\lambda.$$

This procedure has been carried out for the case actually photographed, as in figure 10.25, but only for the 10 lines/mm resolution level. Figure 10.28 shows the variation of focal position with color, which represents a substantial change. Figure 10.29 shows a graph of B_λ, which we can take directly as being equivalent for our purposes to E_λ. Figure 10.30 is the spectral sensitivity curve obtained from
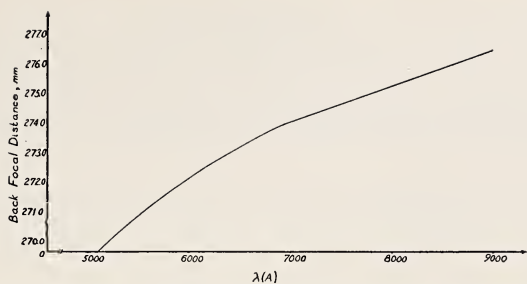
146

FIGURE 10.28. *Variation of back focal distance with color.*
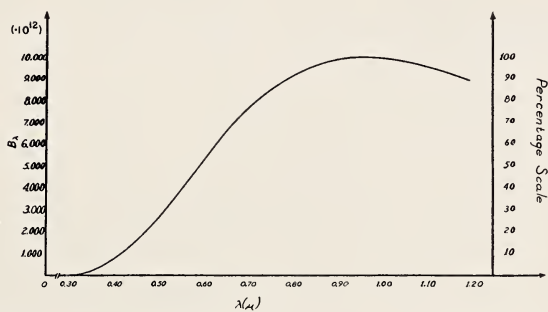

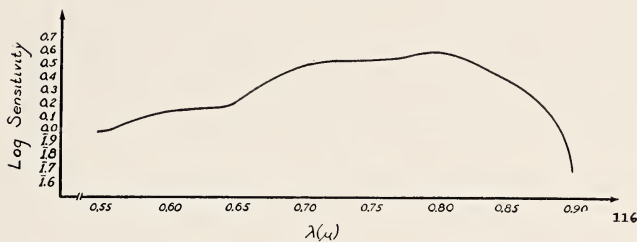FIGURE 10.29. *Black-body radiation curve for 3,000° K.*

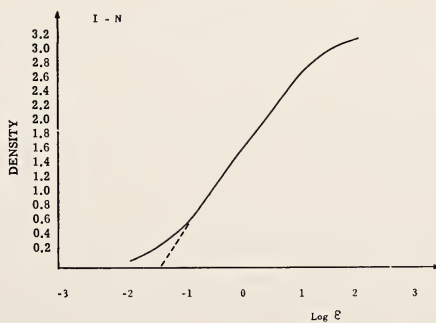FIGURE 10.30. *Spectral sensitivity curve for I–N emulsion.*

FIGURE 10.31. *Characteristic curve adopted for I–N emulsion.*

147

the Kodak Handbook of Photographic Plates. Figure 10.31 is the characteristic curve adopted for type I–N infrared emuslion. Figure 10.32 is the product of the filter function F_λ (minus-blue), S_λ, and E_λ. The total area under the curve represents the maximum exposure. Integration was performed by planimeter, in convenient units. The logarithm of the area was then set at density 1.6 on the characteristic curve.

Figure 10.33 represents typical curves for the indicated points on the pattern. The solid curve represents the illumination at point A on the inserted diagram, which is within a bright line. The peak illumination comes then from the in-focus 0.8-micron light and neighboring colors. Similarly, at point B where there would be no illumination if the lens were good, one can see from the dashed curve that at 0.8-micron there is indeed no illumination. Other outlying colors do contribute to the surface brightness, however, mostly from colors not far distant.

It is to be noted that the blur circles include all three lines of the pattern. The integration is performed by moving the chief ray or point of the planimeter over the outline of the bright areas of the best pattern within the requisite distance of $d\sigma'$. The color aberrations of this particular lens are so large as to have the entire out-of-focus pattern contributing to the photographic image at each point.

Figure 10.34 is the final result, where the calculated densities are tabulated on an enlarged diagram of the three-line pattern. The largest calculated density is 1.20, whereas the adopted density for macroscopic areas is 1.6. However, a density of 1.20 still represents

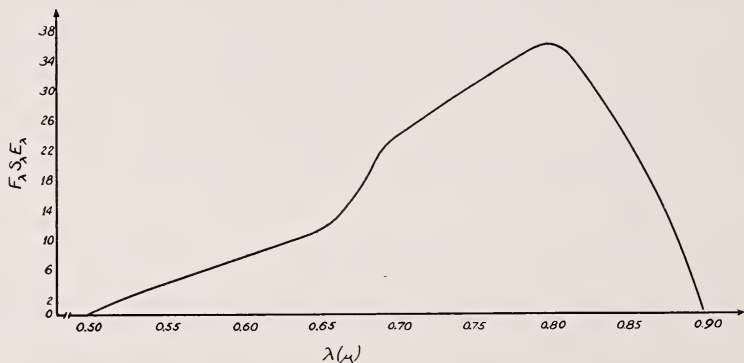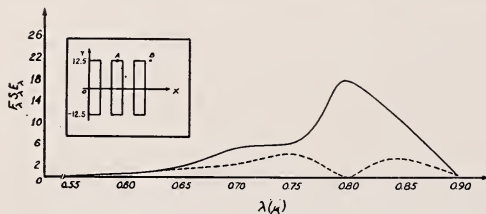FIGURE 10.32. *Maximum effective exposure for filter, emulsion and light source.*

FIGURE 10.33. *Sample curves of calculated illumination versus wavelength.*

Back focal distance=275.27 mm at $\lambda=0.8\,\mu$. —, exposure A at $X_0=12.5$ ul., $Y_0=12.5$ ul.; - - -, exposure B at $X_0=27.5$ ul., $Y_0=12.5$ ul. 1 ul=1 unit length=10 microns   Lines and spaces of resolving-power pattern are each 50 microns wide.

148

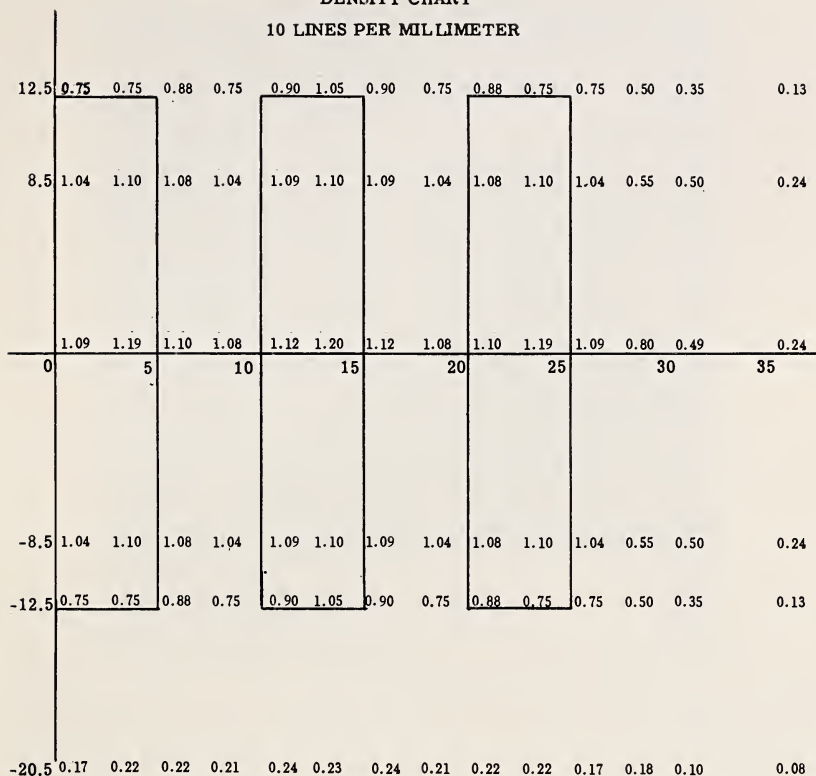| 12.5 | 0.75 | 0.75 | 0.88 | 0.75 | 0.90 | 1.05 | 0.90 | 0.75 | 0.88 | 0.75 | 0.75 | 0.50 | 0.35 | 0.13 |
8.5	1.04	1.10	1.08	1.04	1.09	1.10	1.09	1.04	1.08	1.10	1.04	0.55	0.50	0.24
	1.09	1.19	1.10	1.08	1.12	1.20	1.12	1.08	1.10	1.19	1.09	0.80	0.49	0.24
0		5		10		15		20		25		30	35	
-8.5	1.04	1.10	1.08	1.04	1.09	1.10	1.09	1.04	1.08	1.10	1.04	0.55	0.50	0.24
-12.5	0.75	0.75	0.88	0.75	0.90	1.05	0.90	0.75	0.88	0.75	0.75	0.50	0.35	0.13
-20.5	0.17	0.22	0.22	0.21	0.24	0.23	0.24	0.21	0.22	0.22	0.17	0.18	0.10	0.08

FIGURE 10.34. *Calculated photographic image of a three-line resolution pattern at 10 lines/mm level for test case of unachromatized simple lens.*

good blackness of the image, and indeed this density can be changed by exposure time and development.

The density differences between the spaces and the centers of the bright lines seem small, but experience with measured targets indicates that the eye would call this pattern well resolved. The microscopic contrast would be low, but the three lines could easily be seen.

It is difficult to determine whether the calculated and observed patterns agree. We plan to repeat the experiment with a photomicrographic enlargement as well, in order to have some control on the photographic factors. Also, in redoing the work, we plan to determine the actual best-focus position of the emulsion against millimeters of back focal length, as calibrated from the sodium image. Also, the lens will be stopped down enough to minimize the added effects of spherical aberration.

An Objective Method of Testing

During the summer of 1944 in the wartime laboratory at Harvard, the author initiated a form of testing designed to eliminate the personal equation from evaluation of lens performance. Special cardboard charts were printed for the purpose and copied to target size

149

FIGURE 10.35.   *A view of the large testing tunnel.*

at varying contrasts onto lantern slide plates.   These charts contained lines of thoroughly scrambled block letters, the letters in each line being smaller than in the preceding line in a geometrical progression. Test negatives from the tunnel were read off by an uninformed observer against a check list.   The lens performance was scored according to the line where the observer made fewer than 50 percent errors.   This particular testing method was used for awhile in 1944 and then set aside for other more urgent activities.

In recent months the author has taken up this type of test procedure once again, the purpose being to determine what image properties of a lens facilitate recognition of various types of objects at varying contrasts.   Time has not permitted more than a preliminary series of observations, but enough can be presented here to indicate the nature of the work.

Figure 10.35 shows a view of the testing tunnel in the author's laboratory.   The tunnel is 28 feet long, 4 feet square, and so constructed that the sides of the tunnel may be fastened to the ceiling or lowered at will.   A 16-inch paraboloidal mirror is set up as a collimator at the far end of the tunnel.   This mirror delivers the collimated rays from a target about in the center of the tunnel to the test lens at the other end.   Ordinary focal-plane test photographs or photomicrographic enlargements can be made for any lens up to 100-inches focal length or so.

In order to avoid small photographic targets with uncertain properties for the smallest resolution lines, the writer has prepared to use a *live* target, shown in figure 10.36.   The target cabinet is made of oak and houses 52 light bulbs of *daylight* type rated at 60 watts. There is an intermediate diffusing screen and then an interchangeable target panel.   The present target measures approximately 24 by 24
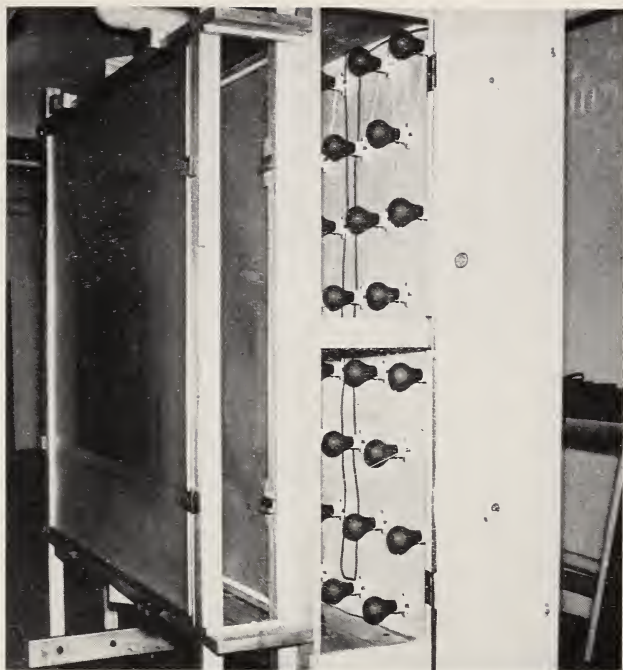
150

FIGURE 10.36.   *A view of the test target and cabinet.*

inches and is handpainted onto a 48- by 48-inch ground glass.   The unpainted surface serves as a macroscopic photometric standard. The surface brightness corresponds to that of the ground on a cloudy-bright day.

In testing lenses of short focal length one can photograph the target directly from distances up to 50 feet in the laboratory itself. For testing lenses of greater focal length, one reimages the target by means of a suitable reducing lens of high quality onto the focal plane of the collimator mirror.   The resulting small image is collimated by the mirror and the parallel rays sent to the test lens.   In this way one can be sure that the surface brightness of even the smallest target will remain the same as for the macroscopic photometric areas.

Varying contrasts are introduced into the collimated beam by means of a beam-splitter placed over the reducing lens near the focus of the mirror.   The lowered contrast will then be brought about by a superposition of uniform illumination onto the whole target area including the photometric area and test patterns.   Temporarily, the writer has had to approximate the equivalent by adding fogging exposures to the target exposures and for this paper has made use of a 3-inch f/3.5 Tessar test lens.

Because the target panels are interchangeable, one can copy large-scale transparencies as *live* objects down to the smallest size used for test purposes.   Thus, different surface objects of interest, such as roads, houses, streams, railroads, etc., can be copied at varying scale with a given lens and the range of its effectiveness for each kind of target determined.   Variable haze and color of haze can easily be introduced.

151

Figure 10.37 shows the target used in the present preliminary investigation. The picket-fence pattern is intended for microdensitometry. The standard three-line patterns decrease in size according to the cube root of 2, although admittedly, the sixth root of 2 is more desirable. Similarly, the letters are selected for uniqueness from the alphabet and in each line have been scrambled to confuse the observer. The block letters are as large as the three-line patterns, and the line width of the letter in every case is equal to the line width of the corresponding three-line pattern. There are 10 *shape* objects selected for distinctness but basic to the problem of *recognition*. Also, there are 10 *figure* objects, selected for recognition tests of orientation and shape.

Every test line and shape in the entire target has clean-cut edges and sharp intersections, produced by ruling pen and hand correction. Thus, every edge and corner delivers a pure *square wave* to the test lens via the collimator. There is no intermediate photographic process to cause a crumbling away of edges and corners of the smaller patterns, and the minifying lens used at the focus of the collimator at f/15 gives results of microscopic quality. The only real deficiency in this *live target* set-up arises from small air turbulence within the test tunnel. Experience to date indicates that the air in the tunnel limits performance to approximately 1 second of arc.

Five test films have been taken. The results of the fifth run are given in figures 10.38, 10.39, 10.40, and 10.41. The author reduced

FIGURE 10.37. *The test target for letters, shapes and figures as used at present.*

the negatives personally, making use of a 60X microscope and calling off each shape, letter, or figure to a recorder who scored against a key. The order of reading was altered frequently, nor did the author examine the key prior to the microscopic examination of the test negatives. Nevertheless, it was apparent that slight effects of memory do interfere with the complete objectivity of the test, and arise in the unavoidable process of elimination.

In figures 10.38 through 10.41 the first column gives the resolving power at the image plane of the test lens, corresponding to the pattern number given in the second column. However, these resolving-power figures are not the observed values but are the progression used. Row 1 is the finest pattern, and also represents the finest patterns among the shapes, figures and letters. The scale, then, runs from 5 to 40 lines/mm. The check marks mean that the particular test object was read correctly against the key.

Roll 5 - SUPER-XX, 18 min. Microdol$_3$, 70°C, tray.

R.P.	Row	Shapes	Figures	Letters	#
40	1			x x x x x x x x x x	5-1
32	2			x x x x x x x x x x	
25	3	x x x		x x x x x x x x x x	f/11
20	4	x x	x x x x x x	x x x x x x x x x x	
16	5	x x x x x	x x x x x x x	x x x x x x x x x x	1/500
13	6	x x x x x x x	x x x x x x x x	x x x x x x x x x x	40/32
10	7	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	0.45 / 0.06
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x x x x x x x	5-2
32	2			x x x x x x x x x x	
25	3	x x x	x	x x x x x x x x x x	f/11
20	4	x x x x	x x x x	x x x x x x x x x x	
16	5	x x x x x x x x x x	x x x x x x	x x x x x x x x x x	1/250
13	6	x x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	40/32
10	7	x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	0.53 / 0.07
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x x x x x x x	5-3
32	2			x x x x x x x x x x	
25	3	x x	x	x x x x x x x x x x	f/11
20	4	x x x x x x x	x x x	x x x x x x x x x x	
16	5	x x x x x x x x	x x x x x x	x x x x x x x x x x	1/125
13	6	x x x x x x x x x x	x x x x x x	x x x x x x x x x x x	40/32
10	7	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	0.51 / 0.06
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x x x x x x x	5-4
32	2			x x x x x x x x x x	
25	3			x x x x x x x x x x	f/11
20	4	x	x x	x x x x x x x x x x	
16	5	x x x x x x	x x x x x x x	x x x x x x x x x x	1/61
13	6	x x x x x x x x x	x x x x x x	x x x x x x x x x x	40/25
10	7	x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	0.80 / 0.07
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	

FIGURE 10.38. *A tabulation of individual test results, 5–1 to 5–4.*

Roll 5 - Super-XX, 18 Min., Microdol$_3$, 70°C, tray.

R.P.	Row	Shapes	Figures	Letters	#
40	1			x x x x x x x x x x	5 - 5
32	2			x x x x x x x x x x	
25	3	x x		x x x x x x x x x x	f/11
20	4	x x x x x x x x	x x x	x x x x x x x x x x	
16	5	x x x x x x x	x x x x x x x x	x x x x x x x x x x	1/31
13	6	x x x x x x x x x x	x x x x x x x x	x x x x x x x x x x	40/25
10	7	x x x x x x x x x x	x x x x x x x x x 'x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	1.12 / 0.08
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x x x x x x .x	5 - 6
32	2			x x x x x x x x x x	
25	3	x x x	x x X	x x x x x x x x x x	f/11
20	4	x x x ·x x	x x x x x x x	x x x x x x x x x x	
16	5	x x x x x x x x	x x x x x x x x	x x x x x x x x x x	1/16
13	6	x x x x x x x	x x x x x x x x x	x x x x x x x x x x	32/32
10	7	x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	1.46 / 0.33
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x	5 - 7
32	2			x x x x x x x x	
25	3			x x x x x x x x x x	f /11
20	4			x x x x x x x x x x	
16	5			x x x x x x x x x x	1/16
13	6	x x x x x x		x x x x x x x x x x	25/25
10	7	x x x x x x x x x	x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x	x x x x x x	x x x x x x x x x x	
6	9	x x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	1.57 / 1.42
5	10	x x x x x x x x	x x x x x x x x	x x x x x x x x x x	
40	1			x x x	5 - 8
32	2			x x x x x x x x	
25	3			x x x x x x x x x	f/11
20	4			x x x x x x x x x x	
16	5			x x x x x x x x x x	1/31
13	6	x		x x x x x ·x x x x x	25/20
10	7	x x x x x	x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x	x x x x x x x x x x. x	
6	9	x x x x x x x x x x	x x x x x x x x	x x x x x x x x x x	1.17 / 1.00
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	

FIGURE 10.39.   *A tabulation of individual test results, 5–5 to 5–8.*

In the last column on the right, we have from top to bottom in each group the film number and picture number. The f/number is then given, and next the exposure time as marked on the lens. It is obvious that these exposures are not correct, because there is no recognizable difference in density between 1/500th and 1/125th. However, the first three observations can be compared for they should be equivalent. 40/32 means that the lens resolved 40 lines in the horizontal line pattern, and 32 lines in the vertical pattern. The pictures indicated that the lens possesses either a slight decentration or that the shutter introduced vibration, even though the camera was clamped to a concrete pier. The next figure is the distance to the target in feet, focused by rangefinder. The figure 0.45/0.06 represents the observed densities of the high light and low light as measured by visual densitometer, and therefore represents the contrast in terms of density differences.

154

Roll 5 - Super-XX, 18 Min., Microdol$_3$, 70°C, tray.

R.P.	Row	Shapes	Figures	Letters	#
40	1			x x x x x x x	5 - 9
32	2			x x x x x x x x x	
25	3			x x x x x x x x x x	f/11
20	4			x x x x x x x x x x	
16	5			x x x x x x x x x x	1/61
13	6	x x x x	x x x x x	x x x x x x x x x x	32/16
10	7	x x x x x x x x x x	x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	
6	9	x x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	0.92 / 0.64
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1'			x x x x x x x x x x	5 -10
32	2			x x x x x x x x x x	
25	3			x x x x x x x x x x	f/11
20	4			x x x x x x x x x x	
16	5	x x		x x	x x x x x x x x x x	1/125
13	6	x x x x x x	x x x x x x x	x x x x x x x x x x	32/25
10	7	x x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	0.68 / 0.44
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x x x x x x x	5 - 11
32	2			x x x x x x x x x x	
25	3			x x x x x x x x x x	f/11
20	4			x x x x x x x x x x	
16	5	x	x x x		x x x x x x x x x x	1/250
13	6	x x x x x x x	x x x x	x x x x x x x x x x	32/25
10	7	x x x x x x x	x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	0.76 / 0.49
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
40	1			x x x x x x x x x	5 - 12
32	2			x x x x x x x x x x	
25	3			x x x x x x x x x x	f/11
20	4			x x x x x x x x x x	
16	5	x x	x	x x	x x x x x x x x x x	1/500
13	6	x x x x x x	x x x x x x	x x x x x x x x x x	32/25
10	7	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	20'
8	8	x x x x x x x x x x	x x x x x x x x x	x x x x x x x x x x	0.74 / 0.47
6	9	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	

FIGURE 10.40.   *A tabulation of individual test results, 5–9 to 5–12.*

Roll   3-   Plus -X,   13 min. Microdol$_1$, 70°C, tray.

R.P.	Row	Shapes	Figures	Letters	#
40	1				3 - 12
32	2			x x x x x x x x	
25	3			x x x x x x x x x	f/11
20	4			x x x x x x x x x x	
16	5			x x x x x x x x x x	1/100
13	6			x x x x x x x x x x	20/20
10	7	x x x x x		x x x x x x x x x x	20'
8	8	x x x x x x x x	x x x x x	x x x x x x x x x x	0.89 / 0.75
6	9	x x x x x x x x	x x x x x x x x	x x x x x x x x x x	
5	10	x x x x x x x x x x	x x x x x x x x x x	x x x x x x x x x x	

FIGURE 10.41.   *A tabulation of individual test results ,3–12.*

155

The first six pictures of roll 5 (5–1 through 5–6 in fig. 10.38 and 10.39) are for high contrast and were taken by single direct exposures of the infinite contrast target at the marked exposure times given. The next six pictures of roll 5 (5–7 through 5–12 in fig. 10.39 and 10.40) were taken in reverse order of exposure time, and contained the direct exposure as marked and a fogging exposure of the photometric area, also of the exposure time marked. Hence, in this case the total exposure given the emulsion was in the ratio of 2/1, if we neglect the superposition of exposures. The recorded densities of high light and low light show what happened, and again we note from 5–10 to 5–12 no change in low light density in spite of a supposedly shorter exposure time for the last picture.

Figure 10.41 shows a similar set of observations made with a single direct exposure and two fogging exposures made for 3–12 at 1/100th of a second. The observed densities are 0.89/0.75. Figure 10.42 reproduces another low-contrast picture of densities 0.62/0.47. The patterns look well resolved in spite of the lowered contrast, but the picture *quality* is far from pleasing.

Table 10.3 gives a summary of the test results on roll 5 and on 3–12, and represents therefore three contrast groups. In order to make this tabulation the author arbitrarily assumed a criterion from figures 10.38 through 10.41 that at least 6 out of 10 test patterns had to be read correctly. The line so qualifying is then recorded in table 10.3 as the resolving-power equivalent. Thus, on 5–2 under *Shapes*

FIGURE 10.42. *Reproduction of a sample low-contrast target photograph.*

156

TABLE 10.3. *Target resolving power corresponding to 60-percent score*

Densities	Shapes	Figures	Letters	R.P.	Number
0.45/0.06	13	20	>40	36	5–1
0.53/0.07	20	16	>40	36	5–2
0.51/0.06	20	16	>40	36	5–3
0.80/0.07	16	16	>40	32	5–4
1.12/0.08	20	16	>40	32	5–5
1.46/0.33	16	20	>40	32	5–6
0.74/0.47	13	13	40	28	5–12
0.76/0.49	13	10	>40	28	5–11
0.68/0.44	13	13	>40	28	5–10
0.92/0.64	10	10	40	24	5–9
1.17/1.00	8	8	32	22	5–8
1.57/1.42	13	8	32	25	5–7
0.89/0.75	8	6	32	20	3–12

the equivalent resolving power 20 is recorded. This means that the three-line pattern at 20 lines/mm has a single line width equal to 0.025 mm and a height to each line of 0.125 mm. The corresponding *shapes* all have a longest dimension of 0.125 mm on the test negative and an average width of 0.025 mm. Under these conditions at least 6 out of 10 were recognized while at the same time the mean resolving power in the two directions on the same test negative 5–2 was observed to be 36 lines/mm. One would conclude that when the photograph resolves 36 lines/mm, shapes measuring 0.125 by 0.025 mm can be recognized with at least a 60-percent probability of correctness from among 10 varying objects of similar nature if the contrast is high on the photograph.

Table 10.3 indicates that on the whole the *Shapes* and *Figures* correspond well for equal difficulty of recognition. The letters on the other hand are so easily recognizable as to require a complete revision of the target.

It was obvious to the author in reducing the test negatives that there is a considerable variation in difficulty of recognition among the various shapes and figures in the same rows. After a few thousand readings have been made from test negatives taken with the best lens available at different scales, one could make up a new target with the shapes and figures on a line arranged for equal statistical difficulty, and similarly for the letters. This new target would then serve for testing inferior lenses and the criterion of a 60-percent score would become much more meaningful and more sharply defined between lines. As it stands, the diamond figure with its longest diagonal tilted at 45 degrees was recognizable down into the smallest lines of patterns. Likewise, among the shapes the so-called *truck* was recognizable, even when only a grain or two represented the wheels. Practice of the observer is also a real factor. The author became rapidly more skillful in interpreting the test negatives as he worked. Figures 10.38 through 10.41 were observed after the author had practiced for about an hour in a preliminary reading. The most difficult patterns to distinguish were those having *rounded* corners versus those with sharp corners, such as the shapes representing the basic profiles of a boat versus a canoe. The decay in the image quality caused by the lens-film combination does round off corners and therefore interferes with recognition.

Recognition Factors

Different types of objects have certain characteristics that lead to their recognition, as distinct from resolving power and detectability. Near the limit of the emulsion where the recognition factors become suppressed, recognition begins to fail. An imperfect lens in general will require a larger scale compared to the perfect lens before recognition can be achieved. Similarly, an increase in contrast of the image will improve the chances for recognition, whether this increase in contrast arises from use of a better lens or from higher target contrast or both. For example, the good lens may be able to show recognizability for a truck in the shade of a tree. The bad lens will fail here, but may nevertheless be able to provide recognizability for the truck at exactly the same scale, if the truck is out in sunlight. The laboratory target discussed above contains *recognition factors* for each of the shapes and figures involved.

Figure 10.43 shows these *recognition factors* in the case of standard block letters. The particular choice of the factors in the second line was made by the author, whereas other workers might adopt different criteria. However, an attempt is made here to have an arc as short as can be distinguished from a straight line segment, and to give only a minimum number of factors. Every letter can be distinguished clearly from every other letter in the scheme adopted.

The third line shows a message that appears entirely unfamiliar, and seems to be only a hodge-podge of letters. The fourth line shows a message that becomes legible upon close examination. The fifth line reproduces the fourth with fewer recognition factors, but is still recognizable once the fourth line has been read. Finally, the third line is simply the fourth line reproduced backwards without spaces. A trained interpreter could easily read the message in the third line. The point is that training is necessary in deciphering various types of objects, and the trained observer needs fewer recognition factors.

ABCDEFGHIJKLMNOPQRSTUVWXYZ

FIGURE 10.43. *Recognition factors for letters of the alphabet.*

The last double line presents a statement familiar to most people. These are ample clues for almost anyone to determine the meaning. A trained observer could read this line with fewer than half the clues. Context is so effective that clues required to decipher letters or words can be suppressed in the entire message. Hence, if the message is viewed as a unit, there will exist a minimum number of recognition factors for the message, as distinct from the individual letter. That is, each object or unit of information required will have this minimum number of recognition factors in context with the surroundings. A blur on a railroad track is likely to be a locomotive or a car. The same blur on a road is likely to be a truck. If on a runway, the same blur is likely to be an airplane, etc.

The type of lens testing described above is to be applied toward measuring the ability of the lens to produce recognition of standard objects and forms. Once correlations have been established with ordinary three-line patterns, the latter may continue to serve as a quick measure of lens performance. In the final analysis it should be possible to draw up a list of object types, and tabulate the scale versus contrast any given lens must employ in order to assure recognizability. A good lens should be able to provide such information at a smaller scale than a poor lens. Hopefully, the correlations between resolving power and contrast on the one hand, and the objective targets on the other, will be so distinct as to render the laboratory study sufficient for a broad range of applications in the field.

Discussion

Mr. D. P. Feder, National Bureau of Standards, Washington, D. C.: This question is addressed to Dr. Coleman. What about the difficulty of calculating the interferometer patterns that I notice seem to fit quite well to the observed pattern? How long a calculation is this?

Dr. Coleman: That takes about 8 hours.

Mr. Feder: Is that for any degree?

Dr. Coleman: Yes. You would have to know the aberrations to know the magnitude.

Dr. Baker: Is this with a test system?

Dr. Coleman: This is to calculate.

Dr. Baker: On a test instrument?

Dr. Coleman: Any system we considered.

Chairman: Are there other questions, please?

Dr. K. V. Pestrecov, Bausch & Lomb Optical Co., Rochester, N. Y.: This question is to Dr. Baker. You mentioned a beautiful wide-angle lens that resolves 60 or 70 lines/mm. Why isn't it used for aerial photography?

Dr. Baker: That is a spherically symmetrical lens with a curved field that I developed for aerial photography in 1941. The system has been used successfully. However, the curved field introduces awkward problems and as a consequence the lens has been little used. The first one was designed and built at the Harvard Observatory in 1941. Two more were built at Harvard during the war along with a rectifying printer projection lens designed by Dr. Grey.

Dr. Pestrecov: Were the calculations on the photographic image resulting from large color aberrations published?

Dr. Baker: The work will be published.   It is part of our activities at Harvard Observatory right now.

Dr. Pestrecov: Is it contract work?

Dr. Baker: Yes, it is contract work.   This portion of our work is declassified, as most of it probably will be.   It is of interest, I think.

Dr. Pestrecov: It is very interesting.   Do ou hope to publish it soon?

Dr. Baker: Well, I think so, as part of the proceedings of this Symposium.

Prof. B. O'Brien, University of Rochester, Rochester, N. Y.: I would like to ask Dr. van Heel if he has been successful in using ordinary photographic materials for recording these interesting patterns.   Granted the color film is ideal but there are inconveniences as has been emphasized; can you not make rather successful measurements from black and white photographs?

Dr. van Heel: I did not understand you.

Dr. Pestreeov: Can you not make rather successful measurements in black and white?

Dr. van Heel: I am sorry.   I have not been clear.   The whole method depends on the fact that you can discern colors.

Dr. Pestrecov: Must this necessarily be true?   For example, suppose you admit two narrow wavelength bands.   Can you not use these as appropriately?

Dr. van Heel: Assume that you use two filters.   The image of a red line and of a green line are next to each other on the plate.   If the limit of resolution is infinitely small, you can get anything you want, but in practice it is hardly feasible to obtain that precision in black and white.   I tried it with monochromatic light and with the photographic plate.   It is difficult to locate the right position.   I don't contend it is impossible, but why not use your ability to discriminate colors?

Dr. Pestrecov: I was thinking only of the practical difficulties of the time—was it two weeks?—required for the return of the processed films from Paris.

Mr. J. M. Naish, Optics Section, Instrument and Photographic Dept., Royal Aircraft Establishment, Hants, England: Would Prof. van Heel please say a little about the actual making of the grating? How accurate is it necessary to maintain the spacing and thickness of the wires?

Dr. van Heel: It should be perfect—a grating of that spacing would be like this (illustrating).   I ought to know all of these positions within 2 microns, and it would be fine without recalibrating every time, but I have not succeeded in making such a grating so I have to calibrate for all positions and all the lines and to use my calibration table.

Mr. Naish: And what is the ratio of line to space?

Dr. van Heel: It is unimportant.   In any case you get colors.

11. Geometrical and Interferential Aspects of the Ronchi Test

By G. Toraldo di Francia [1]

The Ronchi test, when performed with a low-frequency ruling, presents a purely geometrical character; ray optics is then fully sufficient to explain the appearance of the patterns. On the contrary, when a high-frequency grating is employed, the result is undoubtedly a phenomenon of interference and must be dealt with by means of wave optics. Still the interference patterns retain a close resemblance to the geometrical ones. It is interesting to investigate the cause of the resemblance and this is done very easily by means of the eikonal function. As the frequency increases the transformation of the shadows of the grooves into interference fringes is followed step by step. The theory is applied to the third-order aberrations and in that case it is found that, apart from a displacement, the interference pattern is identical to the geometrical one.

Introduction

One of the most sensitive tests for the quality of the image of a point is the Ronchi test. As is well known, this test consists in studying the light wave transmitted by an optical system by means of a diffraction grating. In spite of the fact that diffraction plays an important role in the observed phenomena, there can be no doubt that the ultimate result is the production of an interference pattern. For this reason it is customary to give the apparatus the name of "grating interferometer". On the other hand there are many cases when the test can be considered from a third point of view, that is, the merely geometrical one.

Since the early days of the Ronchi test there has been much speculation regarding its theory and interpretation. In this connection the question arose as to the simplest way of approaching the problem of interconnecting the three different aspects we have just mentioned. Many bulky calculations were carried out on this purpose, especially by the researchers of the Istituto di Ottica at Arcetri.

In almost all previous researches on the subject the simplifying assumption was made that the grating merely splits the impinging wave into many identical waves, rotated through a given angle with respect to one another. However this procedure did not bring about an actual simplification; on the contrary it led invariably to extremely lengthy calculations.

The purpose of this paper is to show how much the problem is simplified by a more rigorous approach that takes into account the actual phenomenon of diffraction. It will become apparent that in this way a better insight can be obtained of the passage from the pure geometrical phenomenon to the interferential one.

---

[1] Istituto Nazionale di Ottica, Arcetri, Firenze, Italy.

Some Geometrical Relations

The shape of the wave under test will be determined by means of the mixed eikonal function, or what amounts to the same, by means of its wave aberration at infinity. In other words the wave will be thought of as coming from infinity and its wave aberration will be computed with respect to a spherical wave of infinite radius, whose center we shall call C_r. Taking the axis of the wave as the x-axis, any point on the wave front at infinity can be defined by means of the two last direction cosines β,γ of the corresponding ray. The wave aberration $w(\beta,\gamma)$ will also be a function of these two cosines.

As is well known, from the theory of the eikonal function, if we call y and z the coordinates of the point where a given ray meets the plane through C_r perpendicular to the axis, we have

$$y=\frac{\partial w}{\partial \beta}, \quad z=\frac{\partial w}{\partial \gamma}. \tag{1}$$

From these relations it follows easily that, if w has the form

$$w=\frac{x}{2}(\beta^2+\gamma^2), \tag{2}$$

the wave is spherical and C_r is located at a distance x apart from the center of the wave.

Astigmatism corresponds to the wave aberration

$$w=\left(\frac{x}{2}+\frac{a}{4}\right)\beta^2+\left(\frac{x}{2}-\frac{a}{4}\right)\gamma^2. \tag{3}$$

If we assume the plane $\gamma=0$ to be the meridional plane, a is the distance from the radial focus to the tangential one and x is again the distance of C_r from the middle focus.

Third-order coma for a centered system is given by

$$w=\frac{x}{2}(\beta^2+\gamma^2)+c(\beta^3+\beta\gamma^2), \tag{4}$$

the plane $\gamma=0$ being taken as the meridional plane and C_r being at a distance x from the paraxial focus.

Third-order spherical aberration is given by

$$w=\frac{x}{2}(\beta^2+\gamma^2)+s(\beta^2+\gamma^2)^2, \tag{5}$$

where x has the same meaning as before.

Now let us take a thin rectilinear wire and put it in the path of the light wave. If the observer places his eye directly behind the wire, he sees the shadow of the wire projected at infinity. The shape of the shadow is an indication of the type of aberration that is present.

Now it is very easy to find an angular equation for this shadow. Let us suppose that the wire is parallel to the y axis and is placed at a distance z_0 from the x axis. We take C_r to be the point of the x

axis that has minimum distance from the wire. According to eq 1 the wire will meet with all rays satisfying the equation

$$\frac{\partial w}{\partial \gamma} = z_0. \tag{6}$$

This is the angular equation of the shadow. In practice, after writing eq 6 explicitly, one can replace β and γ by two rectangular coordinates Y and Z, respectively, and obtain the equation of the shadow as projected on some very distant screen. This is only an approximation, but a very good one, meeting all the requirements of practical application. Accordingly we shall rewrite eq 6 in the form

$$\frac{\partial w(Y,Z)}{\partial Z} = z_0. \tag{7}$$

In this way one obtains for a spherical wave from eq 2

$$xZ = z_0, \tag{8}$$

which is a straight line parallel to the y axis. This is represented in figure 11.1a, where the circle represents the exit pupil of the instrument.

In the case of astigmatism we assume that the meridional plane of the wave is rotated through an angle φ with respect to the y axis, that is, with respect to the wire. Thus we replace β with $\beta \cos \varphi - \gamma \sin \varphi$ and γ with $\beta \sin \varphi + \gamma \cos \varphi$. With this substitution the wave aberration (3) takes the form

$$w = \left(\frac{x}{2} + \frac{a}{4} \cos 2\varphi\right) \beta^2 - \frac{a}{2} \beta\gamma \cos 2\varphi + \left(\frac{x}{2} - \frac{a}{4} \cos 2\varphi\right) \gamma^2.$$
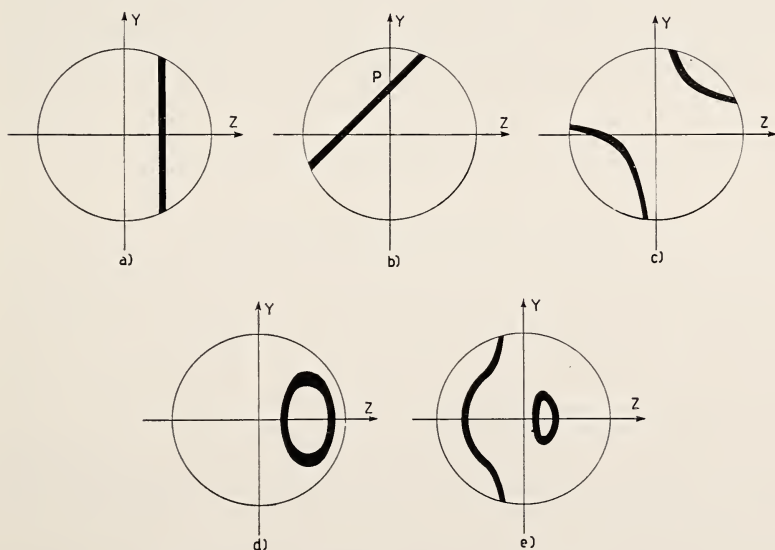
FIGURE 11.1.   *The shadow of a thin straight wire.*

a, Perfect spherical wave; b, astigmatism; c, coma (wire parallel to the meridional plane); d, coma (wire perpendicular to the meridional plane); e, spherical aberration.

163

By applying eq 7 we easily arrive at

$$Y = \frac{2x - a \cos 2\varphi}{a \sin 2\varphi} Z - \frac{2z_0}{a \sin 2\varphi}. \tag{9}$$

This equation represents a straight line. When x is made to vary, the line turns around a fixed point P of the Y axis (fig. 11.1, b).

Let us consider the case of coma. By applying (7) to (4) one easily gets

$$Z(2cY + x) = z_0. \tag{10}$$

The shadow has, therefore, the shape of an equilateral hyperbola, with the asymptotes parallel to the Y and Z axes, respectively (fig. 11.1, c). The center of the hyperbola is located on the Y axis (that is on the meridional plane) at a distance $-x/2c$ from the axis. It is interesting also to consider the case where the wire is perpendicular to the meridional plane, instead of being parallel to it. One has only to write (4) in the transposed form

$$w = \frac{1}{2} x (\beta^2 + \gamma^2) + c (\gamma^3 + \gamma\beta^2), \tag{11}$$

where the roles of β and γ have been interchanged, and to apply the eq 7. Thus, one obtains

$$xZ + 3cZ^2 + cY^2 = z_0. \tag{12}$$

as the equation of the shadow. This is an ellipse, with its axes parallel to the Y, Z axes (fig. 11.1, d.) The center of the ellipse is located on the Z axis at a distance $-x/6c$ from the axis of the wave.

Finally in the case of spherical aberration, we use eq 5 to obtain the equation

$$Z[x + 4s(Y^2 + Z^2)] = z_0. \tag{13}$$

This represents a cubic, very well known in optics (fig. 11.1, e).

The Grating Interferometer

Let us now consider the grating interferometer. If the spacing of the ruling is large, we can apply ray optics, considering the pattern as consisting of the shadows of the grooves. It would then be very easy to discuss the patterns, by means of the results of the preceding section. However, we will take another way of approach that will prove much more instructive.

According to Huygens principle, each point of the wave front at infinity sends out a spherical wave, which arrives at the center of the reference sphere as a plane wave. The phase of this plane wave can be computed by means of the wave aberration at infinity and is evidently $-2\pi w/\lambda$.

Thus, we arrive at the decomposition of the impinging wave into a bundle of plane waves, each having its own orientation and phase. Each one of these plane waves, when incident upon the grating is

164

split into a set of diffracted plane waves. We assume the grating to be perpendicular to the x axis and the grooves to be parallel to the y axis.

According to the laws of diffraction by a grating, a diffracted wave of the k'th order will have the direction cosines β, γ if the original wave had the direction cosines β', γ', given by

$$\beta' = \beta$$

$$\gamma' = \gamma - k' \frac{\lambda}{p},$$

p being the period of the grating. In the same direction we find the diffracted wave of k''th order, corresponding to an incident wave whose direction cosines are

$$\beta'' = \beta$$

$$\gamma'' = \gamma - k'' \frac{\lambda}{p}.$$

According to what we have stated above, the phases of these two waves will be, respectively,

$$\varphi'' = -\frac{2\pi}{\lambda} w \left(\beta, \gamma - k' \frac{\lambda}{p} \right)$$

$$\varphi'' = -\frac{2\pi}{\lambda} w \left(\beta, \gamma - k'' \frac{\lambda}{p} \right).$$

There will be positive or negative interference, accordingly, as the difference of the phases is a multiple of 2π or an odd multiple of π. Therefore, the angular equation of the nth bright fringe at infinity will be

$$w \left(\beta, \gamma - k'' \frac{\lambda}{p} \right) - w \left(\beta, \gamma - k' \frac{\lambda}{p} \right) = n\lambda.$$

If only the two waves considered were present, an observer, placing his eye directly behind the grating, would see an interference pattern whose nth bright fringe on an infinitely distant screen would have the equation

$$w \left(Y, Z - k'' \frac{\lambda}{p} \right) - w \left(Y, Z - k' \frac{\lambda}{p} \right) = n\lambda. \tag{14}$$

This is the fundamental functional equation of the grating interferometer.

The Passage from Geometrical Optics to Wave Optics

Let us now assume that the spacing p is very large with respect to the wave length. In this case $k'\lambda/p$ and $k''\lambda/p$ are very small and we can expand the functions w of eq 14 in Taylor series, retaining

165

only the first derivatives. Thus we obtain

$$\frac{\partial w}{\partial Z} = \frac{np}{k' - k''}.$$

For two consecutive waves $k' - k'' = 1$, and we get

$$\frac{\partial w}{\partial Z} = np. \tag{15}$$

A comparison with eq 7 shows that the nth bright fringe has the same equation as the projection of the nth transparent groove of the grating; the dark fringes are therefore coincident with the shadows of the opaque groove of the grating.

This result is valid for the interference of two consecutive waves. But we can easily extend it so as to take into account all the waves diffracted by the grating. Indeed, we can divide the entire set of diffracted waves into pairs of consecutive waves, each pair forming its dark fringes on the geometric shadows of the opaque grooves. By superimposing the interference patterns of all these pairs we get evidently as a net result the shadows of the grooves.

In this manner, although we started from wave optics, we have accounted for a purely geometric phenomenon. We can therefore utilize the results of our discussion of the shadow of a wire. The patterns of figures 11.1 now become, respectively, the patterns of figure 11.2.

Let now the spacing p become smaller, so that a better approximation is needed. In the Taylor expansion of the functions w of eq 14
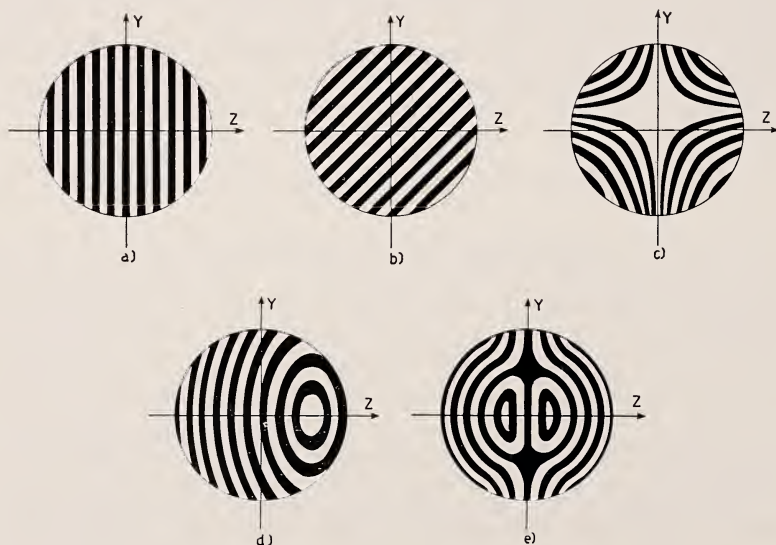
FIGURE 11.2.   *Interference patterns with a low-frequency grating.*

a, Perfect spherical wave; b, astigmatism; c, coma (meridional plane parallel to the lines of the grating); d, coma (meridional plane perpendicular to the lines of the grating); e, spherical aberration.

one must retain also the second derivatives. Assuming for the sake of simplicity $k'-k''=1$, one obtains

$$\frac{\partial w}{\partial Z} - \frac{k'+k''}{2} \frac{\lambda}{p} \frac{\partial^2 w}{\partial Z^2} = np. \tag{16}$$

Let us put

$$\frac{k'+k''}{2} \frac{\lambda}{p} = \delta. \tag{17}$$

We see at once that in the present approximation eq 16 can be written

$$\left[\frac{\partial w}{\partial Z}\right]_{(Y,Z-\delta)} = np, \tag{18}$$

where the left side is $\partial w/\partial Z$ computed at the point $Y, Z-\delta$. It is easy to interpret eq 18 if we notice that δ is the arithmetic mean of the angles of diffraction $k'\lambda/p$ and $k''\lambda/p$ of the two waves k' and k'', respectively. A comparison of eq 18 with eq 15 shows that the entire pattern has been shifted by the amount δ along the Z axis with respect to the preceding approximation. This happens because the scattering of the different waves with respect to the central wave has now become apparent and the field of interference of the waves k', k'' is centered around the point $Y=0$, $Z=\delta$. The situation is illustrated in figures 11.3a, 11.3b, 11.3c, 11.3d, and 11.3e, each case corresponding to that designated with the same letter in figures 11.1 and 11.2. Apart from the shift, the patterns are still identical to those of the shadows of the grooves. Of course, they are repeated many times; once for each possible pair of waves. If the successive interference fields partially overlap, the result is a very complicated phenomenon of multiple interference. However, in figure 11.3 we have represented the case where p is already sufficiently small for being out of this intermediate stage of confusion.

Finally we pass to a third approximation. First we shift the origin to the point $Y=0$, $Z=\delta$; owing to the definition (17) of δ, eq 14 is replaced by

$$w\left(Y, Z+\frac{k'-k''}{2}\frac{\lambda}{2}\right) - w\left(Y, Z-\frac{k'-k''}{2}\frac{\lambda}{p}\right) = u\lambda. \tag{19}$$

We put $k'-k''=1$, and

$$\epsilon = \frac{\lambda}{2p}. \tag{20}$$

Then we expand the functions w of eq 19 in a Taylor series, up to the third derivatives inclusive. We find thus

$$\frac{\partial w}{\partial Z} + \frac{\epsilon^2}{6} \frac{\partial^3 w}{\partial Z^3} = np. \tag{21}$$

The new feature with respect to eq 15 is represented by the term containing $\partial^3 w/\partial Z^3$. This term has no influence when only such aberrations are present as are described by eq 2, 3, and 4, where no power of γ higher than the second appears. In the casd of eq 11, the term
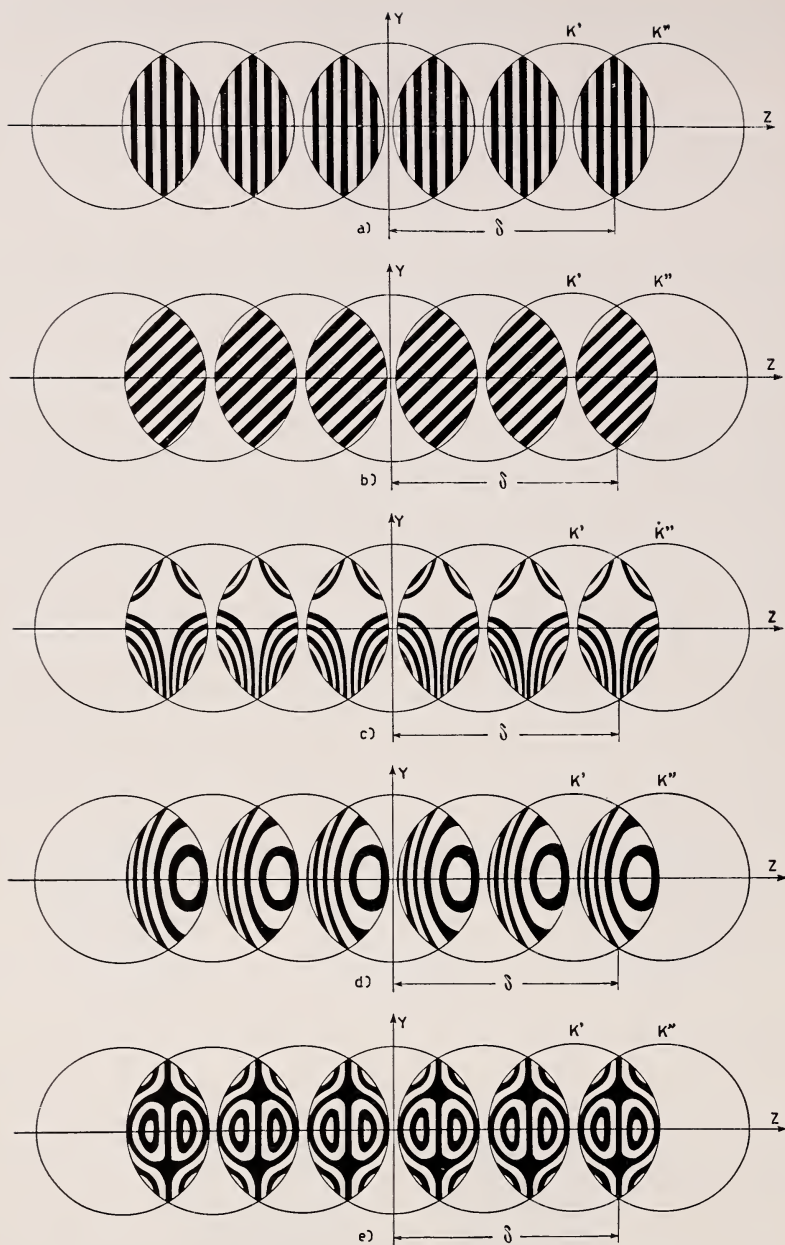
FIGURE 11.3. *Interference patterns with a high-frequency grating.*
a, b, c, d, and e, correspond respectively to the same cases of figure 11.2.

$c\gamma^3$ would only bring a constant in the left side of eq 21; this would amount to the same as changing n, i. e., the order of each fringe, and would by no means alter the structure of the pattern.

We conclude that, apart from the lateral shift, the interference patterns of the third-order aberrations up to coma are identical to the corresponding shadow of the grating, repeated as many times as there are pairs of consecutive waves. This result evidently holds good in any further approximation, depending on higher derivatives of w.

As regards spherical aberration, insertion of (5) into (21) brings to

$$Z[x + 4s\epsilon^2 + 4s(Y^2 + Z^2)] = np. \tag{22}$$

A comparison with eq 13 shows that the difference is represented by the constant term $4s\epsilon^2$ in the square brackets. This term has the same effect as a change of origin for x. It follows that in the present approximation we have exactly the same patterns as in the preceding ones; but they occur at different locations of the grating along the axis of the wave. This result holds in any higher approximation.

Conclusion

We have thus given all the elements that are needed for a discussion of the Ronchi test as applied to third-order aberrations. We have purposely not dwelled on the many rules that can be given for the quantitative evaluation of actual aberrations by means of the interference patterns. These rules are very easy to derive once the basic equations given in the present paper are known and their meaning is understood.

We have shown that both the geometric and the interferential versions of the Ronchi test give rise to the same type of patterns; but this does not mean that they are equivalent. Of course the interferential one is much more sensitive and must be applied in all the cases where a high degree of accuracy is required. One must not be deterred from making use of it by the fear of all the fine adjustments that are inherent in every interferential test. Indeed, the grating represents the only interferometer that need not be adjusted; it could be termed a self-adjusted interferometer.

12. A Combined Test Procedure for Camera Lenses, and Photoelectric Examination of Intensity Distribution in Line Images

By Erik Ingelstam and Per J. Lindberg [1]

Lens Tester

The first part of this paper deals with a simple and convenient instrument for testing lenses, which we believe to be novel. By means of this instrument it is possible to determine the tangential and radial resolving power (estimated from a photographic negative), the corresponding image surfaces, and the distortion, by a method that is rapid and inexpensive. For a more complete description reference should be made to an earlier report from our laboratory [1].[2]

Apparatus

This instrument is shown in figures 12.1 and 12.2. The characteristic feature is a collimator provided with a "tilted test plate" making an angle of 60° with the focal plane and shaped as shown in figure 12.2. Its function is evident from the figure. An advantage of the method is that the test is made with the lens mounted in the camera to which it is fitted, with its own film holder and in conjunction with a filter if desired. The test can be made at different field angles, β, from the axis and on the resulting negative one can read the point of greatest sharpness. This point is related to the distance from the emulsion to the position of greatest sharpness by the quadratic equation shown in figure 12.1. The patterns of line groups on the plate shown in figure 12.2 have been distorted to produce the known square contour of the Foucault groups on the test negative as shown in figure 12.3. The determination of the x positions is very rapid, and in the legend comments are made concerning the presence in this case of one region of best resolution and one of best contrast as discussed by Kingslake in this symposium.

This type of test plate has been of good use for different purposes. The requirements of the Royal Swedish Air Force brought about the construction of most of the apparatus reported here, and it has been used for most of the qualification tests of new lens designs as well as for a routine collimator to be used in the over-all program of the Swedish Air Force for controlling and adjusting the cameras for air reconnaissance. The authors are indebted to O. Hagsten, Chief of the Photographic Department of the Royal Swedish Airboard and his staff, for excellent collaboration and advice concerning the many practical problems in aerial photography.
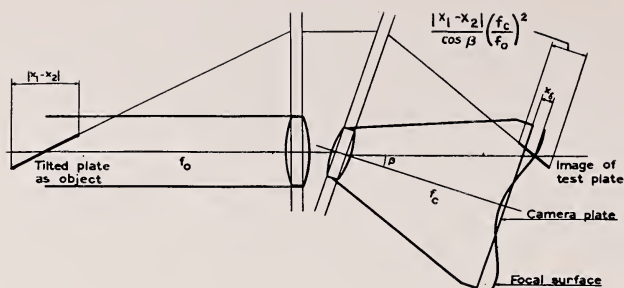
FIGURE 12.1. *Arrangement with the "tilted test plate" in the collimator.*
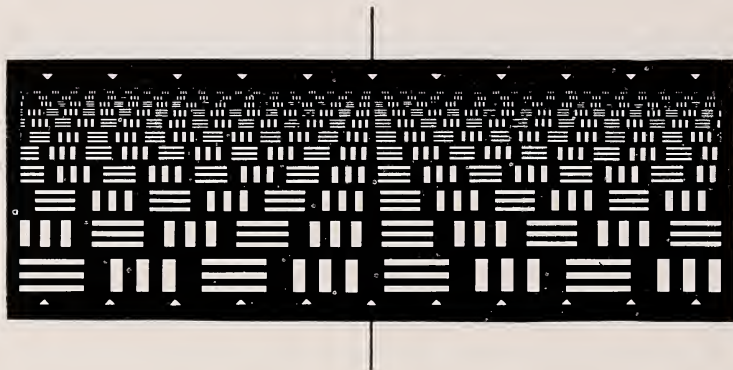

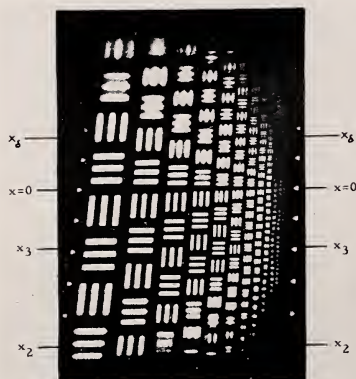FIGURE 12.2. *Shape of the "tilted test plate," the consecutive line groups being in the scale ratio of $2^{1/8}$.*

FIGURE 12.3. *An enlarged record of a tilted test plate with interrupted line groups.*

The smallest groups are best resolved at $x\delta$, which coordinate indicates the intersection between the test-plate image and the camera plate, so the focusing defect is the distance between $x\delta$ and $x=0$.

At the point x_2 out of focus there arise line groups with reduced number of lines, which is sometimes a form represented in the record. At x_3 large groups are imaged without halation, and at first one would say that the best sharpness is situated there, but $x\delta$ indicates the best "micro-sharpness."

172

Another application of the tilted test plate is the testing of the optics of fluorographic cameras, which is reported elsewhere [2].

In order to obtain the distortion of a camera simultaneously with the test of its resolving power, we have made use of a support, of which the main features are shown in figure 12.4. Not having the several fixed collimators used in the larger laboratories, we turn the camera support around an axis, the angles being read with the precision of a few seconds of arc. Besides the plate (fig. 12.2), the collimator is provided with indices in the focal plane designed to photograph well with different resolutions. In the series of records taken as the camera is turned to equally spaced fixed angular positions, the positions of these indices can be exactly measured in the comparator. In order to avoid reading the micrometers on the precision scale every time, the support has fixed angular positions, maintained by the ball and spring positioner shown in detail in figure 12.5. It has been proven that the positions determined by a ball-bearing ball positioning into conical holes are, because of freedom from hysteresis, reproducible within less than 1 micron. Therefore, in certain cases, when it is not necessary to know the absolute value of the distortions, but
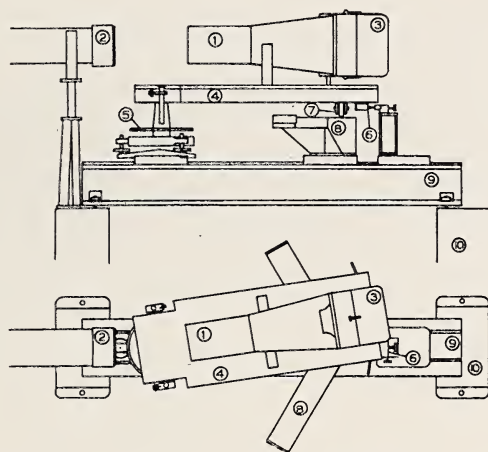
FIGURE 12.4. *Side and upper view of rotatable camera support.*

1, Camera-lens system; 2, collimator lens; 3, film or plate holder; 4, rotatable camera support; 5, goniometer precision scale; 6, ball and spring positioner; 7, ball bearing; 8, V-shaped path for the ball bearing; 9, optical prism bench; 10, box-like base.

FIGURE 12.5. *Ball and spring positioner.*

only relative displacements across the image field, these positions form the necessary fixed points. This is the case when lateral chromatic aberration is examined; parallel adjacent rows of records are taken with six different interference filters inserted between the test plate and the light illuminating the collimator, the successive rows being displaced by slightly tilting the collimator. A cross-table measurement of the indices obtained in this way is sufficient to evaluate this type of aberration. The corresponding axial chromatic aberrations in the radial and tangential directions are, of course, at the same time determined from the records of the tilted test plate.

It should be clear that all photographically achievable information about a camera can be recorded on one single photographic plate by means of this procedure. The values of the resolving powers obtained are relative ones photographically recorded and visually estimated. Because of this lack of information the instrument described in the second part of this paper was constructed.

Image Scanner

In our laboratory we have found it valuable also to construct equipment for measuring intensities in optical images, and accordingly, a photoelectric scanning device has been constructed. The arrangement is suitable for measuring across a line, as the light passes through a linear slit to the phototube. The sensitivity is sufficiently great to permit the use of a pinhole aperture, e. g., of a few tenths of a micron in diameter if this should be necessary, but the "indefinitely long" linear slit is appropriate for most purposes. Of former constructions of this class the interesting method of Jones and Wolfe [3] may be mentioned. The intensity distribution perpendicular to the slit was obtained by photographing through a gray wedge, thus obtaining logarithmic curves that gave much information about image intensity distribution in a lens system. Further, Herriott [4] has devised a registration device utilizing a photocell, its current being coupled to the y-coordinate of an oscillograph whose x-axis was synchronized with the movement of the entrance slit across the image, the movement being maintained by the electromagnetic system of a loudspeaker. In this way the oscillograph directly shows the intensity curve. Surely, other laboratories in optics must also have devised similar equipment utilizing photoelectric measurements of aerial images.

The main characteristics aimed at in our construction are : (1) Wide range of sensitivity with true response (10^6), which makes it necessary to have the power supply for the photomultiplier tube, the amplifier contained in the system, and the recorder, all well and precisely constructed and quickly adjustable by switches to different ranges of sensitivity; (2) freedom from extraneous light in the system itself, a feature that is absolutely necessary in order to cover the large range to which reference has been made (achievable only by avoiding the use of any type of subsequent magnifying system after the image, such as a microscope); (3) very narrow entrance slits, as we still want a true shape of the intensity function perpendicular to the slit.

Photocell carriage. The photomultiplier tube is an RCA 931–A, its mounting with base above being shown in figure 12.6. The adjustment of the device before inserting the phototube is obtained by the
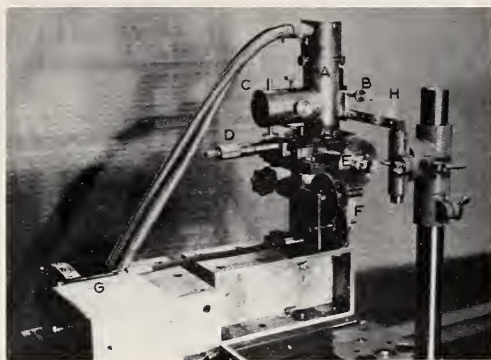
FIGURE 12.6. *Photoelectrical unit.*

A, Tube housing; B, screw for slit adjustment; C, tube for insertion of adjusting microscope; D, microme -
ter screw for axial slide displacement; E, micrometer screw for transversal slide displacement; F, gear box for
E screw; G, synchronous motor; H, glass under test.

microscope inserted into a tube at the rear. The upper part of the
carriage is adjustable in all directions; if the slit is to traverse a plane
which is obliquely cut by the light rays, the tube housing can be
turned before clamping. A micrometer serves for the displacement
in the axial direction. A precision advance laterally is obtained
by a micrometer screw pushing a slide. The synchronous motor is
coupled to the gear box giving to the slide a variety of speeds easily
changed during the recording. The most commonly used speeds are
25, 5, 1, and 0.2 microns per second. Of course, everything must be
very substantially built in order not to be disturbed by occasional
vibrations in the building during the recording.

Production of the slit. The slits are formed by shadowed metal evap-
oration onto glass plates. Before evaporating the aluminum the in-
tended slit area is covered by a quartz thread drawn as described in
Strong's Handbook [5], and fixed in position by tapes. A selection
of such slit plates from 1-micron slit width up to about 10 microns for
use as will be reported later was prepared. Made by this method they
are precisely uniform, the measurement of their widths being deter-
mined by viewing the positions of their Fraunhofer diffraction minimas.
One eye looks at a slit illuminated with monochromatic light, the other
eye observes where the diffraction minima fall on a metric scale. In
this way the uniformity of a slit is rigorously controlled, and a precise
measure of its width is achieved. The slit side of the glass plate is
always turned toward the incident light. Between the slit plate and
the photomultiplier tube a diffiusing screen is mounted to equalize
the intensity over the photosensitive area thus securing a constant
gain.

Electric equipment. The potential applied to the multiplier tube is
supplied by a rectifier and stabilizing unit shown in figure 12.7 as P.
It gives variable dynode voltages, of 50, 70, and 100 volts. The photo-
multiplier current is amplified by a balanced d-c amplifier, A, with
decade input resistances, making it possible to vary the sensitivity in
decades by means of simple switching. The output of the A unit
leads to the Speedomax recorder, ordinarily adjusted, however, for a
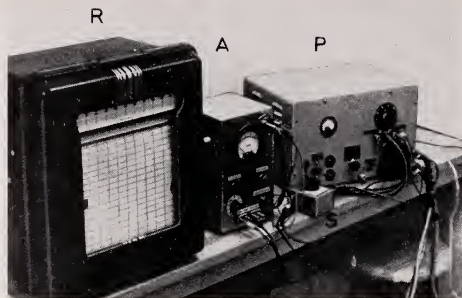balancing time of 1.5 seconds.

FIGURE 12.7. *Electric equipment.*

P, Power supply; S, switch; A, amplifier; R, recorder.

Records of test line figures, etc. We need not go in any detail into the troublesome questions related to resolving power in this symposium. It should be sufficient to point out that most of the practical men dealing with and buying photographic lenses still have nothing more than a specification that the resolving power is e. g., 38 lines/mm at the image center, though we know that such a specification is so incomplete as to have little value because it is based on a subjective estimation of the limit of visual recognition of structure by means of magnification on a photographic film, possibly also with unknown graininess. In view of the large amount of basic research by Selwyn, Arnulf, Coleman, and others we feel that there is now a great need to base measurements upon real physical quantities and to separate a *physical resolving power* which only deals with the intensity distribution in the aerial image.

Even in this idealized case of an infinitely narrow long object slit the specifications are, as well known, not unambiguously defined.

In spectroscopy the conventional measure of resolving power is, in spite of criticism, the well-known Rayleigh criterion. The application of this criterion in spectroscopy generally requires that the natural line width be substantially infinitely small. Therefore, very narrow spectrum lines and appropriate slit widths are used. The criterion corresponds to a drop in the intensity curve between the two maxima of 19 percent, this figure being partially conventional. As is well known, the eye easily resolves closer line objects, which often leads to confusion. The Sparrow criterion [6] which demands that between two or more lines, no drop shall be present in the intensity curve, or the slope shall be zero at one intermediate point, thus defines a limit of resolving power less severe. The criterion has also some further advantage.

When one tries to define physical resolving power for line images for photographic and similar optical apparatus, it seems to be proper to consider the infinitely narrow line. From this ideal object, by integration, one can derive the common test figures with a rectangular intensity shape, as the Foucault targets, and also the desirable Selwyn test objects with a sinuoidal intensity function. This is possible by means of known mathematical procedures [7, 8]. The criterions mentioned, however, relate to the shape of the curves at the top, or to the slope, and they are insensitive to the presence of stray light.

176

The stray light is taken into account in Selwyn's tests but only when the tests are performed in such a manner that the whole area or a known area of the light field is occupied by the test figures, which in practice is often neglected. The test figures with rectangular shape and decreased contrast now often in use are no doubt valuable for routine practice, but there is a link missing between what they give and the problem of the influence of stray light on a physically ideal image.

If one can obtain a well-defined physical conception of the resolving power, and by means of recording devices such as the one reported here analyze experimentally the true intensity shape of lines, it would be possible to relate this physical resolving power to visual resolving powers and the photographic resolving power under the various circumstances as investigated by Arnulf, Selwyn, and others. In the general case it demands recalculations which may possibly be inconvenient, but there is still a large interest in exactly knowing the intensity distribution in one-dimensional optical images. In designing and constructing this highly precise image scanner, the chief purpose has been to provide a means for making objective physical measurements of resolving power and contrast. The experimental results here presented illustrate its suitability for this purpose.

Special Records

Experimental results. Several series of records have been made with test plates of the ordinary rectangular shape in our device. Figure 12.8 shows records taken with lines in groups of three in geometric progression. They were taken with commercial photographic objectives. Without any large uncertainty, the resolving power according to the Sparrow criterion can be determined from the first group where the undulations at the top disappear. It must, however, be understood that one deals here with broad lines, their widths being equal to the space between them, and not with an ideally narrow line.
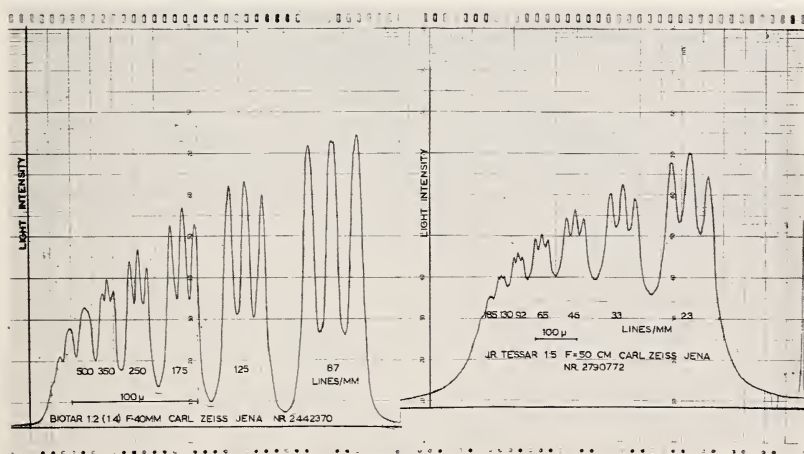

FIGURE 12.8. *Records of three-line Foucault groups.*

Accordingly, a recalculation of the ideal narrow line is necessary; such integration procedures being possible by means of mechanical or optical devices. Collimators with different focal lengths were used, the objects being placed in their focal planes. When the focal length of the objective to be tested was short, the object was placed about 15 meters distant without a collimator. In these records white light was used.

A better solution for the study of a physically ideal line is achieved by examining a single line imaged by the lens under test. A special purpose in taking up this research is the examination of the influence of poor polish of glass surfaces on the intensity distribution.

When examining the image intensity distribution in the outer part of the line image where the intensity has dropped to 10^{-5} or 10^{-6} complete freedom from scattered light and other disturbances must be achieved. For example, in the plate forming the entrance slit of the photocell there may be some very small pinholes in the evaporated layer that permit the passage of stray light. In order to be free from this, an additional opaque screen is placed between the slit and the source with a slit approximately 20 microns wider than the definitive slit.

The apparatus is arranged for this purpose as shown in figure 12.9. The object slit consists of a uniformly broad straight slit. It is illuminated by a monochromator of special type. This monochromator consists of a plane grating and a lens, this device imaging a monochromatic part of the light from a ribbon-filament lamp onto the slit. This arrangement provides many times the illumination that is provided by interference filters, which have also been used. The lens whose intensity distribution was primarily studied was a plano-convex lens of $\dot{F}=1,000$ mm and a diameter of 45 mm, the edges being shielded to prevent reflection of any noticeable amount of light.

The special investigation taken up is the influence on the intensity distribution of different polishes on a glass surface situated near the focal plane. For this purpose two nearly equal thick plano-parallel plates, one of smooth, one of rough polish, were subsequently placed as shown in figure 12.9. Because of the short distance between surface and focal plane, only about one-hundredth of the focal length, a bad polish causes mainly stray light; when the distance is larger, the
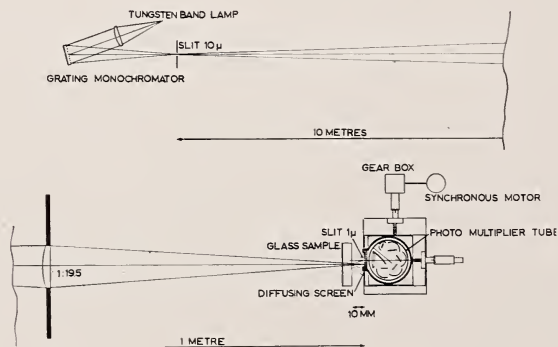

FIGURE 12.9. *Apparatus outline.*

central part of the image will be affected causing loss of resolving power. The additional spherical aberration caused by the plate, as well as the original spherical aberration, is of minor importance for the total shape of the curve and can be calculated with sufficient precision. The area where the rays traverse the glass plates is precisely determined by records on photographic paper to permit their examination with multiple beam interferometry; they were in advance selected by examination in a Michelson interferometer and had sensibly perfect planeness and homogeneity.
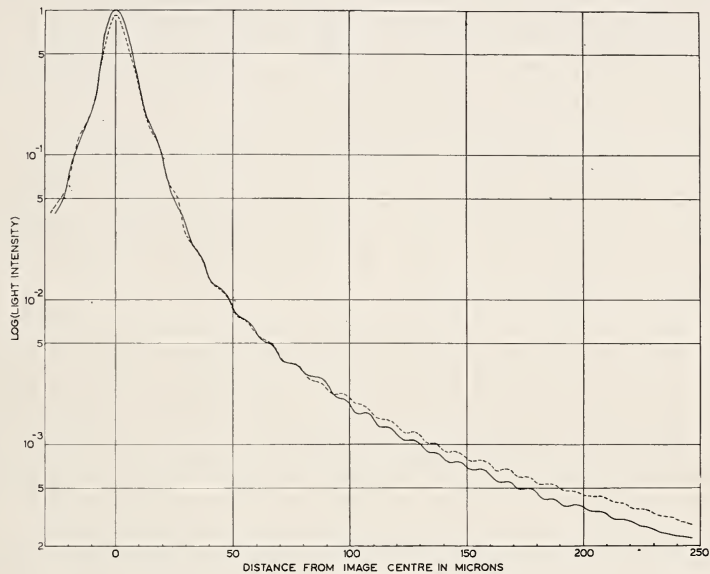
FIGURE 12.10.   *Intensity distribution.*

Solid curve, fine-polished glass; dotted curve, rough-polished glass.
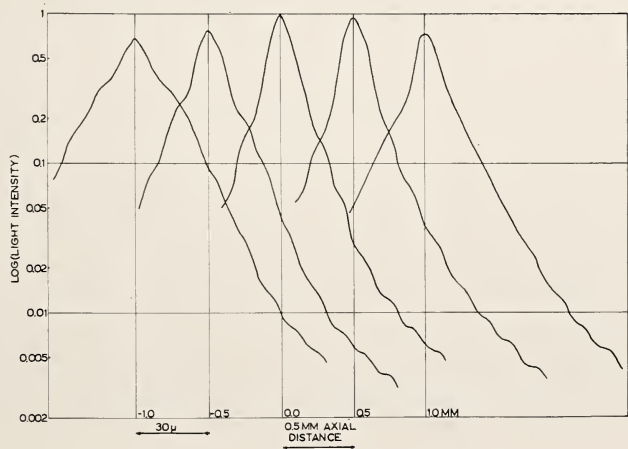
FIGURE 12.11.   *Different axial cuts of image.*

179

Figure 12.10 gives two typical curves for the central part of the image transformed to a logarithmic scale of intensity. The x-coordinate is linear. An interesting feature is the true record of the outer parts of the image, the undulations up to about the 20th being clearly recorded. A test of the correct performance of the carriage micrometer drive is afforded by the coincidences of the fringes. In figure 12.11 the central image is represented, down to about $10^{-2.5}$ of the maximum intensity. The different curves are displaced laterally in this diagram for different positions along the axis. There are different parallel sections of the ray about the plane of the best focus. Figure 12.12 shows an original record for illustrative purposes. The sensitivity range has been changed three times, namely, in the ratios of 10, 10, and 19, the exact values of these ratios having been determined separately. The fluctuations of the last record is due to electronic noises and thus limit the accuracy.

In order to pass over to still larger angles, the intensity was increased about 100 times by widening the slits. This affords less detail and accuracy in the central part, which is of no great importance as this second record covers a region extending some hundreds of microns beyond the first record. These recordings have been extended to approximately 4 mm from the center, thus covering scattering angles up to about 20 degrees.

In figure 12.13 the corresponding curves for the best focal position in the two cases, with high-polished glass and with rough-polished glass are traced, logarithmically in the x-direction. The curves without glass are, within the accuracy, close to those for the highly polished glass. A comparison shows that the centers are somewhat changed in the sense that the rough-polished glass gives a lower intensity. The outer regions are still more different, the light spread resulting from the fine structure of the wavefront due to the undulated profile curve is quite evident and well measurable within desired accuracy.

Quantitative relations between the amount of stray light and the shape of the wavefront as studied by means of the multiple-interference or phase contrast may, we hope, be established when more data are gathered and treated in detail. The interest of such an investigation seems to be indicated by, on one side, the possibility nowadays, thanks to phase contrast and other methods, to examine precisely
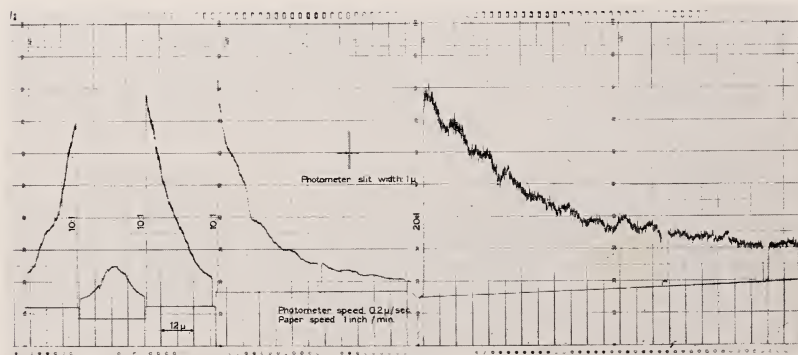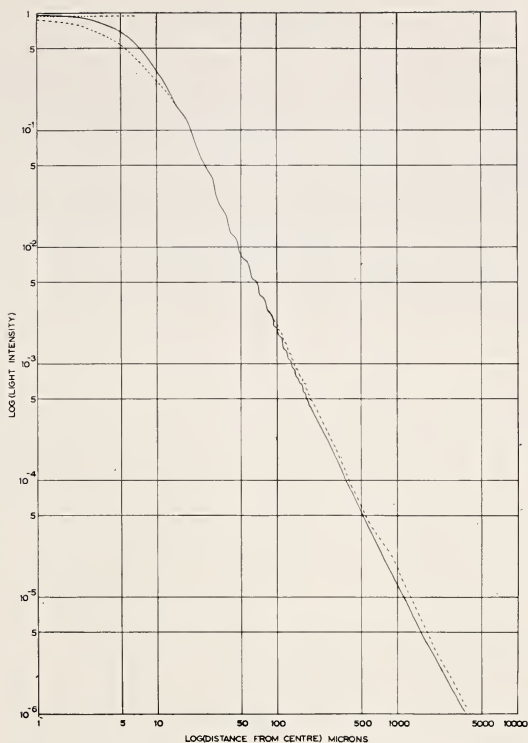

FIGURE 12.12.   *Original record.*

FIGURE 12.13. *Intensity distribution far from the center.*
Solid curve, fine-polished glass; dotted curve, rough-polished glass.

such fine structure of wavefronts, and, on the other side, the troubles in some instruments due to scattered light for the reason mentioned. Illustrative examples to be studied are the front lenses of microscopes, and spectroscopic instruments, where this type of stray light is often known as *foot light*. A recent investigation with phase contrast on such types of irregularities in the wavefront [9] should be supplemented by recordings of the intensity distribution as performed here.

The experimental procedure used here is a means for studying the real performance of a lens system comparing the intensity with that predicted from theoretical calculations. The curves shown here have not yet been analyzed in this way, but the slope of the curves, figure 12.13, seems to agree with the corresponding functions in this region (W, more than one hundred diffraction units) remote from the line center.

Note added in proof (Oct. 1953): A subsequent development of this scanner is reported in P. Lindberg, Measurements of Contrast Transmission Characteristics in Optical Image Formation, Optica Acta (in press).

181

References

[1] E. Ingelstam and P. J. Lindberg, Rep. Lab. Optics; Roy. Inst. Technol. (Stockholm) No. 9 (1951).
[2] E. Ingelstam and P. J. Lindberg, J. Opt. Soc. Am. **41,** 346 (1951).
[3] L. A. Jones and R. N. Wolfe, J. Opt. Soc. Am. **35,** 559 (1945).
[4] W. Herriott, J. Opt. Soc. Am. **37,** 472 (1947).
[5] J. Strong, Modern Physical Laboratory Practice (Blackie & Son, Ltd., London, 1938).
[6] C. M. Sparrow, Astrophys. J. **44,** 76 (1916).
[7] A. Gray, G. B. Mathews, and T. M. MacRobert, A Treatise of the Bessel Functions and their Applications to Physics (McMillan Co., London, 1939).
[8] G. Lansraux, Rev. Opt. **26,** No. 24, 278 (1947).
[9] E. Ingelstam and E. Djurle, Arkiv Fysik **4,** No. 27 (1952).

13. Measurements of Energy Distribution in Optical Images

By R. E. Hopkins,[1] Howard Kerr,[1] Thomas Lauroesch,[1, 2] and Vance Carpenter [1, 3]

Introduction

The Photographic Branch of the Signal Corps is constantly faced with the problem of how to write specifications for lenses to be used in their various applications. They were well aware of the shortcomings of the standard methods of specifying a lens' performance, and requested that our laboratory study the problem and recommend testing and specification procedures. In contrast to the Air Force they were primarily interested in lenses of moderate focal length ranging from 1 to 20 inches.

The first part of this study included an investigation of the literature on the subject. One point that seemed to crystallize out of the study was that most of the work had been done by users of equipment. The tests performed were usually designed to measure the final performance of the lens, and most of the interest was centered around the validity of the testing procedures.

During and after the second war several workers began describing methods for testing the lens itself [1,2,3].[4] Jones and Wolff [4] described a method for determining the distribution of energy in an image with a photographic process. Herriott [1] described a photoelectric method for measuring the intensity contour of a line image. Hansen [5] performed an important basic experiment. He studied an Elmar 8-cm focal-length lens at several apertures by photographing a periodic grating and a transparency of a building. He found that the lens resolved the best when used at the full aperture of f/3.5. On the other hand, observers when asked to pick the sharpest picture of a building selected pictures taken at f/6.3. The spherical aberration of the lens was measured on an interferometer, and showed that the minimum diameter of the blur circle at f/6.3 was less than for f/4.5.

Prior to reading this article Dr. Robert Wolfe at Eastman Kodak suggested to us during a conference a similar approach to the problem of image evaluation.

S. Huber [6] wrote a paper in 1943 that described a logical approach. He designed and constructed a lens with known spherical aberration, and calculated the distribution of light in the image. He then took photographs of a test chart and correlated the results. He found that the resolving power could be related to the diameter of the central core of the image that contained 25 percent of the total light.

It was clear from the trend of papers that more information on the image energy distribution of lenses was needed. It was also evident that purely physical measurements on the image would not be sufficient. A physical measurement would answer many of the problems of resolution and contrast but, in order to set up tolerances on energy concentration, experiments like Huber's and Hansen's would be needed.

With this background it was decided that the problem would be studied as follows. (1) A purely physical method for measuring the characteristics of the lens would be developed; (2) studies similar to Huber's and Hansen's would be made to relate image characteristics with performance on film.

Another point that came out of the literature was that many of the papers on the subject dealt with the problem of how to treat unsymmetrical off-axis images. If one accepts the concept that the object is made up of a series of points and that it is the function of the lens to form a faithful image of the point, the problem then is to measure how well the lens does form a point image and no particular weight should be given to structure of the image in anything but a radial direction. Therefore, it was decided to measure only the radial energy distribution. This approach is probably valid only if the objects being photographed have random orientation of all edges. It is obvious that if one intends to photograph railroad ties or picket fences then radial and azimuthal distribution should be known.

Methods for Measuring Energy Distribution

Interferometer

The first approach was to use the interferometer. The instrument and testing procedures have been adequately described in the literature [7, 8]. In general, the interferometer has been rejected as a testing instrument on two counts, which are believed to be not entirely justified.

Several authors have described detailed methods for analyzing an interferogram to determine the coefficients of the various Seidel third-order and high-order coefficients. These methods are difficult and time consuming. However, if one merely uses the interferogram to determine the radial energy distribution the interpretation is simple and straightforward. For example, figure 13.1 shows the interferometer pattern of an f/2 photographic lens. The radial energy distribution of energy can be determined by marking off equal slope contours. Measuring the areas included within the contour lines and dividing them by the total area gives a radial energy distribution plot. This can be done accurately if necessary, but for many purposes merely a rough estimate of area is sufficient. For example, figure 13.1 shows at a glance that there is considerable vignetting at 20° off-axis and that only about 30 percent of the energy is well concentrated into a spot.[5]

The other complaint about the interferometer is that it is too sensitive. This is not justified if one attempts to gather no more information than other less sensitive devices like a lens bench test. The interferometer is sensitive because it is capable of giving detailed

---

[5] Coleman [8] made the valuable suggestion of using the area of the largest inscribed circle free from fringes as a measure of quality.

PARAXIAL FOCUS   INTERMEDIATE FOCUS   5° OFF AXIS
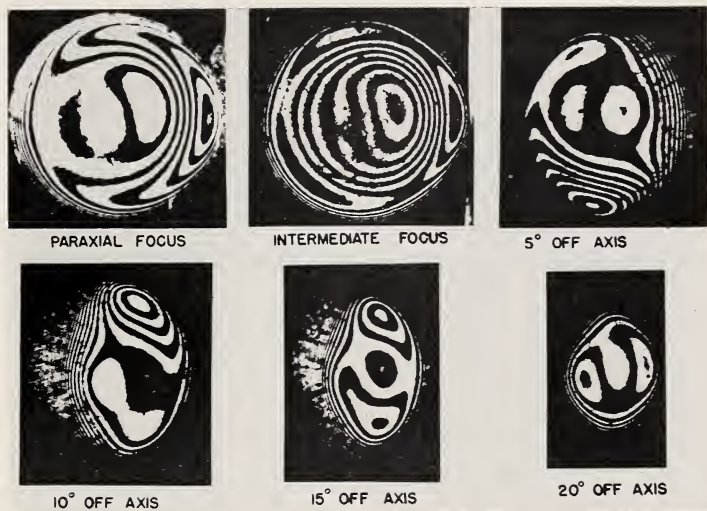
10° OFF AXIS   15° OFF AXIS   20° OFF AXIS

FIGURE 13.1. *Interferogram of F/2–50-mm focal-length photographic lens.*
Wave length 5461 A.

information much of which is not needed for analyzing a photographic lens.

The interferometer is the perfect instrument from a lens designer's viewpoint.   It dissects the image for him.   It gives him a quantitative picture of the wave front.   It gives him a feel for the design.

It does have the following disadvantages: (1) It does not indicate directly the effect of diffraction; (2) the interferogram does not show the effects of scattered light.

Further studies relating interferograms and the diffraction effects are needed in order to make the interferometer more useful.

Several measurements of energy distribution were made on lenses with the interferometer.   Reasonably good agreement was obtained between this method and a photoelectric method to be described. However, the measurements on the interferometer and the photoelectric device were made on separate pieces of equipment and it was not possible to be sure of the focus accurately enough to establish whether the differences were real or due to diffraction.

Photoelectric Method

There are several methods described in the literature for measuring the performance of a lens photoelectrically [1, 2].   Most of the methods described use a slit source and the image formed by the lens is scanned across the narrow dimension.   If one accepts the concept that the radial energy distribution is a measure of the performance of the lens, then the image of a point source may be photometered by allowing the light to pass through successively larger circular apertures.

The experimental arrangement to be described measures the radial energy distribution.   A schematic arrangement of parts is shown in
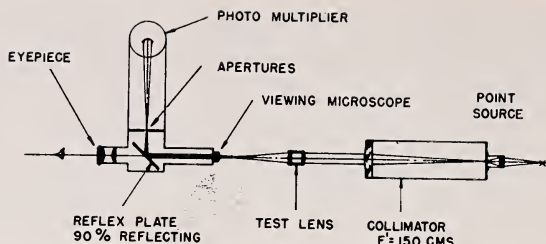
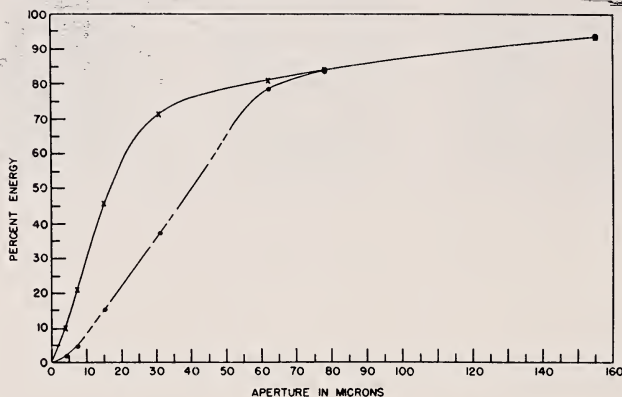FIGURE 13.2. *Schematic arrangement for photoelectric testing of photographic lenses.*

FIGURE 13.3. *Method of plotting energy distribution*

Lens name, Sonnar 50 mm; focal number, F/2; focal setting, best visual; filter, Wratten 58. $\times$ —, 0°; ● - - - , 15°.

figure 13.2. The lens to be tested is set up to form an image of a point source, imaged at infinity by a collimator. The image is then reimaged with a microscope objective onto an aperture slide. The aperture slide contains a series of small apertures that may be slid into the image plane. The light that passes through the aperture is received by a photomultiplier. A measurement of the energy distribution is made by sliding in the smallest aperture. The current through the photomultiplier is recorded. Larger and larger apertures are then slid into place until there is no further increase in photomultiplier current. The percentage of flux for any particular aperture is then obtained by dividing the current for each aperture by the current for the largest aperture. The data obtained for a lens is then plotted as shown in figure 13.3. The interpretation of this plot for the 0° image is as follows: 30 percent of the total energy in the image of a point source is contained within 10 microns; 80 percent is confined to 60 microns. Note that these curves are not intensity-distribution curves. In order to find the intensity at any radius it is necessary to differentiate the curve.

Experimental Details

1. The photomultiplier used was an RCA 931A. This tube is definitely red insensitive, but no attempt was made to balance the

186

photomultiplier with photographic response. Most of the measurements reported in this paper were made using filtered light. The light source was a 50-cp automobile-headlight bulb.

2. The photomultiplier was battery operated and was followed by a low-noise a-c amplifier. The final signal was measured with a Ballantine volt-meter. The light was chopped at 900 c/sec.

3. Special Bausch and Lomb [6] coated microscope objectives were used to view the image. These objectives were painted dull black and special care was taken in blackening the mount and spacing rings in order to eliminate scattered light.

4. Ground glass was used between the apertures and the photomultiplier.

5. The size of the holes and their apparent sizes when in use by the microscope objectives is shown in table 13.1.

6. The reflex viewer was used to provide a view of the image being photometered.

TABLE 13.1. *Size of holes in aperture plate versus size appearing in focal plane of lens being tested*

Apparent size with —	Size of hole (μ)						
	100	200	400	800	1600	2000	4000
16-mm objective_____(μ)__	7.8	15.5	31.0	62	124	155	310
8-mm objective_____(μ)__	3.9	7.8	15.5	31.0	62	78	155

Test on Reliability of Equipment

The linearity of the amplifier was checked electrically by putting in a calibrated electrical signal from a Jackson signal generator. The overall equipment was checked as a photometer in the following manner. A well corrected telescope objective was used as a test lens. The largest size aperture was then put in place. A green Wratten No. 58 filter was inserted in front of the point source. Neutral filters were then also inserted in front of the point source and their transmissions were measured with the equipment. The filters were calibrated previously using a Bausch and Lomb Martens photometer. The results are shown in table 13.2. The errors included in the table are the average deviation from the mean of four readings. The

TABLE 13.2. *Transmission measurement made with photoelectric equipment and compared with measurements made on a Martens polarizing photometer*

Martens polarizing photometer	Photoelectric photometer
%	%
78.2 ± 2	73.9 ± 0.7
62 ± 1	62.9 ± 0.4
46 ± 1	44.7 ± 1
36.6 ± 0.7	38.6 ± 1
28.9 ± 0.6	28.3 ± 0.1
25.3 ± 0.5	26.0 ± 0.4
18.5 ± 0.4	18.4 ± 0.3
14.9 ± 0.3	15.2 ± 0.4
9.1 ± 0.2	10.4 ± 0.6

[6] Recommended for our use by James Benford.

greatest discrepancy occurs at the high transmission values. The error may be due either to the Martens or the photoelectric reading. The largest discrepancy between the two methods amounts to 4 percent.

The complete reflex viewer and photoelectric equipment was checked by measuring the energy distribution of the image formed by a perfect lens. A well corrected telescope objective stopped to f/15 was set up directly on axis. Several attempts were made to measure the image from a point source and when compared with the Airy disk the measured image was always larger. Several refinements were made. The lens was centered more accurately, the tungsten source was replaced with a mercury source with a filter isolating the 5461 A line. The energy distribution was measured at several positions of focus. The pinhole was made as small as feasible. With all these refinements the agreement with the theoretical curve improved. The best agreement found is shown plotted in figure 13.4. The results are tabulated in table 13.3. The curve is made up of points using a 16-mm and an 8-mm objective. The agreement with the theoretical curve is only fair but we believe the difference is not due to an error in measurement. There are two reasons why the experimental curve could be broader than the theoretical curve. (1) There is scattered light. The lens is an air-spaced doublet. The surfaces were not tested but they may be imperfectly polished or imperfectly cleaned. (2) The air between the source and lens was not completely homogeneous. One could observe slight image boiling. This effect would broaden the image.

Attempting to fit the measured curve with the theoretical curve actually strengthened our confidence in the reliability of the equipment. Each improvement made in technique showed up as a better check. In order to get still better agreement it would be necessary to check the lens surfaces and probably repolish them. Also steps would have to be taken to ensure perfectly homogeneous air between the source and the lens.

Since we were interested in testing lenses and it was not known at the time what accuracy of measurement was needed it was decided not to check this point any further. As can be seen, there are differ-
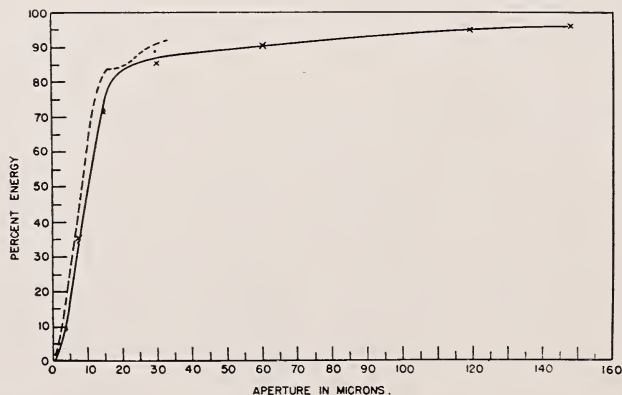

FIGURE 13.4. *Energy-distribution curve for an f/15 telescope objective.*

Lens name, Telescope Objective; focal number, 15; focal setting, best visual; filter, Mercury 5461. - - -, Theoretical; ———, experimental; x, 16-mm objective; ●, 8-mm objective.

TABLE 13.3. *Energy content for a perfect telescope objective*

The theoretical energy content for ξ point source on axis is compared with the measured energy content for a well-corrected F/15 (F′=35 cm) air spaced telescope objective. Monochromatic 5461 A light was used for this test.

Spot diameter	Energy		
	Airy disk theoretical	Measured with 8-mm objective	Measured with 16-mm objective
μ	%	%	%
3.75	14.5	9.6 ± 0.04	----------
7.50	47.2	35 ± .6	35 ± 1.5
15.0	83.0	73 ± .9	71 ± 1.5
30.1	90.8	89 ± 1.3	85 ± 1.0
60.0	94.0	92 ± 3.1	90 ± 2.0
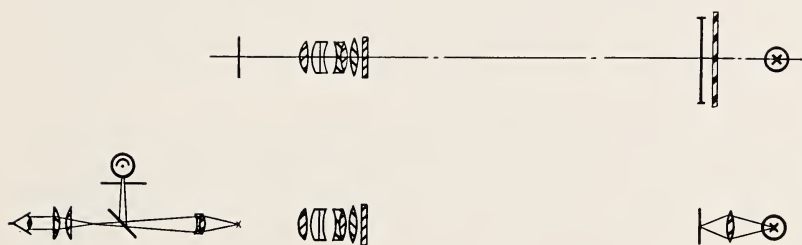
FIGURE 13.5. *Experimental setup for image evaluation.*

ences between the theoretical curve and the measured curve as large as 12 percent, but it is our belief that the difference is not entirely an error.

Image Evaluation Experiment

Following the suggestion of Wolff, Huber and Hansen the following experiment was performed to relate the energy-distribution curve with pictures taken on film. The basic experiment performed and the equipment is illustrated in figure 13.5.

The point source image of a 7-inch Aero Ektar lens was photo-metered. The lens was then shifted to a new position and a transparency of a scene was photographed. The film plane was located in the same position as viewed by the photometer. The scene that was photographed is shown in figure 13.6. Finally, on a separate film, a photograph of a $\sqrt[6]{2}$ line test chart was taken. In this manner the energy distribution in the image, the resolving power, and a picture of a scene were obtained for several focal settings and f-number.

Experimental Details

1. The 7-inch Aero Ektar was used because it provided a large variation of spherical aberration with change in f-number.

2. The scene was placed 8 feet from the lens, and it subtended 5°. The image of the scene on the film was 13 mm by 11 mm. The long focal length was used to obtain a reasonable size image on the film without using the lens more than 2½° off-axis. The image quality over the entire picture was then uniform.

189

3. The Aero Ektar was mounted on a Kine Exacta camera body. A special stiffened focal plane was built into the camera. For each picture the film was pressed against the film plane by a pressure plate that could be screwed down tight from the outside of the camera.

4. The lens was always focused at a definite focus position. The settings used were 4.5, 5.5, 6.5, 7.5, 8.5, 9.5, 10.5, and 11.5. The distance between each setting was 80μ.

5. Eastman Super XX film was used for all the tests. The Aero Ektar was provided with an accurately polished 4-mm thick Corning Sextant green filter. This filter has a peak transmission at 5200 Å and transmits 12½ percent at 4900 A and 5580 A. All of the picture-evaluation data was taken in green light with this filter.

6. A Wild[7], $\sqrt[6]{2}$ light lines on a dark background, high-contrast resolution chart was used for resolution data. A 1-inch clear area was used as a control area. The films were always developed to a gamma of .8 ±.1. For each focal setting several exposures were made. The resolution was read by at least two observers by viewing the negatives with a 40X microscope.

7. The scene shown in figure 13.6 was photographed with an 8 by 11 camera using a Bausch & Lomb 12-inch Protar. The positive transparency has a range of densities from .3 to 1.4 and a gamma of .8. U. S. A. A. F. charts were placed alongside both sides of the print. These were used as a check on the enlarging step. On the top and bottom of the scene grey scales were used to control the exposure and development. All the negatives were exposed and developed in

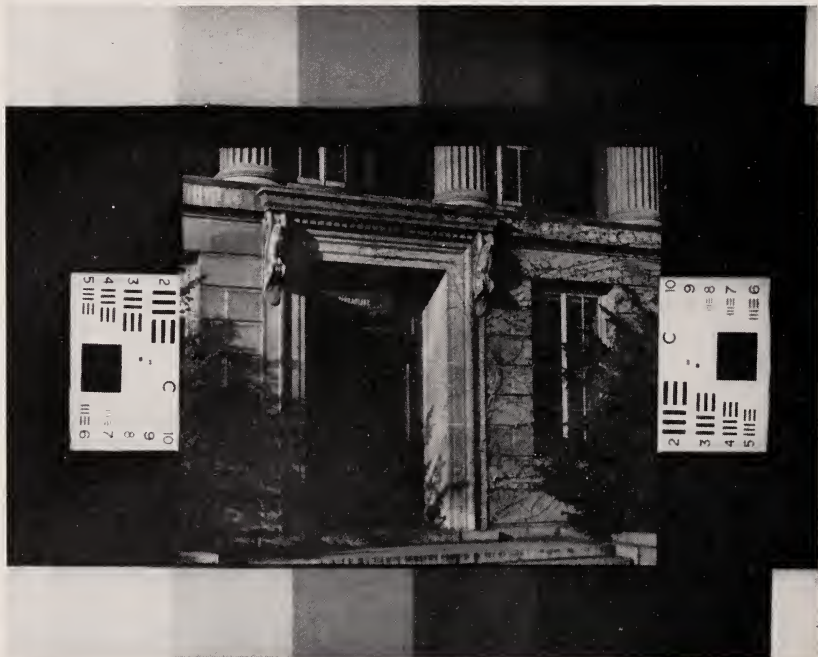[7] Supplied by Alan Murray, Bausch & Lomb Optical Co.

FIGURE 13.6. *Scene photographed for image evaluation.*

microdol to a gamma of .8 ±.03 with a maximum density of 1.8 ±.02.

The negatives were then enlarged eight times in a Leitz condenser enlarger, using a special lens loaned by the Signal Corps. This lens is a Zeiss Biotar 10-cm focal length f/4, stopped down to f/6.3. The lens is nearly perfect at this aperture ratio. The white-light image to the first bright ring measured to be 7μ.

The prints were made on Kodabromide F.3. Exposure and development were controlled carefully to make the reproduced grey scales alike. The grey scales were measured on the Capstaff reflection densitometer. Any picture that had a density step differing by more than ±.1 in the highlights and ±.02 in the shadows was rejected. In the first series of pictures made (f/4 set no. 1) the tolerance was ±.02 but the tolerance was relaxed in order to obtain some results. It was extremely difficult to make prints to fall within this tolerance by using carefully controlled methods. The method finally adopted was to make several prints and then select the ones that fell within tolerance.

An extensive series of tests were performed to determine the loss in resolution from the enlarging process. With a negative resolving power of 37 to 40 lines per mm the prints on the average resolved 30 to 32 lines per mm. With a negative resolving power of 20 lines/mm there was less than 1-line/mm loss in resolution due to the enlarging process.

The above series of experiments provided the following data for each position of focus of the lens: (1) Energy distribution curves; (2) lines/mm resolution for several exposures; (3) glossy prints of a scene.

Energy-distribution curves. Figures 13.7, 13.8, and 13.9 show the energy-distribution curves for aperture ratios of f/2.5, f/4, and f/5.6.

Resolving-power data. The resolving-power data is summarized in figure 13.10. The measured resolving power indicated by the dotted lines is plotted against positions of focus. The first results showed that the resolving power did not change with exposure if the density of the control area remained between 1.4 and 2.0. The resolving
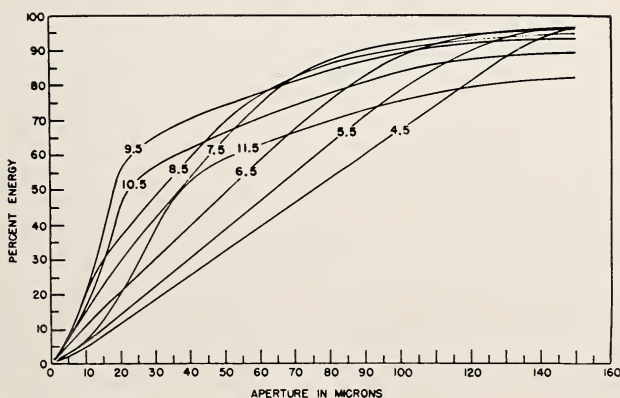

FIGURE 13.7.   *Energy-distribution curves for Aero Ektar at f/2.5.*

Lens name, Aero Ektar; focal numbers, 2.5; focal setting, 4.5 to 11.5; filter, Corning Green.
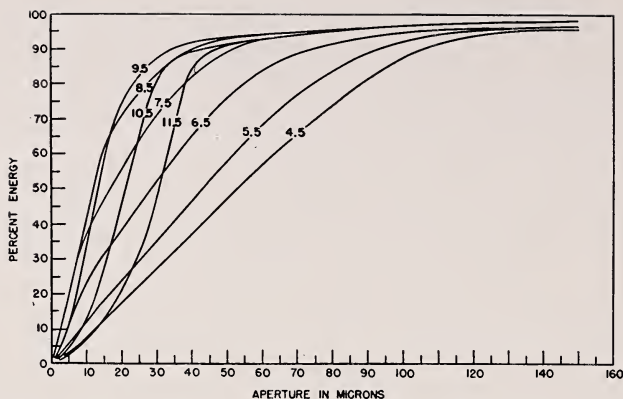
FIGURE 13.8. *Energy-distribution curves for Aero Ektar at f/4.*
Lens name, Aero Ektar; focal number, 4; focal setting, 4.5 to 11.5; filter, Sextant Green.
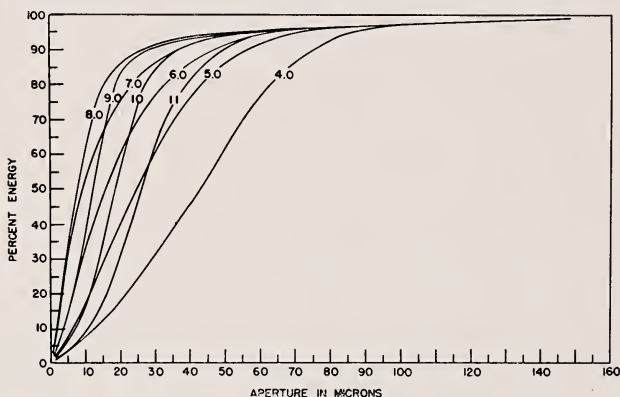

FIGURE 13.9. *Energy-distribution curves for Aero Ektar at f/5.6.*
Lens name, Aero Ektar; focal number, 5.6; focal setting, 4 to 11; filter, Corning Sextant Green.

powers were, therefore, averaged if they fell within this range. At least six independent film strips were used. It is well known that reading resolving power is difficult. The result that one obtains depends on the criterion used. The two observers who read the data agreed on a criterion and then they read resolving power to within an average deviation of not more than 3 lines/mm. It is our opinion that the *R. P.* figure given is subject to a large error, as much as 10 or 15 lines depending on the criterion used. The resolving-power data does, however, offer an interesting relative measure of the image quality.

The solid curves in figure 13.10 are resolving powers calculated from a formula based on the energy-distribution curves. At each focal setting the size of spot that contains 30 percent of the energy was determined. These spot sizes were used in the following formula to predict resolving power.

$$R.P. = 605/(d_i^2 + d_f^2)^{\frac{1}{2}}, \tag{A}$$

192

where

d_l = spot size for 30-percent energy in microns
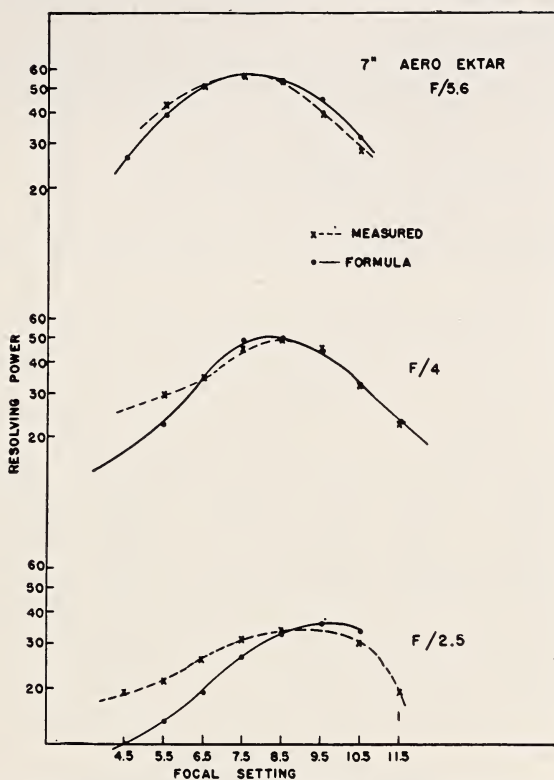d_f = spot size for film (10 microns for Super XX).

FIGURE 13.10. *Resolving power for several focus settings of the 7 in. Aero Ekton at f/2.5, f/4, f/5.6.*

The formula is empirical and is an oversimplification. It does, however, indicate that the resolving power is determined by a very small percentage of the energy. The agreement is good at the large focal settings but discrepancies show up at the small focal settings. At focal settings 4.5 and 5.5 the images contain small cores of light and a large flare. This small core of light accounts for the high resolution. This type of distribution is due to the spherical aberration in the lens. The effect is most pronounced at f/2.5, where the spherical aberration is considerable. Inspection of the energy-distribution curves indicates that less than 30 percent of the total energy is contained within the small core. If the above formula were to predict the high resolution shown at focal settings 4.5 and 5.5, it would be necessary to use a spot size that contained less than 30 percent of the energy, possibly as low as 10 percent.[8] It would then also be necessary to combine this spot size with one at a higher energy level in order to fit the curve at the higher focal settings. No attempt was made to obtain

---

[8] We should recall that Huber [6] found that 25 percent of the energy determined the resolving power.

better fittings for several reasons. The curves measured are not accurately determined at such extremely small images. At the 10-percent level, for example, the spot sizes run between 2μ and 14μ in diameter. There is also some question that the limiting resolving power is significant. Finally, much more data would be needed to obtain a reliable empirical formula.

Formula A is clearly incomplete because it limits the resolution to 60.5 lines/mm for Super XX. Resolution higher than this is possible. A much simpler formula using the first power is as follows:

$$R.P. = \frac{700}{d_e + d_f}. \tag{B}$$

A comparison between these two formulas is shown in tables 13.4,a, 13.4,b, and 13.4,c.

TABLE 13.4. *Comparison of calculated resolving powers, using formulas A and B*

d_1	A	B	Measured	d_1	A	B	Measured
(a) $f/5.6$				(b) $f/4$			
μ	Lines/mm	Lines/mm	Lines/mm	μ	Lines/mm	Lines/mm	Lines/mm
21	26	25	--------	33	17.5	17.5	--------
12	38.8	36.8	43.2	25	22.5	21.9	29.4
6	51.7	53.8	51.5	14	35.0	33.0	34.6
4	56.2	63.6	55.9	7	49.6	50.0	44.8
6	51.7	53.8	54.1	7	49.6	50.0	49.2
10.5	42.7	40.0	38.3	9	44.8	43.8	44.8
16	32	30.4	28.1	15.5	32	31.1	32.4
				24.0	23.3	22.6	22.5

d_1	A	B	Measured
(c) $f/2.5$			
μ	Lines/mm	Lines/mm	Lines/mm
47	12	12.2	19
40	14.7	14	21
30	19.1	17.5	26
20	27	23.3	31
15	33.6	28.0	34
13	36.7	30.4	34
15	33.6	28.0	30

It should be pointed out that these formulae do give a lens designer some idea of the relation between resolving power and image size. The examples are limited but it is the type of data a lens designer needs to know, for how else is he to design a lens to meet specifications that are invariably written in lines/mm. The formula gives a pessimistic answer that may be advantageous.

Picture evaluation. The prints of the scene were observed by two of the authors. Several combinations were obviously different in sharpness. The remaining combinations that were difficult to distinguish between were selected for inspection by several observers. The observers were asked to compare the pictures by pairs and to pick the picture that they believed contained the more information. Tables 13.5, 13.6, and 13.7 show the results of this survey. The first

<center>TABLE 13.5. *Image evaluation data at f/2.5*</center>

<center>Picture evaluation F/2.5</center>

Focal Setting	R. P.	d_1	V. B. T. Set 1	Comments
	Lines/mm	μ		
9.5	34	13	5	Very difficult to differentiate.
8.5	34	15	2	No significant difference.
9.5	34	13	6	}Significant difference.
10.5	30	15	1	
8.5	34	15	6	}Significant difference.
7.5	31	20	1	
8.5	34	15	6	Significant difference due to less than 30-
10.5	30	15	1	percent energy concentration.
7.5	31	20	3	No significant difference. Depends on
10.5	30	15	4	observer's method of viewing.
7.5	31	20	7	}Significant difference.
11.5	19	26	0	
6.5	26	30	7	
11.5	19	26	0	}Significant difference. Reason not clear.

<center>TABLE 13.6. *Image evaluation data at f/4*</center>

<center>Picture evaluation F/4</center>

Focal settings	R. P.	d_1	V. B. T. Set 1	V. B. T. Set 2	Comments
	Lines/mm	μ			
8.5	49	7	8	3	
9.5	45	9	1	5	
8.5	49	7	8	4	}All these pictures are similar. There is a possi-
7.5	45	7	1	3	bility that 8.5 is better than 7.5.
9.5	45	9	2	4	
7.5	45	7	7	3	
10.5	32	15.5	2	5	}Difference less than errors in printing.
6.5	35	14.0	7	2	
10.5	32	15.5	7	6	}10.5 significantly better than 5.5.
5.5	29	25.0	1	1	
11.5	23	24.0	6	3	}No preference, however, 5.5 definitely has more
5.5	29	25.0	3	4	detail.

<center>TABLE 13.7. *Image evaluation data at f/5.6*</center>

<center>Picture evaluation F/5.6</center>

Focal settings	R. P.	d_1	V. B. T. Set 1	Comments
	Lines/mm	μ		
8	56	4	6	}Significant difference.
7	55	5	0	
9	50	8	0	}Significant difference.
7	55	5	6	
10	38	14	3	No significant difference indicating influence
5	36	15	3	of low energy concentration.
6	47	9	6	}Significant difference.
10	38	14	0	
6	47	9	4	}No significant difference.
9	50	8	2	

column is the focal setting.   The second column gives the measured resolving power on the negative.   The third row contains the 30-percent energy spot size.   The fourth row contains the "Votes Better Than" for the final prints.   An arbitrary value of 75 percent of the votes was called a significant difference.

The data at f/4 is the most complete. Print set No. 1 was made with the closest tolerance in tone quality. A complete duplicate set of prints No. 2 were made to check the enlarging process.

Summary of Results

The survey may be summarized briefly as follows:

1. There was only one case of a significant difference where there was disagreement between lines/mm and the 30-percent spot-size criterion. This occurred at f/2.5 for focal settings 6.5 and 11.5. The curves for the two focal settings cross below 25 percent of the energy. The lines/mm criterion agreed with the picture valuation. This indicates that less than 25 percent of the energy is determining the resolution and picture quality. This value may be lowered in this case because the curves cross again at 65-percent energy content.

2. In one case (f/2.5) there was a significant difference between focal settings 8.5 and 10.5 in evaluation, no difference in the 30-percent spot size but a difference of 34 and 30 lines per mm. This also indicates that the energy concentration below 30 percent is influencing the results.

3. There are cases, for example, at f/4, focal settings 11.5 and 5.5 where the observers divided their opinion, and the resolving power and the 30-percent spot-size criterion were opposite. Careful observation of the prints reveal definitely more detail in some parts of the picture with the high resolution. If the observers viewed the pictures from a distance they picked picture 11.5. However, if they were asked to view the pictures closely with a low-power magnifier they would agree 5.5 had more detail. This observation agrees with what Hansen [5] pointed out. His observers picked the pictures taken at f/6.3 even though the pictures at f/4.5 resolved more. The observers either did not look for the fine detail, were not able to see it, or there was a small percentage of fine detail in the picture and most of the observers did not notice it.

4. The survey does show that very slight differences in resolving power are significant in the final picture. The pictures also show that as small as a 20-percent difference in spot-size results in a significant difference in performance. No attempt was made to compare the different f-number sets. Table 13.8 shows comparisons between resolving powers for different focal settings and f-numbers when the spot sizes were common. The agreement is fair but disagreement does occur when spot sizes are similar and focal settings are on opposite

TABLE 13.8. *Comparison of resolving powers for equal spot sizes obtained with different conditions*

Spot size	R. P.	Condition	
		f/No.	Focal setting
9	47	f/5.6	6
9	45	f/4	9.5
14	38	f/5.6	10
14	35	f/4	6.5
15	36	f/5.6	5.0
15	34	f/2.5	8.5
15	30	f/2.5	10.5
15.5	32	f/4	10.5

sides of focus.   The pictures taken at the best setting at f/5.6 were definitely better than those obtained at f/4, which in turn were better than at f/2.5.   The pictures also show very clearly the focus shift between f/2.5 and f/4.   There is a focal shift from 9.5 to 8, which is approximately 120μ.   This particular lens should be focused for best focus at f/2.5.   It is surprising how well the lens performs at f/2.5. The image of a point source viewed visually is very large.   The flare extends out to 200μ.

It is clear that the energy-concentration requirements of a lens depend on viewing conditions and interests of the observer.   One might generalize in the following manner.   If a lens is required to resolve fine detail then between 20 and 30 percent of the energy must be concentrated to within the appropriate spot size.   On the other hand, if the prints are to be viewed so that this fine detail will never be resolved by the observers then it would be better to allow a larger spot size and put more energy within the spot size.   Two opposite cases are illustrated between the interests of a photointerpreter and an amateur photographer.

The photointerpreter will view the prints with a magnifier looking for every bit of information available.   A lens for him then should provide as much resolving power as possible, and the designer should attempt to design a lens to put 30 percent of the energy to within as small a spot as feasible.

The amateur photographer may take his pictures on a 3¼ × 4¼ negative, make a contact print, and then view them from 10 inches. From this distance he will probably not notice any detail less than 14 lines/mm.   The lens, therefore, needs to have a spot size 40μ (Formula B) in diameter.   The designer has the alternative of placing more than 30 percent of the energy in this spot or simplifying the design and allowing just 30 percent.

It is agreed that one can be carried away by oversimplified formulae. The data obtained so far is scant.   More data on more different types of energy distributions should be studied.   The data is confined
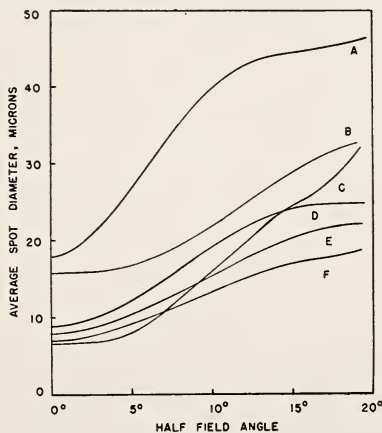

FIGURE 13.11.   *Plot of area weight-average spot size for several lenses.*

Weighted-mean spot diameter of 50-mm objectives (50-percent concentration).
A, Ektar f/1.9; B, Sonnar f/2; C, Ektar f/4; D, Elmar f/3.5; E, Sonnar f/4; F, Elmar f/6.3.

to Super XX and green light. Similar studies should be made with red and blue light separately and then added together.

It is our belief, however, that much of the concern about the accuracy of resolution data is splitting hairs. In actual practice a lens must cover an appreciable field of view. Over this field of view the spot size for the 30-percent energy concentration may vary by more than a factor of 10. It would appear, therefore, that a lens can be appraised most profitably by analyzing the energy concentration for several points in the field for several positions of focus, and to place the film in the proper position to obtain the minimum area–weighted-average spot size. Figure 13.11 contains plots of area–weighted-average spot sizes for several lenses. This type of analysis is extremely lengthy and a simple criterion like a 30-percent energy spot size is needed. We believe a simple formula, even though an oversimplification, is what optical designers need now. With it they can make improvements in overall field performance of existing lenses.

Conclusions

The above experiments show that the energy-distribution curve does indicate the performance of a lens. If the requirements on a lens call for maximum resolution, as little as 25 or 30 percent of the energy may be concentrated into as small a spot as possible.

It is recommended that lenses should be tested independent of film by measuring the energy distribution in the image of a point source. It is recommended that the lens should be checked photographically. Line test charts are satisfactory if backed up with energy-distribution curves. Anomalies may show up but they can be explained and advance the knowledge of the subject.

It is also urged that the lens designer's interests should be considered, by specifying lenses by their energy distributions. A single spot size may be sufficient; two spot sizes would probably be even better. Almost all military specifications are written today in terms of lines/mm resolution. There is so little data available connecting lines/mm and lens characteristics that the lens designer can go only on his past experience and theories.

The photoelectric method described in this paper is probably not suitable for production testing of lenses. The task of making all the required measurements is formidable. The method suggested by Schade [2] may offer a more convenient method of measurement, but we urge that specifications should be written in terms of energy concentration. Schade's response factors can be converted to energy concentration.

References

[1] W. Herriott, J. Opt. Soc. Am. **37,** 472 (1947).
[2] O. H. Schade, RCA Publication ST–3BD.
[3] J. D. Hunter, Proc. Roy. Soc. [A] **82,** 307 (1909).
[4] L. A. Jones and R. N. Wolfe, J. Opt. Soc. Am. **35,** 559–569 (1945).
[5] G. Hansen, Zeiss Nach. Heft **6,** 129–137 (Nov. 1942).
[6] S. Huber, Z. Instrumentenk. **63,** 333–341, 369–383 (1943).
[7] F. Twyman, Trans. Opt. Soc. **22,** 174 (1920–21).
[8] H. Coleman, J. Opt. Soc. Am. **37,** 671 (1947)

Discussion

Dr. K. V. Pestrecov, Bausch & Lomb Optical Co., Rochester, N. Y.: Dr. Hopkins, I really appreciate what you said at the very end about too much emphasis put on resolution and perhaps some other test could give you adequate results. But one who gets involved in these data comes to a conclusion because of oversimplification. I would like to challenge this statement. After all, if you have to choose between photographs and one at first looks a little better to you—perhaps because there is a little better contrast—although the resolution is a little lower, the picture with higher resolution still is better when you look for finer detail. I think the interpreter looks for recognition of detail. Resolution is my pet subject although I don't have much time to perform all these excellent experiments.

For example, if you consider a steeple on a church in several photographs, on one of the superficially worse looking photographs the steeple may look a little better. The picture with the highest resolution would have the peak of the steeple clearly defined. The photointerpreter looks for recognition of the steeple. Perhaps in high-resolution photography some corners will be better defined although the overall detail may be hazy; nevertheless, the picture may look a little better and one would have an easier job of recognition. I am not defending my position, I am just questioning. I think there is some work pointing in that direction. It may be recognition of detail but not mathematical detail that is so important. The photograph may actually be better for functional use, aerial interpretation, or something like that. I should like your comments.

Dr. Hopkins: I think there is a great deal in whether you are trying to recognize something or whether you are trying to detect what is there. I think it is splitting hairs. I think basically that if you want to get a lens to satisfy all photointerpreters, you have to put at least 25 percent of the light within as small a spot as is practicable. That, in itself, is theoretical. When you start to design the lens, the minute you get off axis you really have a problem. I am not so sure that anything more refined is necessary at this time. The experience we have had is that if there is high resolution the people can pick up more detail. We would ask them to look at the ivy in this picture and tell whether it had little leaves or little buds on it. They would be able to tell in the high-resolution picture but not in the others.

From the results we had, resolution is very important. Also, I have heard a lot of comments about people saying they can show me a picture that does this or does that, but I haven't seen them. We have a set of pictures you can look at.

Mr. R. V. Shack, National Bureau of Standards, Washington, D. C.: When you get off axis do you recenter your spot?

Dr. Hopkins: Yes, when you have coma it means that you will be pushing the aperture around each time you change it.

Dr. M. Herzberger, Eastman Kodak Co., Rochester, N. Y.: Prof. Hopkins, I think your method and the method that we apply can be very easily coordinated since we have in our calculations a quantity that is equivalent to the light distribution. One can very easily count the number of points within an area, and this determines the area that contains 20 or 25 percent of the light and one

can easily see by inspection where to look for such an area. I think, if you have a lens, it would be quite interesting to try to see whether we could duplicate your results from the data of the calculation.

Dr. Hopkins: It would be very interesting. We actually have the design data on that F/2.5. We have retraced it and checked the energy-distribution curves. There is quite a difference. We don't know whether the lens was not made right or the patent fails to include the latent constructional details.

That actually is a very difficult problem to work out. I would not want to attempt it on anything more than a simple doublet.

Dr. Pestrecov: I would like to state that we have recently found with some of our work that there is a negative correlation between limiting resolving power and what we have been choosing to call picture sharpness. This business of using a resolving-power limit can get you into very serious trouble. It does not tell you about picture sharpness. In a great many cases they do run parallel, but in many significant cases they do not.

Dr. Hopkins: I covered myself by saying that if we had high-resolution data and energy data we would be able to explain this.

Mr. A. H. Katz, Wright Air Development Center, Dayton, Ohio: In connection with the points raised by Dr. Pestrecov and in earlier papers, I notice that a number of people have been gleefully trying to kick the three-line resolution target to death. I want to point out again—and I have done this in other meetings,[1, 2] that it has served its purpose well. This purpose, simply stated, is to serially grade lenses in a manner that will correlate with their photograph-making rank. I have yet to be shown that our use of the three-line target in the judging of lenses to be used for aerial photography has led to any error, let alone consistent error.

Now, we have lots of data, most of which is not neat and packaged. The exigencies arising with the working conditions in the Air Force are such as to effectively preclude the careful running of planned experiments. We substitute large numbers of airplane flights and tests, and after a number of years we come to pretty definite conclusions—by statistical osmosis, if you will. We know by now that when we get a lens that performs well in the laboratory (on the much maligned three-line high-contrast target) it will take high-quality photographs in the air on good days as well as bad days. The converse is also true. Laboratory tests enable us to predict the quality of actual aerial photographs. I can't expect much more of a laboratory test. Let us not forget that it is only within the last 10 years that lens performance began to be specified in terms of resolution requirement over the field, and that manufacturers began to use these tests, and it is only within the last couple of years that photointerpreters have begun to hear of lines per millimeter as a measure of performance.

A word about the photointerpreter, whose name has already been taken in vain. He is the consumer of the photographs produced by the lenses, and if he were here today he'd be speechless. Only lately has he really realized that the scale of a photograph (the altitude divided by the focal-length) is not enough to describe what he needs. The fact that some photographs are sharper than others has only com-

[1] A. H. Katz, Contributions to the theory and mechanics of photointerpretation from vertical and oblique photographs, Photogrammetric Eng. **16**, 339–386 (June 1950).
[2] Panel Discussion, Cameras, lenses, and calibration, Photogrammetric Eng. **17**, 417–420 (June 1951).

plicated his life, and it is not because of him—as would be logical— that we have specified, procured, tested, and standardized lenses, many of which can hardly be significantly proved. What I mean is that we've done this for the photointerpreter, but not because of any enlightened self-interest on his part.

A few years ago the Air Force issued a statement that we could take photographs of railroad ties from 40,000 feet. It so happens that we had previously resolved railroad ties on photographs from 30,000 feet, but we couldn't find these when asked to produce actual photographs. Well, we had to do it. We did it. It was not extraordinarily difficult, but it was not done the first time we tried. One of the requirements on the photography was that it be of the St. Louis area. We used one of Dr. Baker's 40-inch f/5.0 telephoto lenses, mounted on a standard K–22 camera, and flown in an RF–80 reconnaissance aircraft. No moving film magazine, no special mounts were used. Just straight photographic technique—careful exposure, development, and printing. We found the particular stretch of track later. The contrast—on the ground—between ties and ballast was negligible—the main differences being that of texture between well oiled ties and dirty darkened ballast. Now in this series of photographs was one of the Alcoa aluminum plants in East St. Louis; I identified this as such because on the roof of one of the buildings was painted "ALCOA." However, when I started to show these exceptionally sharp photographs to a photointerpreter here in Washington he identified the aluminum plant before I got the photograph within 2 feet of his eyes. His identification was based on the large residue lake near the plant, a characteristic feature of aluminum production plants, I haven't the vaguest idea how to put this into an equation.

I think many of you here have either cooperated on or seen this standard—Military Standard–150.[3] It represents the best government-industry-designer agreement we have ever had on the subject of definitions and test methods. In this standard we adopted the area–weighted-average resolution of radial and tangential resolution

$$AWAR=\left[\frac{\sum R_i T_i A_i}{A}\right]^{1/2},$$

where R_i and T_i are the radial and tangential resolution in the zone of area A_i, and A is the total area of the field used. Some of us like to ascribe some meaning to this average, although what happened is that we needed a number and we produced the most reasonable and justifiable we could think of.

But how about the difference between two lenses that have the same AWAR? Suppose we have a lens whose resolution varies from 40 lines/mm at the center to 10 lines/mm at the edge of its field, and that the average is 20 lines/mm. Let us say we have another lens that varies from 22 to 18 lines/mm, center to edge, and that also averages 20 lines/mm. Which lens is preferable, assuming uniform probability of occurrence of objects of interest in all parts of the field? I strongly suspect that we really want the second lens.

Other questions dealing with evaluation of two lens systems are even more complicated. Suppose for Air Force purposes, we have,

[3] Photographic lenses, Mil-Standard–150 (U. S. Government Printing Office, Oct. 23, 1950) 25¢.

let's say (with our test, which can be disputed the rest of the evening of course) a lens that will average 22 lines/mm with the film used in aerial photography. Let's also suppose it weighs 3 times as much and costs 10 times as much as another lens of the same focal length and angular coverage, but that the second lens averages only 14 lines/mm. (By the way, we never seem to be faced with the converse of this situation. Here's a good problem for the lens designers.) Let's further suppose that the detail we're interested in is caught with the lighter, cheaper, lower resolution lens perhaps but 10 percent of the time it is used, but the better lens records this detail in interpretable form 90 percent of the time it is used. How then compare the value of the lenses?

From this standpoint it is clear that relative figures of 9 to 1, or some function of these figures, is more reasonable than actual cost or resolution figures. I am not throwing these out as answers but am only suggesting that there are large numbers of questions that are unanswered, even after resolution figures or other performance indices are given. Unfortunately, the more we learn of this business, the more questions arise. Progress here seems to lie in the direction of awareness of relevant questions.

I also feel strongly that we are really looking for some measure of the information content of an optical (or photographic) image. This question has not been thoroughly explored, and as far as measuring information content, I feel that all proposals—U. S. Air Force resolution measurements, Howlett's doughnuts, Cobb Charts, Eastman's acutance, etc.—while clearly having some relation to information content, are equally poor measures of information content. This is no criticism of these interesting systems. They weren't supposed to measure information content, and they don't. I think this may well be a fruitful area for careful investigation.

This particular problem is already under investigation by Dr. Macdonald's group at Boston University, where the concepts of communication theory and related techniques are being applied. The main difference between this approach and the ordinary methods of measuring lens performance seems to be in consideration of the overall system—including the ground object, the atmosphere, the camera, its platform, the lens, and its windows, the film, processing system, and the interpreter. But since this is such a large subject it provides me a good place to stop.

Chairman: Dr. Baker?

Dr. J. G. Baker, Harvard College Observatory and Perkin-Elmer Corp., Norwalk, Conn.: There is one point I meant to emphasize this morning in connection with aerial photography. It is that an increase in focal length appears to be much more effective in improving ground resolution than an increase in the ultimate quality of a lens. In other words, put it this way. Suppose that you are taking aerial photographs under hazy conditions, which is often the case. If you double the focal length, you may realize up to twice the results in ground resolution, whereas if you double the lens quality of the standard lenses, you may realize only a 20-percent gain. On good photographic days, a doubling of the lens quality may realize up to a 60 percent-gain, possibly, which is likely to be a striking difference. The focal length is still the more important parameter.

This does not mean that quality should not be sought after. Where the focal length must be held to a certain value because of space limitations, the high-quality lens will easily out-perform the low-quality lens. Differences of a few lines/mm are discernible and often of importance. But where the focal length can be increased, that is the more certain way to increase ground resolution. Where quality and long focal length are combined, you have the best answer, and this has been the goal of my own work since 1941.

However, at some point in the quality spectrum the law of diminishing returns sets in, and a lens may become too expensive for the percent gain achieved. The designer should go as far as he can in obtaining quality, while employing practicable constructions. But sometimes the demands of a problem are so severe as to require complicated constructions in order to maintain even a fair degree of quality. When I speak of lens quality, I am thinking of microscopic contrast as well as of resolution.

Dr. R. C. Gunter, Clark University, Worcester, Mass.: These remarks are addressed to Dr. Ingelstam and to practically everyone who has spoken before and some who are going to speak subsequently. Some years ago during a rather heated discussion between some medical men, when it didn't seem that we were going to be able to resolve the question as to whether a new theory was or was not applicable to the human brain and various ones were tearing each other apart mentally, I suggested that perhaps the reason for the difficulty and the lack of agreement between the experimentalists and the theorists was similar to that that had existed in objection between the wave theory and the particle theory. The minute I said that Werner turned around like a flash—I was the only physicist I think in the group and that was the only observation I could make—and he said, "The trouble with you is that you neglect the role of the observer."

Now, it has been touched upon several times as I have said, by Dr. Howlett, by Dr. Hopkins, and by many others, that we should be able to separate the observer, the human being, from the equipment. If we cannot at this time define what we mean by the word "energy distribution", perhaps we had better reevaluate some of the fundamental concepts of physics. On the other hand, it is perfectly possible and is indeed present in every day life—as Dr. Hopkins could undoubtedly tell you though I haven't conversed with him on this—if you ask somebody, "Do you see such-and-such in a picture," if there is any possibility of seeing it there invariably people will say, "Yes." There are many other experiments in which the borderline, where we go from psychology, perhaps, into physics, is so nebulous as to influence our decisions. This is a point, as Dr. Howlett mentioned, actually of rather deep philosophical importance, but I believe that the point that was made in Dr. Ingelstam's abstract—it is in quotations where he says so-and-so. I would like to see some discussion at a later date—it is getting late tonight—on some support of experiments that could be conducted to find out just how much of a part the observer plays from the point of view of seeing what he wants to see. I have measured spectral lines and after awhile I knew I was measuring lines that were not there but I was recording wavelengths for them.

I would like to hear Dr. Ingelstam's further delineation of his ideas of putting down this term "standardization."

Dr. E. Ingelstam, Royal Institute of Technology, Stockholm 70, Sweden: I quite agree with you. One should distinguish what is really physically measurable concerning the intensity distribution. How it should be done I can not say at present. I think there could be some discussion tomorrow after we hear Selwyn's paper and some other papers. However, I think that one should make a separation between the physical resolving power and the intensity distribution in a given instrument, and the influence of the receiver, maybe the eye or the photographic plate. I think that there is sufficient work done now to enable one to make decisions of this type. I quite agree with the logic of your question.

14. Optical Calculations at the National Bureau of Standards

By Donald P. Feder [1]

Since the advent of automatic computing machinery in the past few years, these machines have been increasing in numbers very rapidly. There are at present some six, large-scale, general-purpose, automatic, digital computing machines in operation. In the next few years their number will have increased to about 50. In addition there are a larger number of medium-sized machines in use. In view of the increasing availability of this equipment, which can greatly facilitate optical calculations, you may like to hear something about these machines and the optical calculations which have been programmed for them, at the National Bureau of Standards.

Three words have been used to describe these machines, "automatic", "high speed", and "electronic". By "automatic" one means that the initial data and the constants of the system are introduced into the machine, which solves the equations, stores necessary intermediate results, makes logical decisions about the course of the calculations, and finally prints out the answers in a useable form.

The phrase "high speed" can be best illustrated by an example. The SEAC (Standards Eastern Automatic Computer) has been coded to trace a general ray through a system of spherical surfaces. Through a system of 10 such surfaces, the SEAC can trace a skew ray in 8 seconds. In this time it does 440 additions, 320 multiplications, 10 divisions, and 20 square roots, and it does these things to 44 binary places (about 13 decimal places).

Finally "electronic" simply refers to the technology that has made these machines possible.

In addition to the SEAC there is also available at the Bureau a CPC (Card Programmed Electronic Calculator). This is a machine of more moderate speed and capacity. It has a memory for 23 numbers in one type of storage, and, in addition, a small high speed memory. The use of the high speed memory permits the very rapid extraction of square roots.

By comparison the SEAC has a memory of 1,024 words. In the SEAC, however, some of the storage space always contains instructions. This limits the effective memory considerably. In the CPC the instructions are on punched cards and do not occupy any of the memory cells.

The CPC has been used for routine ray tracing since March 1950. This machine required 34 seconds to trace a skew ray through a spherical surface. It traces one ray at a time, which distinguishes it from certain other machines of this type. It has also been coded for nonspherical surfaces so that now rays can be traced through any system of rotationally symmetrical surfaces. The aspheric surfaces are

[1] National Bureau of Standards, Washington, D. C.

handled by adding a power series to the equation of a spherical surface. The power series may contain as many terms as necessary and this introduces no storage problem because the coefficients of the power series are stored on punched cards. The problem is to find the intersection of the ray with the aspheric surface. A simple iteration has been found, which converges and, what is more important, converges in less than five iterations for any aspheric surface likely to be used optically. The entire calculation requires about 2 minutes per surface, but since not many surfaces of a system are apt to be aspheric, this is not very serious.

Something should be said about the type of formulas used in these calculations. They involve only curvatures and never radii of curvature. Each number used by the computer is bounded in a known way by the linear dimensions of the optical system. Hence, by choosing a proper scale factor it is always possible to fit the problem into the machine. Floating decimal operations are thus avoided. All the numbers involved are of the same order of magnitude and the accuracy is very good.

These same statements might also be made in connection with the other optical calculations discussed in this paper. The formulas have always been converted to such forms that no number has the possibility of becoming infinite. In practice this means that no number becomes very large and so accuracy is maintained.

The other type of calculations presently being done at the Bureau are calculations of the image errors of various orders. The CPC was first programmed for the calculation of the first and third orders (that is the Gaussian and Seidel coefficients). For this purpose two paraxial rays are traced through the system, a paraxial marginal and a paraxial principal ray. The machine runs through the calculation and prints the following data: The height of each ray on each surface; the slope angle of each ray in each medium; the coefficient of spherical aberration, of coma, of astigmatism, of field curvature, and of distortion. In addition it prints the first-order coefficient for the two types of chromatic aberration and five checks, which ought to be zero. These checks may be inspected to verify that no mistake has been made. The entire calculation including the printing requires only 40 seconds per surface.

A somewhat more unusual calculation is that to obtain the fifth-order coefficients. The equations are algebraic equations modified from those published by Wachendorf.[2]

It is too early to say how valuable the fifth-order coefficients will be since our experience with them is extremely limited. It is likely that their chief value will be in optical design rather than in the evaluation of image quality.

With the CPC the complete calculation of the fifth-order coefficients requires 7 minutes per surface. While this may seem a long time, the calculation is very laborious and would require almost too much time to be practical for hand computing. This calculation is being programmed for SEAC and on this machine will probably require only a few seconds per surface.

The calculations mentioned above are all being done routinely at present in this laboratory. It can be seen that with the possible

---

[2] F. Wachendorf, Optik 5, 80 (1949).

exception of the fifth-order coefficients and the tracing of skew rays through aspheric surfaces, these calculations are no more extensive than are commonly performed by lens designers using hand machines. We have seen that there are gaps between the information that the designer gains in this manner and the information that the user of the system would like to have.

In the past the designer has usually obtained curves denoting the aberrations of meridian rays as functions of aperture, and curves showing the position of the sagittal and tangential foci as functions of field angle. These have not been sufficient to answer many questions about the image formed by the system. Such a question, for example, as the calculation of the limiting resolution in lines per millimeter from a particular combination of object, lens, and film could not be answered from such data alone. More generally, one might be interested in the image contrast from an object of given contrast produced by a lens and film of known characteristics. Questions such as these ought to be answered by calculation from the constructional data of the system.

In accomplishing this purpose various approaches might be tried. With a machine such as SEAC, it should be possible to calculate the diffraction disk for a luminous point. On the other hand it may not be necessary to go beyond the geometrical picture, especially for photographic objectives where the aberrations are likely to be many times the Rayleigh limit. In this case it is possible to get an approximate distribution of energy in the image using a method similar to that used by Herzberger.

In this method one divides the entrance pupil into a large number of equal areas. For a particular object point a ray is traced through each of the small elements of area, and its intersection with the image plane is calculated. One supposes that the energy of each element of the wave front arrives at the image plane in the neighborhood of the ray. If the calculation is made for a large number of rays, the distribution of their image points gives an approximate picture of the energy distribution in the image of a luminous point. When the distribution of energy is known, a knowledge of the characteristics of the photoreceptor should enable one to find the appearance of the image.

Another possibility is to calculate the interferometer pattern for various field angles. A large number of rays might be traced by a machine such as SEAC which would retain in its memory the results of the calculations, fit them to some favorable type of function giving the wave surface and then by intersecting this surface with a series of spheres one-fourth of a wavelength apart obtain an interferogram. This could then be interpreted directly to secure an idea of the image quality.

15. Resolving Power of Airplane-Camera Lenses

By F. E. Washer [1]

Shortly after the close of World War I, the National Bureau of Standards began testing photographic objectives for the U. S. Army Air Corps to determine their suitability for use in airplane cameras for mapping projects. The measurements made included determinations of spherical aberration, chromatic aberration, curvature of field, astigmatic difference and distortion.

These measurements were made on a visual optical bench [1][2] and for a time were entirely adequate. However, as the use of airplane photography for both planimetric and topographic mapping increased, it became evident that the volume of work submitted to this Bureau was too great to handle by the visual method. In addition, the evaluation of the quality of definition throughout the image plane together with the location of the plane of best average definition is a laborious and time-consuming process when performed visually. Moreover, legitimate doubt exists as to whether a visual method can be satisfactorily used to determine probable photographic quality of definition.

To meet the increasing demand for a rapid and adequate performance test for photographic objectives, the precision lens-testing camera, shown in figure 15.1, was conceived by I. C. Gardner and designed by F. A. Case [2]. This device in its earliest form permitted the evaluation of the equivalent focal length, distortion, and resolving power from a single negative. Nineteen rows of images were obtained on the negative, each of which showed the quality of imagery at 5° intervals from the center of the field to a point in the image plane 30° from the axis for a specific distance from the lens. Consequently, it became an easy matter to locate the plane of best average definition and to determine the back focal length of the lens associated with the plane of best average definition.

A resolving-power test chart, shown in figure 15.2, was developed for use in this equipment, consisting of a series of patterns of parallel lines with the spacings varying by the square root of two. The contrast between the lines and spaces in the target patterns is high and may be considered as infinite.

With the large number of lenses being tested, it was only natural that a considerable quantity of test negatives would accumulate that could be used to carry out statistical investigations of the quality of imagery inside and outside of focus as well as in the plane of best axial imagery. The first of these studies was reported in 1939 [3] and dealt with three types of airplane camera lenses in most common use at that time. The wide-angle lens had not yet come into extensive use at that time and the investigation was confined to imagery in the angular region between 0° and 30° from the axis for lenses having a focal length

[1] National Bureau of Standards, Washington, D. C.
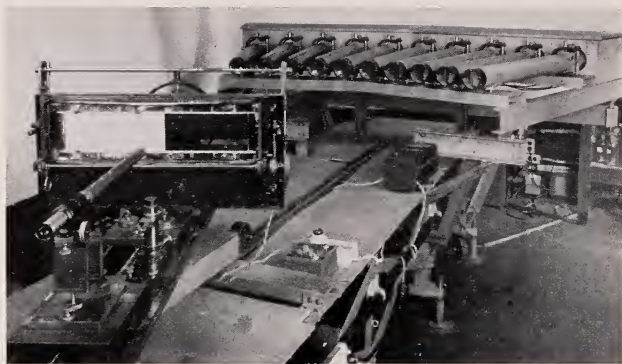[2] Figures in brackets indicate the literature references on p. 218.

FIGURE 15.1. *Precision lens-testing camera.*

The 10 collimators are in the upper right-hand part of the photograph; the lens testing bench and plate holder are in the lower left-hand part; the lens holder is obscured by the camera back and doew not appear in the picture.
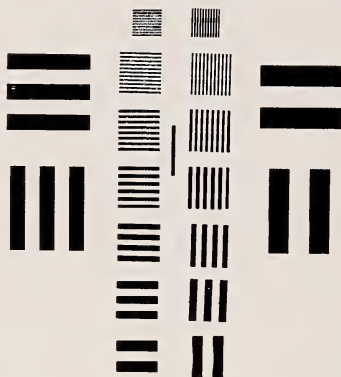

FIGURE 15.2. *Chart for testing resolving power.*

A chart of this type is located in the focal plane of each collimator.

of $8\frac{1}{4}$ inches. Several items of interest were found in this early investigation. First, the resolving-power characteristics of a given type of lens remain remarkably constant from one lens to another. Second, what appeared to be striking variations from one lens to another so long as a single image plane was considered faded into insignificance with proper choice of the focal plane. Third, by proper choice of a focal plane, the performance of a lens in the outer portions of the field could be considerably improved with only a small diminution of quality in the axial regions. This finding led directly to the concept of plane of best average definition, which is now recognized as the best plane to bring into coincidence with the focal plane of the aerial camera. Fourth, the phenomenon of false resolving power was noted. This led to the concept that a pattern should only be counted as resolved when it and all coarser patterns in the series were resolved.

Each of the collimators of the lens-testing camera is provided with one of these charts with the line patterns so proportioned that when a given pattern is imaged by the lens under test the same value of resolving power is indicated regardless of its position in the field. The values of the resolving power in the image plane of the test lens usually ranges from 3.5 to 56 lines per mm for both tangential and radial lines.

Eastman type V-P spectroscopic plates are generally used. This emulsion, while too slow for airplane photography, has much greater resolving power than the panchromatic films used in aerial photography. Consequently, this makes it probable that an indicated deficiency in the observed resolution arises from a deficiency in the lens rather than in the photographic emulsion. The plates are processed in Eastman D19 developer for 3 minutes at 65° F.

When the lens-testing camera was first put in operation in early 1936, perhaps 50 lenses were tested the first year. The number has increased steadily since, and during one recent year more than 300 lenses were tested. It is safe to say that the performance of well over 1,500 lenses have been evaluated on the lens-testing camera alone since 1936.

By 1940, it was evident that the trend was toward the use of wide-angle aerial camera lenses and the precision lens-testing camera was modified to extend its range to cover the region from 0 to 45 degrees from the axis. The second study on resolving power was reported in 1942 [4] and dealt among other things with the resolving-power characteristics of several types of wide-angle lenses.

Throughout the years of making test negatives on the precision lens-testing camera, it was always evident that the device was admirably well suited to investigating image quality throughout the region of usable imagery and in 1945 [5] a study was reported along these lines. The curves, shown in figure 15.3, illustrate the manner in which the resolving power increases and decreases as the image plane moves steadily from a point well inside focus to a point well outside focus. The results are plotted at each 5° interval from 0 to 45 degrees from the axis for both tangential and radial resolving power. The zero line of abscissae corresponds to the point of best axial focus. The effect of curvature of field is strikingly manifested by the change in relative position of the maximum with respect to the plane of best axial imagery. The varying widths of the principal maximum increase with increasing f-number and in the graphs shown for f/16, the heights of the peaks are beginning to show the expected decrease. From curves of this type, complete information concerning the effect of stop opening on resolving power and depth of focus; the change in maximum resolving power with angular separation from the axis; and the effect of change in stop opening on the resolving power in a given image plane may readily be obtained for a given lens.

The possibilities of the precision lens-testing camera as a research instrument in the study of resolving power are by no means exhausted. However, since this device was first put in operation the study of resolving power of aerial camera lens has become the subject of interest in other laboratories both here and abroad. These other investigations brought fresh points of view and new ideas regarding the manner in which the subject should be approached. A wide variety of test charts was developed by the various groups. Chief among these
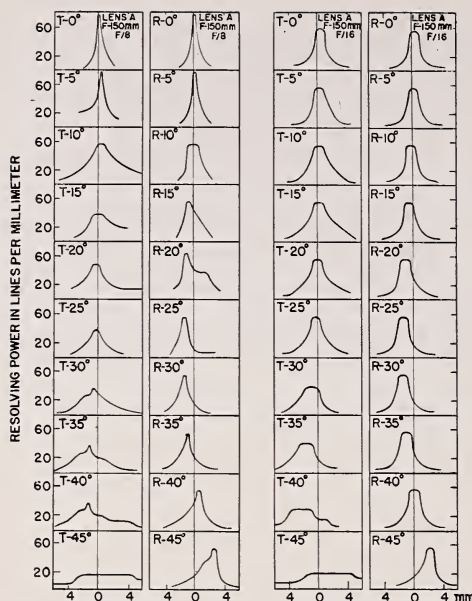
FIGURE 15.3. *Resolving power versus position of the image plane for lens A a apertures f/8 and f/16.*

The resolving power throughout the region of useful imagery is shown for tangential T and radial R lines at 5° intervals from 0° to 45°. The zero of abscissae marks the position of best visual focus on the axis at f/8 and positive values of abscissae indicate positions farther from the lens.

charts are the three-line high-contrast target of the Air Force [6], the low-contrast annulus target of the Canadian group [7] and the low-contrast two-line or Cobb chart of the British group [8].

The principal changes in approach from that followed at this Bureau are the emphasis on low contrast and the use of emulsions having characteristics more nearly approximating the emulsions used in the aerial camera. The low contrast of the test chart was justified on the basis that the difference in brightness between adjacent objects on the ground was small when viewed from an airplane and, consequently, the variation in contrast in an aerial negative is low.

As a result of the findings of these other laboratories, there has been some tendency to question the suitability of the high-contrast chart in the determination of resolving power. It is right and proper for such questions to be raised, but in defense, it ought to be stated that the results of the resolving-power test as used at the National Bureau of Standards have been applied as a go or not-go criterion by the various mapping agencies. A lens to be accepted for use in aerial photography must satisfy certain minimum requirements. These minimum requirements were early established by examining a number of negatives made by different cameras and assessing the negatives as being suitable or unsuitable for use in map making. On checking back to the resolving-power measurements made on these lenses for the high-contrast target, it was found that the unsuitable negatives were obtained with lenses having a resolving power less than 7 lines/ mm, whereas for the suitable negatives the resolving power of the lenses exceeded 7 lines/mm in all parts of the useful field. It is

unfortunate that this correlation was not made the subject of a formal report instead of being informally presented as was done at the time.

This work was done by a committee of the American Society of Photogrammetry and the suggested lower limit of 7 lines/mm first appeared in 1935 in a tentative set of standard specifications [9] prepared by this society. In later specifications the figure was changed to 20 lines/mm for central imagery with the figure left at 7 lines/mm throughout the remainder of the useful field. In addition, there are Federal Specifications pertaining to optical performance that require that the quality of imagery be determined with the aid of high-contrast charts. The Federal Specifications are used extensively in connection with government purchases. The requirements contained therein have been developed over a period of years in collaboration with representatives of industry and government agencies. Consequently, any change that is recommended must be in the direction of a definite improvement that is clear to all concerned. A change in requirements frequently necessitates the development of new test methods, the development of which must precede any actual change in the Federal Specifications themselves. Accordingly, changes in these specifications are not quickly made and even though new requirements may appear desirable, the existing requirements continue in use until the new requirements with appropriate test methods are agreed upon.

Because of the uncertainty that had been created in the minds of those people concerned with specifying quality of definition as a result of the many diverse opinions on the subject, it was decided to initiate further research at this Bureau into the factors affecting definition in the photographic image. A new resolving-power test chart was designed that would simultaneously give information on resolving power and the effect of contrast on resolving power. At the same time, the chart was provided with a greater range of frequencies and capable of yielding greater precision in determining the limit of resolution. The design of this chart was reported by I. C. Gardner in March 1950 [10] at the New York meeting of the Optical Society of America.

The actual making of the new chart was reported at the Cleveland meeting of the Optical Society of America in October 1950 [11]. This new resolving-power chart is so made that the transmittance of the lines and spaces along the y-axis vary in such a manner that the contrast is a linear function of y and the transmittance of the chart averaged over an area embracing several pairs of lines is uniform for the entire plate. Along the x-axis, the widths of the lines and spaces progressively decrease so that the "instantaneous" value of the line frequency is a linear function of the distance measured normally to a "zero" line of abscissal of the chart.

The chart was made in two steps, the first step being the making of the master high-contrast chart, shown in figure 15.4. The range of frequencies on the master target plate is from 0.2 to 2 line/mm, which means that the actual widths of lines and spaces vary from 2.5 mm to 0.25 mm. To cover this range in a continuously varying linear manner there are 200 lines and 200 spaces contained in a distance of 180 mm. The zero line is 20 mm from the origin so that the over-all length of the resolving-power portion of the chart is 200

FIGURE 15.4. *Master high-contrast test chart.*

The line frequency is a linear function of the distance measured across the lines. The zero of the chart is at the inner edge of the first broad line.

mm. It is necessary that the lines be as long as possible to provide a contrast scale of approximately the same length as the frequency scale and, moreover, to simplify the process of making microdensitometric studies of final test images. The length of line used is approximately 185 mm, enabling the entire chart to be registered on a standard 8- by 10-inch plate.

The second step consists in contact printing the high-contrast target on a photographic plate under conditions such that the exposure time varies over the plate in a predetermined manner [11]. The high-contrast target is then removed and a second exposure is given to the photographic plate in such a manner that the transmittance averaged over several lines of the finished negative is a constant. Figure 15.5 shows a typical variable-contrast target produced by this process. For the particular variable-contrast target illustrated in figure 15.5, the scale at the left marked "Contrast" shows that in the original transparency the contrast varies in a linear manner from 0 to 1.5. Contrast as used here is simply the difference in photographic density between dark and light areas as measured on a densitometer. The lower scale gives the line frequency or resolving power in lines per millimeter at a 25x reduction. Since the present print is somewhat reduced in size, the marked scale is correct only when the distance from 0 to 50 is equal to 8.0 mm in the final image. The chart was photographed for use in this illustration and it is unlikely that the contrasts shown on the scale are still correct. However, in the original, the contrast remains constant at a given y-height, whereas the line frequency varies linearly from 0.2 to 2.0 lines/mm.
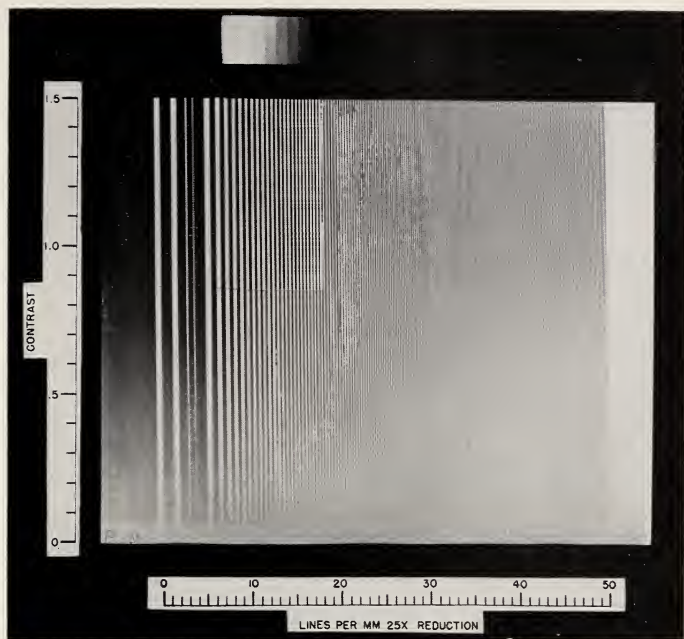
FIGURE 15.5. *Variable-contrast master test chart.*

In addition to the features contained in the master high-contrast chart, the contrast varies in a linear manner from 0 to 1.5. The "moiré" effect appearing in the right-hand portion of the illustration is pro-duced by the interaction of the periodic pattern of the halftone screen and the pattern of continuously vary-ing line spacing in the chart. Consequently, the above figure is not a true reproduction of the chart, al-though it is a fairly close approximation in the left-hand portion.

Some trials have been made with this new target and the results are shown in figure 15.6. The target was set up at a distance of 51 focal lengths in front of the test lens. It was lighted from the rear by a broad uniform light source and negatives were made using three different emulsions with the lens set at an aperture of f/8. Figure 15.6 shows the results obtained with Eastman Spectroscopic Plate, type V–G. The reduction ratio is 50 to 1 so that the frequency that is shown at the bottom of the figure ranges from 10 to 100 lines per millimeter. Values of the contrast obtained from microphotometric recordings at three different heights across the negative image are shown as a function of the line frequency. The contrast in the target at a given height is approximately constant, so that the decrease in contrast in the image results from action by the lens-emulsion combi-nation. Curve 1 shows the results for high-contrast target, curve 2 for a medium-contrast target, and curve 3 for a low-contrast target. The results appear to be in agreement with what one might anticipate on the basis of the literature dealing with the subject of target con-trast and resolving power. However, the chief point to emphasize here is that the data for drawing these curves were obtained from a single negative, with consequent marked reduction in the number of factors that might conceivably make difficult the correlation of resolv-ing power with contrast in the target.
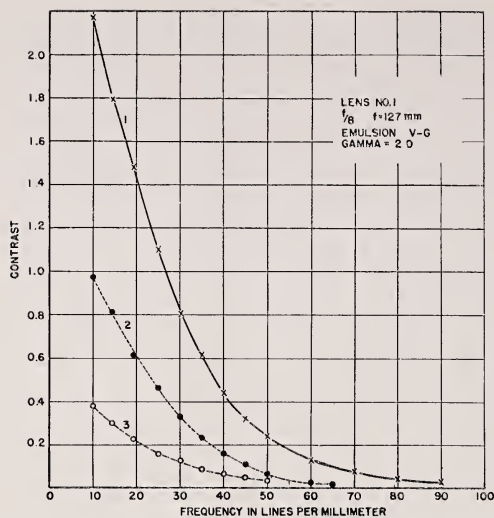
FIGURE 15.6.   *The variation in contrast for increasing line frequency for Eastman spectroscopic type V–G emulsion.*

Target contrast is 1.3 for curve 1; 0.5 for curve 2; and 0.2 for curve 3.   (The average density of the negative is 0.45).

At the present stage of the investigation, it is clear that it would be difficult to determine just what value to assign to the upper limit of resolving power of the lens for a given target contrast.   The resolving power reaches its maximum when the contrast in the image reaches zero.   However, because of local irregularities in the image it is not possible in the present stage of the investigation to state with certainty just when the contrast in the image reaches zero.   As a practical expedient it may be preferable in reporting the performance of a lens under given conditions to state what the value of the resolving power is for a selected value of contrast in the image for a specified value of contrast in the target.

For example, we may wish to report what the resolving power is for an image contrast of 0.1 for several values of target contrast.   This information is readily obtained from these curves.   Table 15.1 shows the values of the resolving power for an image contrast of 0.1 for three emulsions for the three values of target contrast that were shown in figure 15.6.

TABLE 15.1   *Resolving power in lines per millimeter for image contrast of 0.1 for* **three** *emulsions for three target contrasts*

Target	Contrast	V-G	Pan Process	Panatomic X
1	1.3	64	36	30
2	0.6	46	25	20
3	0.2	33	13	13

The development and use of the multiline variable-contrast target have been of primary interest in our work on resolving power during the past year.   Consideration has been given to 3-line, 2-line, and

216
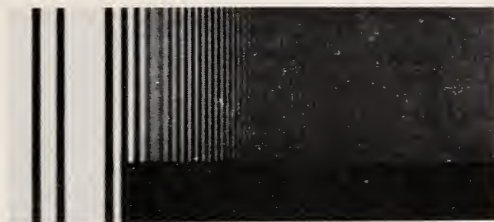
FIGURE 15.7.   *Typical out-of-focus exposure showing spurious resolution.*
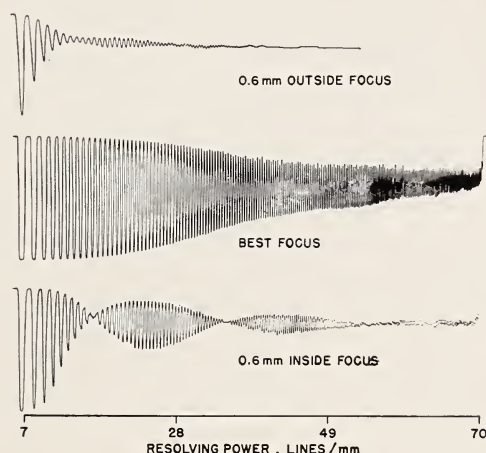
FIGURE 15.8.   *Microphotometer traces of images formed by a lens at f/5.6 on Eastman spectroscopic emulsion type 548 GH.*

1-line targets.   Some of the results obtained with these targets, together with their theoretical interpretation, are reported in paper 20.

Work with high-contrast targets has not, however, been completely neglected.   The high-contrast target with continuously varying frequency has been found to be exceptionally well suited to the study of some phases of spurious resolution.   A report on this phase of the investigation was made by R. N. Hotchkiss [12] at the Cleveland meeting of the Optical Society of America.

The axial imagery was studied throughout a region extending from a point well inside best focus to a point well outside best focus.   A typical out-of-focus image is shown in figure 15.7.   It is readily seen that the contrast between adjacent lines and spaces quickly falls to approximately zero as one proceeds toward the finer patterns.   However, instead of remaining zero, the contrast rises and falls repeatedly as one continues in the direction of the finer patterns.   This waxing and waning effect is even more strikingly illustrated in the microdensitometric traces shown in figure 15.8.   For the cases shown, the phenomena can be explained fairly simply on a geometric basis.   This analysis is given in the September 1951 issue of the Journal of the Optical Society [13].

[1] A. H. Bennett, J. Opt. Soc. Am. **14,** 235 (1927).
[2] I. C. Gardner and F. A. Case, J. Research NBS **18,** 449 (1937) RP984.
[3] F. E. Washer, J. Research NBS **22,** 729 (1939) RP1216.
[4] F. E. Washer, J. Research NBS **29,** 233 (1942) RP1498.
[5] F. E. Washer, J. Research NBS **34,** 175 (1945) RP1636.
[6] C. W. Kendall and B. A. Schumacher, Photo Tech. **3,** No. 4, 51 (1941).
[7] L. E. Howlett, Can. J. Research **24,** Sec. A. No. 4, 15 (1946).
[8] E. W. H. Selwyn and J. L. Tearle, Proc. Phys. Soc. **58,** 493 (1946).
[9] Standard Specifications, 19 and 20, published by the American Society of Photogrammetry Nov. 15, 1935.
[10] I. C. Gardner, J. Opt. Soc. Am. **40,** 457 (1950).
[11] F. E. Washer and F. W. Rosberry, J. Opt. Soc. Am. **40,** 801 (1950); **41,** 597 (1951).
[12] R. N. Hotchkiss, F. E. Washer, and F. W. Rosberry, J. Opt. Soc. Am. **40,** 802 (1950).
'13] R. N. Hotchkiss, F. E. Washer, and F. W. Rosberry, J. Opt. Soc. Am. **41,** 600 (1951).

16. Theory of Resolving Power

By E. W. H. Selwyn [1]

Fundamental Theory

The test object for a measurement of resolving power is usually one in which the demarcation between areas of different brightness is a hard edge. A common and typical example consists of dark and bright bars. When an image of this is produced, either an aerial image or one in a sensitive photographic layer, light spreads from the bright bars into the dark. The diagram representing the brightness distribution in the test object is like the castellations on old fortresses, wherein the light bars are represented by short horizontal lines interposed between other short lines, representing the dark bars, spaced lower and connected by vertical lines. If the degree of spreading of the light is small the corresponding diagram for the image is somewhat similar but the corners are rounded and the vertical lines are replaced by sloping ones. As the spread of light is increased, relative to the spacing of the lines, the corners become more rounded, and the slope of the connecting lines becomes greater until the spread of light is about equal to the space between two light lines. A still greater spread of light then causes the hollows to start filling up and the tops of the diagram to drop. If we wish to be strict in terminology we may say that the first stages represent a diminution in sharpness and the last stage a diminution in contrast. When the last stage has been reached the top of the diagram representing the distribution of brightness approximates in shape to a curve of sines. If we start with a test object in which the graduation from light to dark lines is smooth and represented by a sine curve, only the contrast change can take place, for the effect of spread of light must then always be to raise the hollows and bring down the tops of the curve. Also the curve cannot change its character. It always remains a sine curve, with the same period. That this is necessarily so can be seen from the consideration that two sine curves of the same period always sum to another sine curve whatever the amplitudes of the original curves and whatever the phase difference between them. We may consider the spreading of the light as producing from the sine curve of the original test object a whole set of sine curves of equal period, with a total phase difference equal to the total spread of light and amplitudes varying according to the character of the spread of light. Thus the final curve for the image is also a sine curve with a lower amplitude (but of the same period). It may also be shifted in phase from that which would be appropriate if the image were perfectly sharp.

---

[1] Research Laboratories, Kodak Limited, Harrow, England.

Apart from any difference of phase, which is related to the optical distortion produced by the image forming system and is not of interest in the pure resolving power problem, the only physical variables involved are the period of the sine curve and the contrast. By eliminating the gradual change in shape of the light-distribution curve (by choosing a sine-curve test object in the first place) it becomes possible to proceed with theoretical arguments. This is mainly because the necessity of knowing how the changes of shape affect visibility of the image of the test object is avoided. The contrast in the image depends upon the interval between the bright bands, for as the test object is made smaller and smaller (the image following suit) the smudging-out effect of any unsharpness in image formation becomes more and more pronounced. In theoretical discussion the relation between the contrast in the image and the period of the sine curve is of fundamental importance. We shall find it convenient to consider first an optical system forming an aerial image.

The diagram representing the brightness of the test object or its image may be split into two parts, one representing a uniform brightness and the other the sine-curve portion, consisting of alternate positive and negative sections. The uniform brightness is unaffected by any lack of sharpness, but lack of sharpness diminishes the positive and negative sections of the sine-curve portion equally since they are of identical shapes. In other words, what happens to the amplitude of the sine-curve is independent of what happens to the mean brightness and if we change the mean brightness of the test object without changing the amplitude (in brightness) of the sine-curve component, the mean brightness in the image will be changed but the amplitude of the sine-curve component in the image will remain unchanged, and vice-versa. If a neutral filter is introduced between the test object and the lamp illuminating it the amplitude will be reduced (so also will the mean brightness but this is immaterial) and the amplitude in the image will also be reduced in the same proportion. Thus the ratio of the amplitudes in image and test object is constant for any given test object and optical arrangement. In fact, however, the amplitudes and mean brightnesses change in the same ratio, so that for both test object and image, the important quantity is the ratio of amplitude of brightness to mean brightness. If we call this ratio for the test object a and for the image a^*, then a^*/a is a constant for any one test object and optical system. Now the amplitude in the image is dependent upon the amplitude in the test object, upon the optical system and the spacing between bright stripes in the image. It will be convenient to use the reciprocal r of the last quantity. The above result therefore shows that

$$a^* = af(r), \qquad (1)$$

$f(r)$ being a function characteristic of the optical system, which may be calculated, at least in theory, from the specification of the optical system. The process of calculation starts from the brightness distribution in the image of a long line source. Each linear element of this image is regarded as producing a geometrically perfect image of the test object, a sine wave, that is, with phase appropriate to the distance of the linear element from the center of the image and amplitude proportional to the brightness of the element. All the sine

curves of this type are added. The above is the fundamental formula of the theory of resolving power when the test object is of the sine-curve form.

Visual Resolving Power

The value of r that makes $af(r)$ equal to the least value of $a*$ that can be appreciated by the eye, under the conditions of observation, is the visual resolving power of the system. The threshold value of $a*$ is not a quantity that can be calculated, a priori, but must be determined experimentally. It may depend on the effective aperture of the pupil of the eye, on the angular separation between the bright stripes in the image on the retina, and upon the mean brightness of the image, apart from other conditions such as the presence or absence of a bright surround to the image. Let us take a simple system consisting of a test object, a perfectly corrected lens forming an aerial image of it, a perfectly corrected eyepiece to make the emergent wave fronts flat, and the eye. The aerial image may be replaced by a luminous test object exactly similar to the original but with spacing between the bright stripes changed according to the magnification and with lower contrast (an amplitude in fact of $a*$), owing to loss of definition as a result of diffraction. That this is possible must be so because the wave front arising from any point is spherical, exactly as with the aerial image. But we have to assume that there is no diffraction by the eye, although there may be distortions of the wave front by refraction. The effective aperture of the eye is determined by the exit pupil of the lens and eyepiece system. With such an arrangement, using a test object with variable contrast, and placing the lens at different distances it is possible to determine a wide variety of conditions under which the image just ceases to be visible, as an array of bright and dark stripes. The value of $f(r)$ can be calculated, since the image-forming properties of a perfectly corrected lens are known from Airy's discussion of diffraction by such a lens. Thus it is possible to calculate $af(r)$ at the limit of resolution, the angular separation between bright bands, as seen by the eye, the effective diameter of the pupil of the eye (the exit pupil of the telescopic system), and the relative brightnesses of the various images. The results that were obtained by such an experiment are exhibited diagrammatically in figure 16.1. There appeared to be no influence of
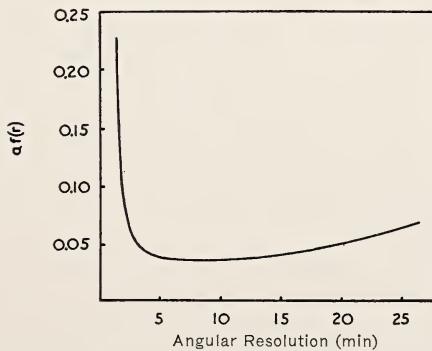

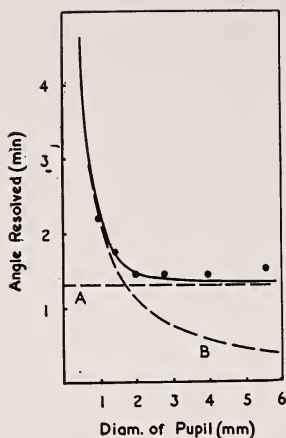Figure 16.1. *Visual threshold as a function of size of detail.*

FIGURE 16.2.   *Comparison of Cobb's experimental results with theory.*

A, Limit set by retina and/or nystagmus; B, limit set by diffraction at pupil.

brightness (over a range of 10,000:1 !) outside the experimental error. The striking feature about figure 16.1 is that the curve approaches an asymptote at around $1\frac{1}{4}$ minutes of arc, but that apart from this the threshold value of $af(r)$ is remarkably nearly constant. The simplest way of accounting for the asymptotic rise is to suppose that the eye averages out detail over a range of $1\frac{1}{4}$ minutes or so, in consequence of coarseness of structure of the retina, or involuntary movements (nystagmus) or both.

There is no evidence in the results that the value of $af(r)$ at the threshold varies with pupil aperture over a range of 0.1 to 2.0 mm and it is probably legitimate to assume that below pupil apertures of 2 mm the eye is perfectly corrected. The light used in the experiments was monochromatic and it may be that this is the reason why the eye appears to be perfectly corrected to an aperture as large as 2 mm.

If it is assumed that the eye is perfectly corrected and that the retina averages out the illumination over $1\frac{1}{4}$ minutes of arc, the value of a^* in the image transmitted by the retina to the brain may be calculated, in terms of brightness, from the formula

$$a\,f_1(r).f_2(r),$$

where $f_1(r)$ is a factor calculated for diffraction by the eye and $f_2(r)$ a factor calculated for the averaging effect of the retina. If now a^* is taken as 0.03, and a unity, corresponding to a high-contrast test object, the results of Cobb on the resolving power of the eye at different pupil apertures may be well duplicated, except that at apertures above say 3 mm the observed resolving power is rather less than that calculated, presumably owing to the greater effect of aberrations at higher apertures (fig. 16.2).

Photographic Resolving Power

For some purposes the theory of photographic resolving power can be very simple. If sine curve distributions of light, of different contrasts, were projected on a photographic material it would be possible to determine the relation between the resolution of the developed

222

image and the contrast of the light distribution. Suppose that the results are exhibited in the form of a curve relating the minimum contrast of incident distribution necessary for the light and dark stripes to be just visible in the developed imaged and r the reciprocal of the distance between neighboring light stripes. Now draw on the same diagram the curve showing the relation between the contrast in the aerial image produced by an optical system and r. At the point where the two curves cross, the optical system is producing an image with spaces between stripes equal to those at which resolution just occurs with an image on the material of the same contrast as that produced by the optical system. The value of r at the common point is therefore the resolving power of the combined optical system and sensitive material.

By this means we have accomplished a synthesis between the purely optical and the purely photographic properties of the combined system. In principle, if the photographic curves for N photographic materials and the optical curves for N^* optical systems were known we could immediately find the resolving power for all the NN^* combinations, for any and every contrast of test object.

There is, however, a difficulty. The above argument is sound if the light is monochromatic. But if it is not monochromatic we may be faced with a situation of the following type. Suppose we have a lens giving an aerial image of a point source consisting of a bright central dot of red light surrounded by a thin blue halo. This image is first supposed to be projected on a sensitive layer that diffuses red light more than blue. The image in the emulsion layer will then consist of a diffuse red central area with a moderately diffused blue halo. Now suppose the aerial image is projected on a sensitive layer, with the same diffusing properties, for white light as a whole, as the preceding one, but which diffuses blue light more than red. The image in the emulsion layer will then be a fairly small central red spot with a diffuse blue halo. When converted to black and white (by the process of development) these two images will be quite different. In other words we must take account of the distribution of color in the aerial image and of the variation with color of the diffusion of light in the sensitive layer. In practice this means that one needs to calculate the amplitude of the sine distribution in the sensitive layer for every wavelength of which the light is composed and then average according to the color sensitivity of the material. Alternatively one may calculate the distribution of actinic energy in the sensitive layer of the image of a long thin line of light and use this for calculating the amplitude of the sine-curve image.

Resolving Power of Sensitive Material

The amplitude of a sine-curve distribution of light projected on the sensitive layer is reduced by diffusion of light in the layer, according to the formula,

$$a^{**}=a^*\varphi(r), \tag{2}$$

analogous to that for an optically formed image. The process of development converts small percentage variations in exposure into the same percentage variations in transparency, multiplied by γ, the

slope of the usual density versus log exposure curve of the material at the density at which the mean exposure is reproduced. The amplitude of the image after development is therefore $\gamma a^* \varphi(r)$. This is true only if the amplitude is small, otherwise there is some distortion of the sine curve. If the photograph is magnified M times the apparent separation between bright stripes in the image is increased in the ratio M and the reciprocal of the separation is decreased, so far as the eye is concerned, to r/M. Thus the amplitude of the signal to the brain, in terms of light intensity, is reduced to

$$\gamma a^* \varphi(r) F(r/M),$$

where $F(r/M)$ is a factor (equivalent to $f_1(r)f_2(r)$, used in discussing Cobb's results on variation of acuity with pupil aperture) to take account of the unsharpness in image formation in the eye and the averaging effect of the retina.

So far, everything has proceeded on the assumption that the image is "smooth" that is to say, without taking any account of the irregular distribution of grains in the photographic material. In order to deal with the effects of this irregular distribution we imagine that the surface of a uniformly exposed and processed material is divided up by a grid into small equal and contiguous areas and that the densities of all these areas are measured. These densities will be different one from the other owing to the irregularity of distribution of grains. We ask, "What is the statistical distribution of density and in particular what is the root mean square fluctuation of density?" We suppose that the variation from the mean of the density of any elementary area of the grid is small enough for the density of a pair of areas taken together to be equal to the average of the densities of the pair.

If the fluctuations of density from the mean for a long set of N alternate areas of the grid be denoted by symbols like δ_1 the mean fluctuation $\Sigma\delta_1/N$ is zero and the mean square fluctuation Δ^2 is equal to $\Sigma\delta_1^2/N$. For the set of areas interposed between the preceding areas, with fluctuations denoted by symbols like δ_2 the mean fluctuation $\Sigma\delta_2/N$ is again zero and the mean square fluctuation is again equal to Δ^2 and to $\Sigma\delta_2^2/N$. The mean fluctuation for the pairs of areas taken together is $\Sigma(\delta_1+\delta_2)/2N$ and is zero and the mean-square fluctuation is $\Sigma(\delta_1+\delta_2)^2/4N$, which is equal to $\Sigma\delta_1^2/4N+\Sigma\delta_2^2/4N+\Sigma\delta_1\delta_2/2N$. If the distribution of fluctuations is symmetrical $\delta_1\delta_2$ is as likely to be positive as negative, so that in a long sequence $\Sigma\delta_1\delta_2/4N$ is zero. Thus the mean square fluctuation for doubled areas is $\Delta^2/2$. From this it follows that the root mean-square fluctuation is inversely as the square root of the elementary area of the grid, or

$$\Delta = G/\sqrt{\alpha},$$

where G is a constant, for the given sample of material, called the "granularity" and α the area of an element of the grid.

There is a peculiarity about density, since it may legitimately be regarded as having a "dimension" in the same way as mass. For in just the same way as a mass may be measured by equilibrating it against a number of arbitrary units of mass, a density may be measured by equilibrating it photometrically against a stack of units of

224

density, such as pieces of neutral glass all of the same density. The choice of unit is arbitrary. In terms of the definition of density, namely

$$D = \log (1/T),$$

(where D is the density and T the transmission) all this means is that we have chosen the base of the logarithm arbitrarily. A much more elaborate argument than the above, based on assuming a dimension for density and assuming that the fluctuation of density is controlled by a constant G and that the densities of neighboring areas are uncorrelated, yields both the statistical distribution and the formula

$$\Delta = G/\sqrt{\alpha}.$$

The question now arises whether it is possible to estimate the nature of the signal to the brain obtained from grainy photographic material. We suppose that the retina constitutes a grid with elements having areas determined by the disk of confusion (mainly the Airy disk) of the optical apparatus of the eye and the averaging effect of the retina, already discussed. The corresponding areas on the material itself may be taken as α^*/M^2, where α^* is the area of an element of the retinal grid projected on the material at unit magnification and M the magnification. The signal to the brain, in terms of density fluctuation, is therefore

$$MG/\sqrt{\alpha^*}.$$

The signal from the sine-curve image to the brain is, in terms of light intensity,

$$B\{ 1 + \gamma a^* \varphi(r) F(r/M) \sin 2\pi r x \},$$

where B is the mean light intensity and x distance across the bright and dark stripes. Now the appearance of almost any object, a picture, for example, of the above sine-curve distribution of light, is practically independent of the light intensity (except for very low and very high intensities). We may therefore legitimately take the logarithm of the above expression and drop the term $\log B$, thus giving the equivalent total density fluctuation

$$2 \log_{10} e \gamma a^* \varphi(r) F(r/M),$$

provided the amplitude of the sine curve is small.

Now it is easy to show, for instance by superimposing a sample of grainy photographic material upon a grainless transparency, that the irregularity in density reduces the visibility of detail, and we therefore have to enquire what the relation is between the sine-curve density fluctuation, and the density fluctuation due to granularity at which the sine-curve density fluctuation is just visible. Since the two terms in the relation have the same dimension, namely of density, it follows immediately by the very simplest of dimensional arguments that [2]

---

[2] If the limiting density difference, δ, that the eye can perceive be introduced into the dimensional argument, the right-hand term must be multiplied by an arbitrary function of $\delta \sqrt{\alpha^*}/GM$. Other calculations show that this function is constant when the parameter is small.

$$2 \log_{10} e\gamma a^* \varphi(r) F(r/M) = Const. \; GM/\sqrt{\alpha^*}.$$

The resolving power is to be calculated from this by solving for r. The value obtained is dependent on the magnification. If $F(r/M)$ were proportional to M/r, the observed resolving power would be independent of the magnification, since M would then disappear from the equation. But if the magnification be chosen to obtain the maximum value of r, the value of r will remain the same for small changes in magnification. Thus the maximum value of r will be secured at the value of M/r at which the curve of $F(r/M)$ against r/M contacts a rectangular hyperbola (i. e., a curve of $Const.M/r$ against r/M). This occurs at the point at which the slope of log $F(r/M)$ against log r/M is $-45°$. This point obviously specifies a constant ratio between r and M. That is to say maximum visibility and maximum resolving power are attained when the magnification is chosen so that the apparent separation between stripes is always the same and of a certain magnitude.

In these circumstances the equation becomes

$$\gamma a^* \varphi(r) = Const. \; Gr. \tag{3}$$

The value of the constant has been estimated, from what is known about the eye, at 0.003. The value of a^* is determined by the contrast of the original test object, and the optical apparatus forming the image, according to eq 1.

Thus the complete formula for the photographic resolving power is

$$\gamma a f(r) \varphi(r) = Const. \; Gr. \tag{4}$$

A graphical illustration of the use of this formula is shown in figure 16.3, for a high contrast test object, for which a is unity. The visual resolving power in the aerial image produced by the optical apparatus is the value of r at the point of intersection A of $f(r)$ with the visual
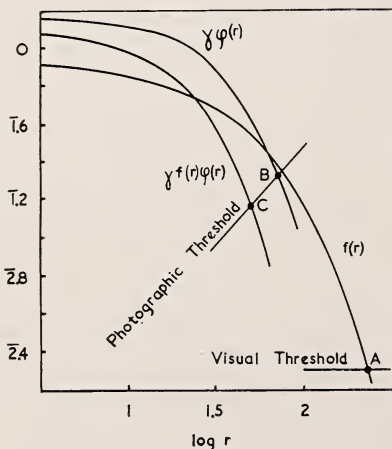
FIGURE 16.3. *Graphical illustration of theory of resolving power.*

A, Visual resolving power of optical system; B, resolving power of photographic material; C, resolving power of photographic material and optical system combined.

226

threshold, the resolving power of the photographic material is determined by the intersection B, at which eq 3 holds, and the resolving power in the negative produced by the optical system by the intersection C, at which eq 4 holds. It is obvious that there is no simple connection between the visual resolving power and the photographic resolving power.

Results

The value of a theory is either that it enables results to be calculated, or that it enables one to understand a phenomenon or group of phenomena. Obviously, the function of a theory of resolving power is to permit one to understand what is happening, for it is much easier to measure the resolving power in any given case than to calculate it. In spite of its defects (which will be mentioned shortly) the theory given here gives an account, of certain features found in resolving-power measurements, sufficiently nearly quantitative to justify the belief that any physical properties other than those discussed have relatively unimportant effects.

The first result to be discussed concerns the variation with aperture of the resolving power in the center of the field of a photographic objective. The aberrations of such a lens are relatively unimportant at small apertures such as F/45. The function $f(r)$ may thus be taken as identical with that for a perfectly corrected lens. When the aperture is increased a little the central core in the image of a point remains about the same in size as the Airy disk appropriate to the aperture under consideration, but there is a greater amount of light in the outer parts of the image than when the lens is perfectly corrected. The main effect of this is to decrease the value of $f(r)$ by a roughly constant factor for moderate and large values of r. As the aperture is still further increased the image becomes more extended and the central core may disappear or almost disappear. Then $f(r)$ is diminished still more, its value at large values of r diminishing greatly. An estimate of its value at various values of r may be obtained from estimates of the visual resolving power of the lens with test objects of different contrasts. It is not difficult, if this be done for the maximum aperture, to estimate graphically a set of curves of $f(r)$ for different apertures. The form of $\varphi(r)$ for photographic materials may be found if it is assumed (as is commonly done in photographic theory) that the sideways scatter of light in the sensitive layer follows a negative exponential law. The curve for any given material may then be obtained from a knowledge of the granularity G and the resolving power, using eq 3. Sufficient information is thus provided to determine, by the method illustrated in figure 16.3, the resolving power at any aperture. Results of such a calculation are shown in figure 16.4. The visual resolving power for high-contrast test objects at moderate and small apertures approximates closely to that expected from Rayleigh's formula, while the decrease caused by diminishing the contrast of the test object is not great. The resolving power at high apertures is markedly below the theoretical. With a material of low granularity the resolving power is much decreased and the maximum occurs at a smaller aperture than that at which maximum visual resolving power is found. With materials of greater granularity the resolving power is still more decreased and the curve against aperture becomes flatter. These results are of a general character
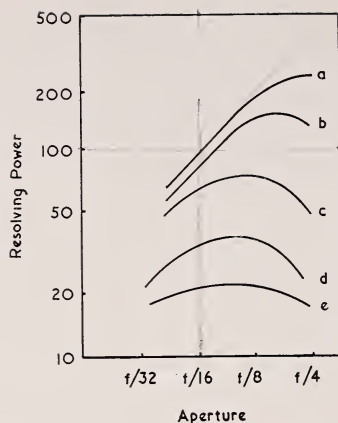
FIGURE 16.4. *Resolving power of camera objective as a function of aperture.*

a, Visual resolving power, high-contrast test object; b, visual resolving power, low-contrast test object; c, photographic resolving power, fine-grain film, low-contrast test object; d, photographic resolving power, medum-grain film, low-contrast test object; e, photographic resolving power, coarse-grain film, low-contrast test object.
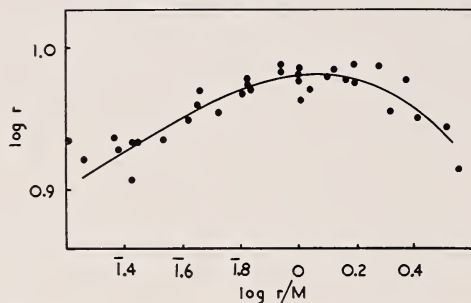

FIGURE 16.5. *Effect of magnification on resolving power.*

and have been observed with all lenses that have been tested in these laboratories. The resolving power observed varies with the density of the processed material. The above results refer to a density of 1.0, at about which density the resolving power is a maximum.

The results obtained by experiments designed to check the theory of photographic resolving power also provided evidence that the observed resolving power depended upon the magnification used. Typical results are shown in figure 16.5. The points shown consist of six sets in which the resolving power varies over a range of 5:1, but in the diagram as shown the six sets of points have been adjusted along the log r direction so as to have the same mean value. It is clear from the diagram that maximum visibility is attained when the apparent separation between bright stripes is 1 mm, at the standard viewing distance of 250 mm. As expected from theory this is a universal condition independent of the actual separation and the granularity, but calculations from $F(r/M)$ suggest a value four times less for the magnification. The reason for the discrepancy is not known, but it is known that different observers prefer different magnifications. The above condition, namely making the apparent size of detail it is

wished to see about 1 mm, is not far from correct for maximum visibility on most photographic records.

If one could provide a diffusing screen with turbidity identical with that of a photographic material, and introduce granularity into the image, the image seen should be identical with that seen on a negative developed so that $\gamma = 1$. A difference of γ from unity can be allowed for by changing the contrast of the test object. Apparatus of this kind has, in fact, been devised, and although the test object consisted of two short bars of the type used by Cobb, and not the sine-wave form, the resolving powers so observed were the same as the photographic resolving powers within a standard deviation of 15 percent. This is not much more than can be accounted for by experimental and observational error. The gain in speed over photographic measurement is substantial.

Evaluation

There are certain defects in the preceding theory. The most outstanding are the disagreement between the calculated optimum magnification and that found by experiment as shown in figures 16.5, and the disagreement between the theoretical estimate of the photographic threshold, involving granularity, and experimental determinations. The latter are shown in figure 16.6. Curve B is the theoretical result (eq. 3). Curve A refers to a few estimates made by determining corresponding values of G and r at which images of Cobb test objects projected optically on samples of uniformly exposed and processed photographic materials ceased to be visible. Curve C refers to exposures made with sine-curve test objects of different contrast analysed in such a way as to give the threshold almost directly. Full details are given in the original publication. There is agreement in order of magnitude and in the general result that the photographic threshold increases with Gr approximately according to a power law with an exponent of the order of unity. Perhaps this is as much as can be expected from photographic experiments of a type that is notoriously liable to both systematic and accidental errors. There are, however, stronger reasons than these for supposing that
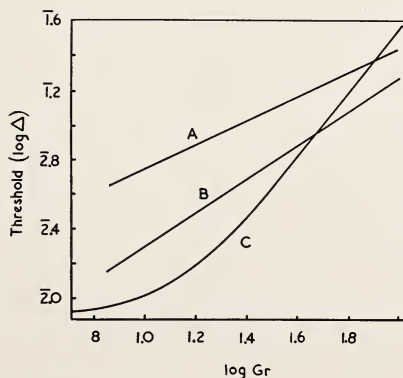

FIGURE 16.6. *Photographic threshold (log Δ) as a function of Gr.*

the theory may be defective. Jones and Higgins have demonstrated that granularity figures, obtained by the method used (theoretically described in principle in this paper) do not correlate well with figures for the graininess, the corresponding quantity estimated by purely visual observation, and have also shown that the theoretical description is not followed when α is small. It seems likely that the density of any given small area on the photographic material is not independent of that of neighboring areas. In other words there is correlation, which may formally be described by a quantity with the dimension of length, corresponding with the distance over which the correlation extends. The theory as given takes no account of such a quantity, nor indeed does any theoretical treatment of granularity yet proposed. Correlation might give the grainy structure a character akin to periodicity and this could well affect both the threshold and the optimum magnification.

References

E. W. H. Selwyn, Phot. J. [B] **88,** 6, 46 (1948).
P. W. Cobb, Am. J. Physiol. **36,** 335 (1915).
E. W. H. Selwyn, Phot. J. **75,** 571 (1935).
E. W. H. Selwyn and J. L. Tearle, Proc. Phys. Soc. **58,** 493 (1946).
W. Romer and E. W. H. Selwyn, Phot. J. **83,** 17 (1943).
L. A. Jones and G. C. Higgins, J. Opt. Soc. Am. **35,** 435 (1945); **36,** 203 (1946).

17. A New System of Measuring and Specifying Image Definition

By O. H. Schade [1]

Introduction

An objective analysis of image definition leads fundamentally to methods for determining the geometrical properties or an equivalent measure of the point or "star" image of the imaging device. When the intensity distribution of the point image is invariable, it is obvious that image quality and apparent sharpness vary inversely with a significant diameter of the point image. When point images having different intensity distributions are compared, however, this relation is no longer true. It then becomes necessary to specify the relative intensity distribution as well as a diameter. The problem is further complicated by the fact that the point image of practical lenses departs considerably from circular symmetry. Figure 17.1 illustrates the changes in the geometry of the axial-point image in successive stages of a motion-picture process. From a study of resolution and detail contract characteristics of photographic and television images it has long been apparent that the sharpness of an image has no fixed relation to the limit of resolution of the system but depends rather on the steepness and form of the intensity or luminance curve representing a unit function transition, i. e., a sharp edge. It is not difficult to see that the total length of the transition is equal to the diameter of the point image (fig. 17.1), while shape and gradient depend on the *energy distribution* in the point image. It must be appreciated that the energy in a point image is a function of its volume (considering intensity of the third dimension), and that a low-intensity disk of light surrounding a small high-intensity center may contain as much or more light flux as the bright center of the light spot and thus produce a gradual transition with rounded corners. The transition curve that is readily generated by displacing the point image over a brightness step is, therefore, a much more sensitive measure of quality than the intensity distribution in the point image. The maximum gradient of a single transition, however, is not a sufficient specification of image quality. Furthermore, the shape of the transition is affected when several edges (or repetitive contours) occur within a distance less than the maximum spot diameter. It is known from communication theory that any complex transition or waveform can be expressed by its Fourier components. The properties of the point image can thus be specified accurately by a response characteristic to optical sine waves ranging in wavelength from infinity to zero.

INTENSITY
DIAMETER   DISTRIBUTION   TRANSITION
CROSS SECTION   AT EDGE

0   POINT SOURCE OF LIGHT

1   CAMERA LENS

2   1 + NEG. FILM

3   2 + POS. FILM
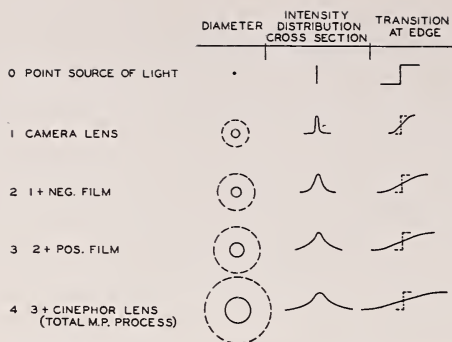
4   3 + CINEPHOR LENS
    (TOTAL M.P. PROCESS)

FIGURE 17.1.   *Point images and edge transitions in a photographic process.*

Response Characteristics of the Point Image

Normal optical-bar test patterns are "square-wave" flux patterns. Accurate sine-wave test patterns can be obtained from variable-density recordings of electrical constant-amplitude sine-wave frequencies on the sound track of motion-picture film. A number of sections so obtained are shown in figure 17.2. It is customary to express the wavelength in these patterns in reciprocal units: the *line number* N. The line number N will be defined in this paper as the number of half-waves or "lines" (dark and light) in a length unit. The line number has the dimension length^{-1}. When the geometrical properties of the point image are known, its response to sine-wave or square-wave flux patterns is readily computed by scanning the flux pattern with the point image, considering it as a scanning aperture. The point image is thus defined as the *resolving aperture* or *sampling aperture* of the image-forming device referred to the image plane, the intensity distribution in the point image is represented by the transmittance τ of the aperture.

Response characteristics of a number of aperture types are shown in figures 17.3 to 17.8. The line number is given in relative units N_δ defined by the equation

$$N_\delta = l/\delta, \tag{1}$$

where l is the unit of length ($l = 1$ mm) and δ is a significant diameter of the aperture. The sine-wave or square-wave response is likewise given in relative units. The sine-wave response factor $r_{\tilde{\psi}}$ is defined as the ratio of the sinusoidal aperture flux $\tilde{\psi}_N$ at a line number N to the sinusoidal flux $\tilde{\psi}_0$ at a line number N approaching zero as a limit as expressed by

$$r_{\tilde{\psi}} = \tilde{\psi}_N / \tilde{\psi}_0. \tag{2}$$

The square-wave flux response factor $r_{\Delta\bar{\psi}}$ is defined similarly by

$$r_{\Delta\bar{\psi}} = \Delta\bar{\psi}_N / \Delta\bar{\psi}_0. \tag{3}$$

The symbol $\Delta\bar{\psi}$ indicates that the square-wave response is measured not by the peak amplitude [2] of the waveform but by the differential

[2] The square-wave amplitude response $r_{\Delta\hat{\psi}}$ is shown for comparison in some figures.
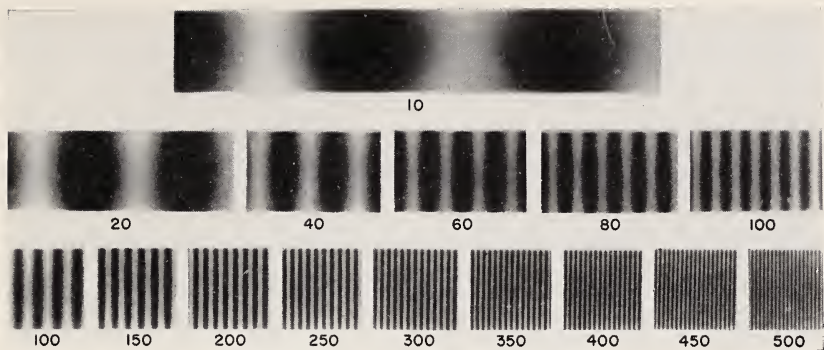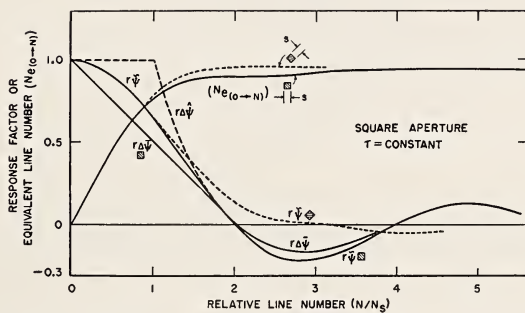
FIGURE 17.2.   *Sine-wave test pattern.*

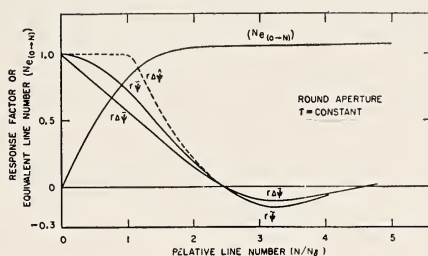FIGURE 17.3.   *Response characteristics of square aperture* $(\tau=1)$.

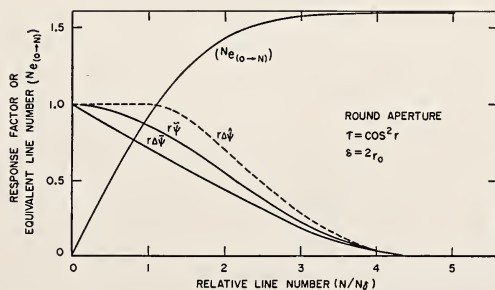FIGURE 17.4.   *Response characteristics of round aperture* $(\tau=1)$.

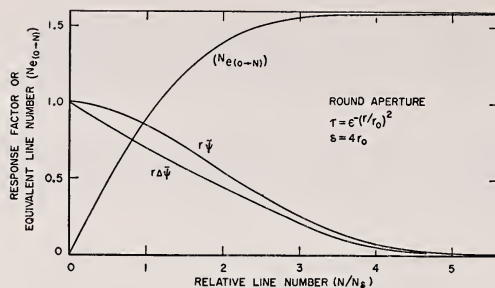FIGURE 17.5.   *Response characteristics of round aperture* $(\tau=cos^2 r)$.

233

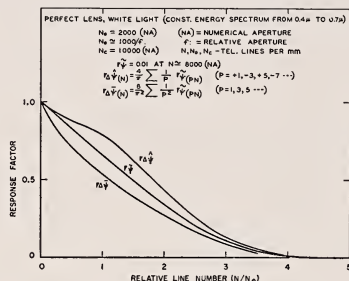FIGURE 17.6. *Response characteristics of round aperture* $(\tau = \epsilon^{-(r/r_0)^2})$.

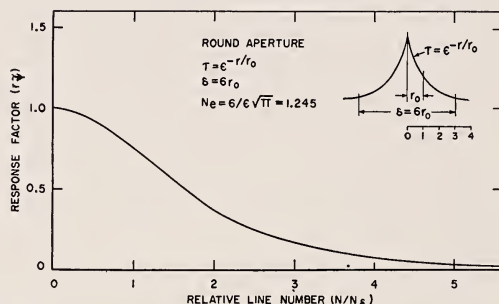FIGURE 17.7. *Response characteristics of theoretical lens (white light)*.

FIGURE 17.8. *Response characteristics of round aperture* $(\tau = \epsilon^{-r/r_0})$.

flux, which is proportional to the mean value or area under the recti-
fied half-waves of the waveform, because the latter changes from a
rectangle to trapezoids and triangles or sinusoidal shapes when N
is varied from zero to infinity. Both response factors are single
valued and independent of contrast, provided waveform distortion is
avoided by the use of small "signals" when nonlinear devices such as
photographic films are measured. The use of sine-wave patterns
has the advantage of permitting a direct and accurate evaluation of
the overall response of systems containing a number of imaging
stages in cascade. The over-all sine-wave response factors of the
system are simply the product of the response factors of the com-
ponents at corresponding line numbers. This advantage is lost for
square-wave response factors, because of the variable harmonic content
of the waveforms. Inspection of the computed characteristics,

234

figures 17.3* to 17.8, however, shows that the ratio $r_{\hat{\psi}}/r_{\Delta\bar{\psi}}$ is substantially independent of the aperture type at given values $r_{\Delta\bar{\psi}}$, permitting a point by point conversion of square-wave to sine-wave response characteristics with satisfactory accuracy.

The response characteristic of a theoretical aberration-free lens limited only by diffraction is of interest as a standard of comparison. The characteristic (fig 17.7) shows the "white"-light response computed for a coaxial superposition of diffraction spots for a constant energy spectrum for 0.4 μ to 0.7 μ.[3] The round aperture with a transmittance $\tau = \epsilon^{-r/r_0}$ (fig. 17.8) is of interest as a mathematical equivalent for the aperture effect of grain structures in which the intensity from an infinitesimal pencil of light decreases exponentially from its center outward because of diffusion, absorption, and diffraction in the grain layer; the magnitude of these optical effects (diameter of the aperture) depends on the layer thickness, transparency, spacing, and size of the particles. The theory is well substantiated by measurements on a variety of film types and kinescope phosphors.[3] Plotted to a normalized scale with reference to the rated resolving power N_{cr} of photographic film[4] at which $r_{\hat{\psi}} \simeq 0.02$, the measured values result in a single curve shown in figure 17.9. A comparison with figure 17.8 shows excellent agreement when N_{cr} is placed at $N/N_\delta = 5.17$.

Another characteristic of special interest is that of the round aperture with $\tau = \epsilon^{-(r/r_0)^2}$ (fig. 17.6), which has the form $r_{\hat{\psi}} = \epsilon^{-kN^2}$. It can be seen from this form that a repetition or cascading of such aperture processes will not change the shape of the characteristic, but only the constant k.

The response characteristics of practical lenses, microscopes, television tubes, and other image-forming devices can be measured accurately by photoelectric methods, either by scanning test patterns with their point image or by using a photoelectric or electronic microphotometer[3] to scan the test-pattern image formed by the device. One of these latter methods developed by the author appears to be particularly simple and flexible for testing lenses and was described and demonstrated.

The quantitative information obtained from response characteristics is demonstrated by table 17.1 showing the changes in the square-
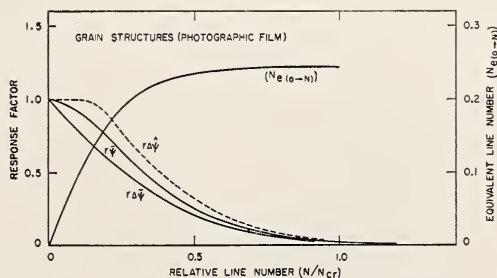

FIGURE 17.9.   *Response characteristics of grain structures (measured).*

[3] Compare first reference on page 249.
[4] The photographic value N_{cr} (lines/mm) is to be multiplied by a factor of two according to the definition of N used in this paper.
* The broken line curve for $N/N_* > 2.8$ should be positive in figure 17.3.

Table 17.1 Square-wave response $(r_\Delta \overline{\psi})$ of 8.3-mm apochromat, NA=0.65 with 10x compensating eye piece

(White light)

N/mm	Theoretical*		$r_\Delta \overline{\psi}$				
	r_ψ	$r_{\Delta\overline{\psi}}$	Cover glass_____0.18 Tube length: (mm) 160	0.127 160	0.127 172	None 160	None 196
100	0.98	0.95	0.925	0.91	0.925	0.76	0.89
250	.94	.88	.81	.78	.81	.49	.76
500	.87	.78	.655	.57	.648	.34	.585
700	.83	.72	.565	.45	.555	.25	.485
1,000	.75	.64	.45	.31	.44	.195	.37
1,500	.64	.53	.30	.17	.30	.13	.225
2,000	.52	.43	.21	.13	.21	.08	.135
2,500	.42	.34	.14	.12	.14	.03	.07
3,000	.32	.26	.095	.11	.095	.01	.033
3,500	.22	.18	.055	.065	.055	.0	.018
4,000	.14	.11	.036	.038	.036	.0	.002
4,500	.07	.06	.018	.018	.018	.0	.0
5,000	.03	.025	.008	.008	.008	.0	.0
5,500	.01	.008	.0	.0	.0	.0	.0

*Computed for 2,870° K source and 1P21 multiplier phototube.

wave response $(r_\Delta \overline{\psi})$ of an 8.3-mm Bausch & Lomb Apochromat (N. A.=0.65) used with a 10x compensating eye piece at a magnification of 300 when cover-glass thickness and tube length are varied. The values have not been corrected for aberrations in the measuring microscope (4-mm Apochromat N. A.=0.95 with 25x compensating eye piece). It is of interest to the lens designer to note that a response characteristic (sine wave or square wave) can be considered as the sum of a number of component response characteristics of known shape thereby permitting a reconstruction of the point image. This process is illustrated on the characteristic of a 40-mm Ciné Ektar lens shown in figure 17.10. The measured characteristic (at 5°) can be regarded as the sum of response characteristics from $\cos^2$ apertures. The first component, curve 2, is obtained by continuing the "tail" end of the curve as a $\cos^2$ response curve. Subtraction of this component from the normal characteristic leaves curve 1, which closely matches the response characteristic of a second $\cos^2$ aperture. The flux division (ψ_1 and ψ_2) between the corresponding coaxial spots is indicated by the response at $N=0$ and happens to be equal for both components. The diameters of the corresponding $\cos^2$ spots follow from the respective N_δ values (eq 1); the relative flux intensities
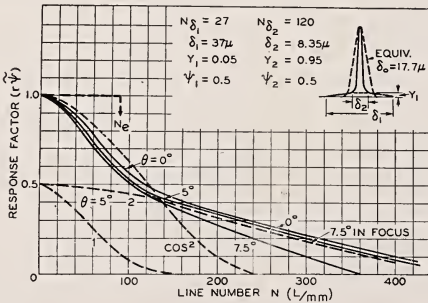

FIGURE 17.10. Response characteristics of a camera lens for 16 millimeter film.
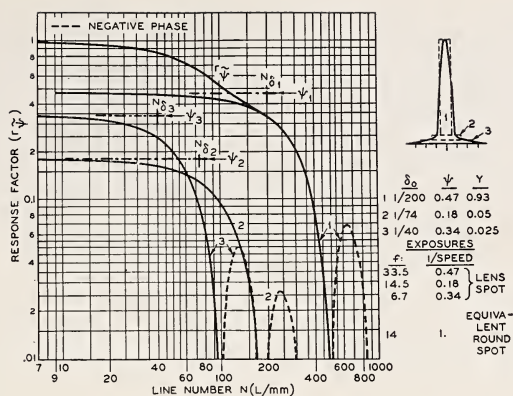
FIGURE 17.11. *Synthesis of response characteristic from three round aperture characteristics* ($\tau = 1$).

(Y_1 and Y_2) at the spot centers are obtained with

$$Y_1/Y_2 = \psi_1 \delta_2^2 / \psi_2 \delta_1^2 \quad \text{and} \quad Y_1 + Y_2 = 1. \tag{4}$$

Addition of the cosine-squared intensity curves of the components furnishes the approximate intensity in the point image from the lens. A practical method for performing this analysis is shown in figure 17.11 illustrating a breakdown into round apertures with uniform transmittance that appear as coaxial cylinders in a topographic representation of the point image. The response characteristic of the lens is drawn in log-log coordinates (curve $r_{\tilde{\psi}}$ in fig. 17.11). The response characteristic of the round aperture (fig. 17.4) is drawn on a separate sheet of log-log paper. The two curves are superimposed and a position matching the tail end of the lens curve is found. This portion (curve 1 in fig. 17.11) is subtracted and the process is repeated. In this particular case, the two additional component curves shown result. Flux values, diameters, and intensities are then found as outlined above and shown in the insert of figure 17.11. It should be noted that the negative-phase portions of the round aperture response must be included with negative sign when making the summation. The process is reversible, illustrating a synthesis of the sine-wave response from a complex symmetric aperture as the sum of three component characteristics. It is not difficult to see that the response of an asymmetric aperture can be obtained by the same process of subdivision into cylindrical sections having displaced axes. This displacement is equal to the displacement of all sine waves between respective response characteristics, and causes a phase shift proportional to N that must be duly considered when the summation is made.

Evaluation of a Measure of Equivalence (N_e) from the Response Characteristic

The response characteristics of practical imaging devices depart in general more or less from theoretical characteristics and may differ considerably in the relative response between low and high line numbers. In the case of optical lenses the response may decrease more

or less rapidly at low line numbers because of aberrations, although the theoretical resolving power may still be measurable in some cases.

The differences in response characteristics and the condition of astigmatism indicates that a comparison of image definition by a single figure of merit requires the evaluation of a measure such as an equivalent response characteristic or an *equivalent aperture* with standard transmittance that can be specified by one significant number. It is important that this equivalent measure agree with a visual impression of sharpness. In an objective evaluation the "quality" of an image or, more general, the quality of a device or system transmitting information (code, sound, or pictures) is specified by three fundamental characteristics: Transfer characteristic, fluctuation or "noise" level, and sine-wave response characteristic. The seemingly corresponding visual impressions of tone range, graininess, and sharpness in optical images, however, are not exact equivalents, because they represent combinations of the above objective characteristics. A judgment of contour sharpness, for example, is influenced by graininess and by the amplitude of the brightness step (contrast), whereas the sine-wave response characteristic is a *relative measure* describing shape and length of the brightness transition only. In an objective system of ratings, amplitudes, transmittance, "flare", wave-form distortion, rectification, etc., are specified or determined separately by the transfer characteristic, which, in general, is not controlled by the same parameters determining the frequency response of a system. A low-reflection coating of lens elements, for example, improves the light-transfer characteristic of the lens, but has no effect on the sine-wave response characteristic. In order to agree with an objective measure of definition, a visual comparison must be arranged for judging the *relative sharpness* of images and requires that image content, transfer characteristic (luminance and contrast) and fluctuation level in the images are made substantially identical. The impression of relative sharpness in an image can then be considered as a visual evaluation of the complex waveforms from a large number of arbitrary cross sections of the intensity distribution in an image frame. In an ideally sharp image the sum of the Fourier components of these cross sections fills an infinite sine-wave spectrum with constant amplitudes. Superposition of all components in random phase relation results in a most general and interesting test pattern, appearing optically as a *random grain structure*. Any cross section of this structure taken by a microphotometer with infinitesimal aperture is a complex wave containing a constant-amplitude sine-wave spectrum
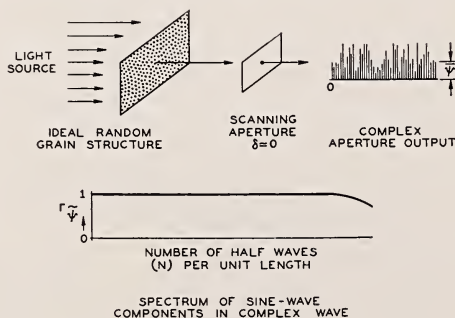

FIGURE 17.12. *Fourier components in ideal random grain structure.*

238

up to a very high line number, as illustrated by figure 17.12. An imaging device with a finite scanning or sampling aperture integrates high-"frequency" components, as illustrated by figure 17.13, a harmonic analysis furnishing the sine-wave response characteristic of the device. The complex wave contains a variational or modulation component $[\tilde{\psi}]$ that can be measured in total by an rms current meter, and a steady component $\bar{\psi}$ that can be measured by a d-c meter. Their ratio, known electrically as the "noise-to-signal" ratio, is *the relative deviation*

$$\sigma = [\tilde{\psi}]/\bar{\psi} = \left[\int_0^\infty \tilde{\psi}_{(N)}^2 \, dN \right]^{\frac{1}{2}} \Big/ \bar{\psi}. \tag{5}$$

When the sine-wave response is normalized according to eq 2 so that $\tilde{\psi}_{(N)} = 1$ at $N = 0$, the mean squared deviation is expressed by

$$[\tilde{\psi}]^2 = \tilde{\psi}_0^2 \int_0^\infty (r_{\tilde{\psi}})_{(N)}^2 \, dN, \tag{6}$$

where $r_{\tilde{\psi}}$ is the sine-wave response factor and $\tilde{\psi}_0$ is a measure of the magnitude of the a-c flux passing through the aperture at a line number N approaching zero. A hypothetical aperture having a constant response ($r_{\tilde{\psi}} = 1$) from $N = 0$ to a line number N_e^{*} [5] where the response drops abruptly to zero, would give a mean-squared deviation

$$[\tilde{\psi}]^2 = \psi_0^2 N_e^{*}.$$

The integral of squared response factors in eq 6 may hence be interpreted as a normalized mean-squared deviation or as an *equivalent passband* of constant amplitude extending to the line number N_e^{*} as defined by

$$N_e^{*} = [\tilde{\psi}]^2 / \tilde{\psi}_0^2 = \int_0^\infty (r_{\tilde{\psi}})_{(N)}^2 \, dN. \tag{7}$$

The measure N_e^{*} has the dimension length^{-1}. Its reciprocal value expresses an equivalent length or diameter of the aperture *in the scanning direction*. Like the aperture response, N_e^{*} depends, in general, on the aperture orientation relative to the direction x of aperture displacement. Apertures with circular symmetry have a
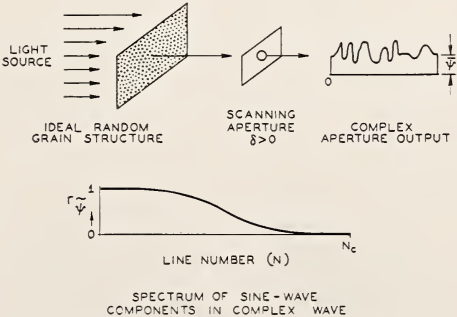

FIGURE 17.13.   *Fourier components in grain structure integrated by a finite aperture.*

[5] The asterisk on the value N_e^{*} is used to indicate that this value is obtained when a random grain structure is scanned. Other values will be introduced subsequently.

239

single effective length proportional to their diameter δ and a single value N_e^*. Elliptical or rectangular apertures can be specified by two values $N_{e(a)}^*$ and $N_{e(b)}^*$ obtained by orienting their major or minor dimensions (a or b) in the direction of scanning. These two values can be combined into a single value

$$\overline{N}_e^* = (N_{e(a)}^* N_{e(b)}^*)^{1/2} \tag{8}$$

representing an equivalent symmetric aperture.

The direct evaluation of the measure N_e^* for an unknown aperture requires a calibrated random grain pattern that must be tested by a harmonic analysis of the complex aperture output. A practical alternative is a synthesis of the sine-wave characteristic from the aperture response to constant amplitude sine-wave patterns of various wavelengths and an evaluation of N_e by eq 7. Optical sine-wave patterns consisting of parallel "lines", however, do not duplicate exactly the sine-wave components in a random flux pattern, but rather in a pattern that is random only in the direction x of scanning and uniform in the direction y, perpendicular to the scanning direction. Figure 17.14 illustrates the difference in cross sections through a random grain structure and a synthetic structure representing a random addition of sine-wave test patterns (random phase relation). The differences resulting from scanning a random grain structure or sine-wave test patterns and the suitability of N_e-values, in general, for the purpose of indicating an equivalent aperture area can be determined by a comparison with an equivalent $\overline{N}_0$ based on the *sampling* of a normalized random structure. The various equivalents N_e^*, N_e, and $\overline{N}_0$ can be computed without recourse to response characteristics when the geometrical properties of the aperture are known.

The effective sampling area of an aperture (pictured as a three-dimensional body, the aperture transmittance τ representing height) may be determined by subdividing the aperture into differential columns (fig. 17.15) with a base area $\Delta a = \Delta x \Delta y$ and constant or varying height representing the transmittance $\tau = f(x,y)$. The relative deviation obtained by taking a large number of samples from a random
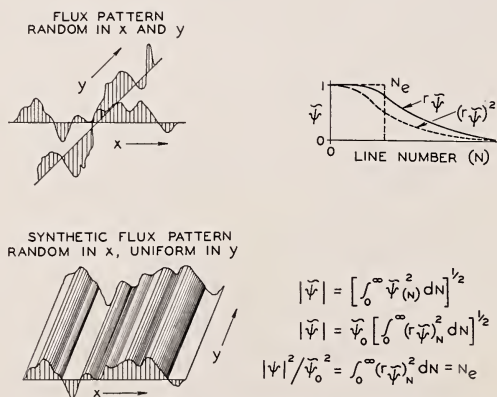

FIGURE 17.14. *Cross sections of random and synthetic grain structures (see text).*

240

grain pattern with one differential column is $\Delta\sigma = (\tau^2 \bar{n}_0 \Delta a)^{1/2}/\tau \bar{n}_0 \Delta a$, where $\bar{n}_0$ equals average number of grains per unit area. For a normalized grain density $\bar{n}_0 = 1$, the above relation becomes $\Delta\sigma_0 = (\tau^2 \Delta a)^{1/2}/\tau \Delta a$. Integration over the aperture yields the normalized relative deviation

$$\sigma_0 = \frac{[\lim \sum \tau^2 \Delta a]^{1/2}}{\lim \sum \tau \Delta a} = \frac{[\int\int f^2(x,y)dxdy]^{1/2}}{\int\int f(x,y)dxdy}. \tag{9}$$

The normalized relative deviation σ_0 has the dimension length^{-1}. The length may be regarded as the geometric mean of the sides of an equivalent rectangular sampling area a_e having constant transmittance $\tau = 1$. According to eq 1 the relative deviation $\sigma_0 = 1/(a_e)^{1/2}$ can also be interpreted as the line number $\bar{N}_0$ of an equivalent square sampling aperture and eq 9 may, hence, be stated in the form

$$\bar{N}_0 = \frac{[\int\int f^2(x,y)dxdy]^{1/2}}{\int\int f(x,y)dxdy}. \tag{10}$$

The measure $\bar{N}_0$ is independent of the aperture orientation for both symmetric and asymmetric apertures and can, hence, be used as a standard for comparison. The equivalent passband N_e^* of an aperture scanning a grain structure random in x and y directions can be computed by subdividing the aperture into incremental sections parallel to the direction of scanning (fig. 17.15). The mean-squared flux obtained is the same as that obtained when the aperture is sampling. The flux $\tilde{\psi}_{0(y)}$ at $N=0$ contributed by each section to $\tilde{\psi}_0$ is represented by the areas $\int \tau dx$ of the sections, and because the flux is random (out of phase) in y, the total flux $\tilde{\psi}_0^2$ is obtained by the sum of the squares $\tilde{\psi}_0^2 = \sum [\int \tau dx]^2$. The measure N_e^* obtained when a random grain pattern is scanned is, therefore,

$$N_e^* = [\tilde{\psi}]^2/\tilde{\psi}_0^2 = \frac{\int\int f^2(x,y)dxdy}{\int dy[\int f(x,y)dx]^2}. \tag{11}$$

The asterisk is used to distinguish N_e^* from the value N_e that will henceforth be used to indicate a sine-wave synthesis.
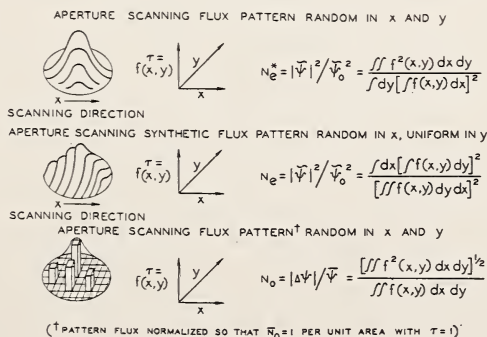

APERTURE SCANNING FLUX PATTERN RANDOM IN x AND y

$f(x,y)$ $\tau = $ SCANNING DIRECTION

$N_e^* = |\tilde{\psi}|^2/\tilde{\psi}_0^2 = \dfrac{\int\int f^2(x,y)\,dx\,dy}{\int dy[\int f(x,y)\,dx]^2}$

APERTURE SCANNING SYNTHETIC FLUX PATTERN RANDOM IN x, UNIFORM IN y

$f(x,y)$ $\tau = $ SCANNING DIRECTION

$N_e = |\tilde{\psi}|^2/\tilde{\psi}_0^2 = \dfrac{\int dx[\int f(x,y)\,dy]^2}{[\int\int f(x,y)\,dy\,dx]^2}$

APERTURE SCANNING FLUX PATTERN† RANDOM IN x AND y

$f(x,y)$ $\tau = $

$N_0 = |\Delta\tilde{\psi}|/\tilde{\psi} = \dfrac{[\int\int f^2(x,y)\,dx\,dy]^{1/2}}{\int\int f(x,y)\,dx\,dy}$

(†PATTERN FLUX NORMALIZED SO THAT $\bar{n}_0 = 1$ PER UNIT AREA WITH $\tau = 1$)

FIGURE 17.15. *Subdivision of apertures for evaluation of the equivalent measures* N_e N_e^* *and* $\bar{N}_0$.

Evaluation of the equivalent passband N_e from a response character-istic obtained by the method of scanning sine-wave test patterns represents the case in which a synthetic structure random in the x direction but uniform (in phase) in the y direction is scanned. The aperture is subdivided into sections parallel to y. The mean-squared flux $[\tilde{\psi}]^2$ is the sum of the squares of the section flux values $\sum[\int \tau \, dy]^2$, and the flux $\tilde{\psi}_0^2$ is the squared sum of the section flux values, furnishing the ratio

$$N_e = [\tilde{\psi}]^2/\tilde{\psi}_0^2 = \frac{\int dx[\int f(x,y)\, dy]^2}{[\int\int f(x,y)\, dy\, dx]^2}. \tag{12}$$

All measures $\overline{N}_0$, N_e^*, and N_e represent dimensionally a length^{-1}, but the formulations appear to have little resemblance to one another. Because the measures N_e^* and N_e depend on the direction of scanning, asymmetric apertures require evaluation of two N_e-values as stated by eq 8. For apertures having circular symmetry, however, the sampling equivalent $\overline{N}_0$ is seen to equal the geometric mean $(N_e^* N_e)^{1/2}$. To evaluate the relative accuracy of the three measures it is of interest to determine how closely the values computed with eq 10, 11, and 12 compare in a number of representative cases. To obtain N_e in relative units N_e/N_δ, the above equations must be multiplied by the ratio of the characteristic lengths δ/u when the length u chosen for computing the measure N_e differs from the length expressing a characteristic diameter of the aperture. In relative units eq 10, 11, and 12 can be written.

$$\overline{N}_0/N_\delta = \frac{\delta}{u} \frac{[\int\int f^2(x,y)\, dx\, dy]^{1/2}}{\int\int f(x,y)\, dx\, dy}, \tag{10}$$

$$N_e^*/N_\delta = \frac{\delta}{u_x} \frac{\int\int f^2(x,y)\, dx\, dy}{\int dy\,[\int f(x,y)\, dx]^2}, \tag{11}$$
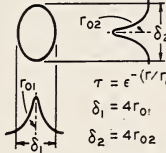
$$N_e/N_\delta = \frac{\delta}{u_x} \frac{\int dx\,[\int f(x,y)\, dy]^2}{\int\int f(x,y)\, dy\, dx]^2}. \tag{12}$$

It must be kept in mind that the length u in eq 10 is the square root of an area and, therefore, independent of the aperture orientation. The length u_x in eq 11 and 12 however, is always the characteristic aperture length in the direction x of scanning.

The measure N_e for a round aperture with $\tau = \epsilon^{-(r/r_0)^2}$, for example, may be computed in terms of a radius length $r_0 = u_x$; the corresponding relative line-number unit N_δ in figure 17.6 represents a length^{-1} measured by the diameter $\delta = 4r_0$. The ratio δ/u_x in this case is, there-fore, four. The relative values in table 17.2 show that the sine-wave equivalent N_e is as good an equivalent as the value N_e^* obtained by the scanning of a random grain structure. Both values are somewhat in error for a round aperture with $\tau = 1$ and for a square scanning diagonally. Practical apertures such as lenses, grain structures, or electron beams have nonuniform transmittances similar to the aperture types 4 to 6 in table 17.1, for which the error is negligible or zero. The definition of N_h^* as the integral of squared response factors given by eq 6 applies also to the measure N_e, which is obtained from a

Table 17.2. Relative passband-equivalents of apertures

APERTURE TYPE	$\tau = f(x,y)$	δ	$\bar{N}_0/N\delta$	$N_e^*/N\delta$	$N_e/N\delta$
1 SQUARE (side s)	$\tau = 1$	s	1	1	1
2 SQUARE (diamond, s)	$\tau = 1$	s	1	$\dfrac{1}{(2/3)\sqrt{2}} = 1.06$	$(2/3)\sqrt{2} = 0.943$
3 ROUND (δ, r_0)	$\tau = 1$	$2r_0$	$\dfrac{2}{\sqrt{\pi}} = 1.13$	$\dfrac{3\pi}{8} = 1.178$	$\dfrac{32}{3\pi^2} = 1.08$
4 ROUND (δ, r_0)	$\tau = 1$, $\cos^2 r$	$2r_0$	1.575	1.56	1.59
5 ROUND (δ, r_0)	$\tau = \epsilon^{-r/r_0}$	$6r_0$	$\dfrac{6}{\epsilon\sqrt{\pi}} = 1.245$	—	1.245
6 ROUND (δ, r_0)	$\tau = \epsilon^{-(r/r_0)^2}$	$4r_0$	$\dfrac{4}{\sqrt{2\pi}} = 1.596$	1.596	1.596
7 RECTANGLE (b, a)	$\tau = 1$	$\bar{N}_0 = \dfrac{1}{\sqrt{ab}}$	$N_e(a) = \dfrac{1}{a}$, $N_e(b) = \dfrac{1}{b}$	$\bar{N}_e = \dfrac{1}{\sqrt{ab}}$ FOR $a = \delta$; $\dfrac{N_e}{N\delta} = \sqrt{\dfrac{a}{b}}$
8 ELLIPSE (r_{02}, δ_2, r_{01}) $\tau = \epsilon^{-(r/r_0)^2}$, $\delta_1 = 4r_{01}$, $\delta_2 = 4r_{02}$	$\bar{N}_0 = \dfrac{1}{\sqrt{r_{01}r_{02}2\pi}}$	$N_e(1) = \dfrac{1}{r_{01}\sqrt{2\pi}}$, $N_e(2) = \dfrac{1}{r_{02}\sqrt{2\pi}}$	$\bar{N}_e = \dfrac{1}{\sqrt{r_{01}r_{02}2\pi}}$, $\dfrac{N_e}{N\delta_1} = 4\sqrt{\dfrac{\delta_1}{\delta_2\,2\pi}}$

sine-wave synthesis; i. e.,

$$N_e = \int_0^\infty (r\tilde{\psi})^2_N \, dN. \tag{13}$$

The results obtained by a numerical integration of the squared aperture response according to eq 13 are illustrated by the curves $N_{e(0\to N)}$ in figures 17.3 to 17.9 that show the growth of the partial integral when the limit is increased from $N=0$ towards $N=\infty$. The accurate agreement of the values obtained by this method is a check on the accuracy of the sine-wave response characteristics as well as the formulation of eq 12. The ϵ^{-r/r_0} aperture is of interest as a mathematical equivalent for grain structures with finite thickness. The line-number scale of this aperture is referred to a diameter $\delta = 6r_0$, which for identical values N_e places the rated resolution N_{cr} of film at the value $N_{cr}/N_\delta = 1.245/0.241 = 5.17$ of the theoretical characteristic. A comparison of figure 17.8 and 17.9 shows an almost perfect agreement of the sine-wave response characteristics. The resolving or sampling aperture of grain structures is, therefore, well represented [6]

---

[6] A finite grain size removes the pointed tip of the aperture transmittance. The effect, however, is negligible because the flux contributed by a transmittance exceeding the value $\tau = 0.65$ ($r = 0.6\,r_0$) is only 2.5 percent of the total flux.

243

by a round aperture with a transmittance $\tau = \epsilon^{-r/r_0}$. The value $\overline{N}_e$ of an asymmetric aperture of width a and height b can be determined accurately when the deformation of the dimensions a or b from circular symmetry does not alter the relative aperture transmittance in the b or a dimension, respectively. In this case the sine-wave measure $N_{e(a)}$ or $N_{e(b)}$ obtained with eq 12 is determined by the dimension of the aperture (a or b), which is oriented parallel to the direction x of scanning, the measure being independent of the aperture scale factor in the y direction. The aperture is thus simply considered first as an aperture with circular symmetry and a diameter $\delta = a$, furnishing the value $N_{e(a)}$, and second as an aperture with the diameter $\delta = b$, furnishing the value $N_{e(b)}$. The geometric mean of these values (eq 8) furnishes the symmetric equivalent $\overline{N}_e$. The corresponding procedure when the sine-wave response of an astigmatic lens is measured, for example, requires orientation of the sine-wave pattern and scanning direction parallel or perpendicular to the direction of astigmatism. The values $N_{e(a)}$ and $N_{e(b)}$ are then determined by numerical integration from the two corresponding sine-wave response characteristics (eq 13). The evaluation of $\overline{N}_e$ is illustrated by two examples in tables 17.3 and 17.4

The numerical evaluation of the measure N_e from a sine-wave response characteristic by means of eq 13 is illustrated by table 17.3 for a 40-mm f:1.6 Ciné Ektar lens measured at f:1.6 and 5° off axis. The value $\bar{a}$ is the mean response factor within the increment ΔN. The equivalent passband N_e is obtained directly in television lines per millimeter; $N_e = 90$ L/mm. Table 17.4 illustrates the evaluation of N_e for grain structures from figure 17.9 in relative units. With reference to the rated resolving power N_{cr} of film, $N_e = 0.241\ N_{cr}$. Hence, for fine-grain positive film (type 5302) with $N_{cr} = 180$ *television* lines per millimeter, $N_e = 43.4$ L/mm.

It may be of interest to the lens designer that the measure N_e can be estimated from the diameter $\delta_{0.5}$ of the physical aperture passing 50 percent of the light flux in the star image. The relation

TABLE 17.3. *Evaluation of N_e for 40-mm Ciné Ektar lens at f/1.6 (5°)*

$N_e = 90$ lines per millimeter

N/mm	$\tau\tilde{\psi}$	$\bar{a}$	$\bar{a}^2\Delta N$	$\Sigma(\bar{a})^2\Delta N$
10	0.98	0.99	9.8	
20	.94	.96	9.2	
30	.90	.92	8.5	
40	.85	.88	7.7	
50	.79	.82	6.75	41.95
60	.74	.765	5.85	
70	.67	.70	4.9	
80	.62	.65	4.22	
90	.57	.59	3.5	
100	.53	.55	3.0	63.42
120	.46	.49	4.8	
140	.42	.44	3.88	
160	.39	.40	3.2	
180	.36	.37	2.76	
200	.33	.345	2.38	80.44
250	.27	.30	4.5	
300	.20	.23	2.65	
350	.14	.17	1.45	
400	.08	.11	.61	
450	.03	.05	.13	89.78

Table 17.4. Evaluation of N_e for grain structures

$$N_e = 0.241 \, N_{cr}$$

N/N_{cr}	$r_{\bar{\psi}}$	$\bar{a}$	$\bar{a}^2 \Delta N$	$\Sigma(\bar{a})^2 \Delta N$
0.05	0.97	0.985	0.049	
.10	.91	.95	.045	0.094
.15	.835	.88	.0385	
.20	.740	.79	.031	.1635
.25	.67	.685	.0235	
.30	.53	.585	.0171	.2041
.35	.44	.50	.0125	
.40	.37	.41	.0084	.225
.45	.30	.335	.0055	
.50	.245	.275	.0038	.2343
.55	.20	.22	.0024	
.60	.16	.18	.0016	.2383
.65	.125	.14	.0010	
.70	.10	.11	.0006	.240
.75	.075	.085	.00035	
.80	.058	.065	.0002	.2405
.85	.04	.045	.0001	
.90	.03	.04	.0001	.2407
.95	.02	.03		
1.0	.018	.02		.241

Table 17.5. Diameter δ and equivalent passband N_e of various aperture types

Aperture type	Relative transmittance	Diameter (δ)	Relation of δ to N_e
Square	$\tau = 1$	s	$s = \qquad l/N_e$
Round	$\tau = 1$	$2r_0$	$\delta_0 = 1.08 \quad l/N_e$
Do	$\tau = \cos^2 r$	$2r_0$	$\delta_{\cos} = 1.59 \quad l/N_e$
Do	$\tau = \epsilon - r/r_0$	$6r_0$	$\delta_\epsilon = 1.245 \, l/N_e$
Do	$\tau = \epsilon - (r/r_0)^2$	$4r_0$	$\delta_{ex} = 1.6 \quad l/N_e$

Table 17.6. Equivalent passband (N_e) and diameter (δ_0) of equivalent round sampling aperture of imaging components

		N_e	δ_0 (microns)	
Theoretical lens	$f/4$	250 L/mm	4.32	See figure 17.7.
	$f/6.3$	159	6.8	
	$f/8$	125	8.65	
	$f/16$	62.5	17.3	
40 mm Ciné Ektar $\{$ at	$f/1.6$ (5°)	90 L/mm	12	See figure 17.10 and
$\quad$ at	$f/2.8$ (5°)	180	6	table 17.2.
50 mm Baltar at	$f/2.8$ (5°)	64	17	
4-inch Super Cinephor	$f/2$	27.3	39.5	
Film	$N_{cr}*$	0.241 N_{cr}	See figure 17.9	*N_{cr} = "rated" resolution at $r_{\bar{\psi}} \cong 2\%$.
Plus X		110 L/mm	40.8	
Fine-grain negative (5203)	220	26.5 L/mm	20.4	
Fine-grain positive (5302)	180	53	25	
16 mm reversal	150	43.4	30	
Square spot	$\tau = 1$	36	See figure 17.3	**N_e = limiting resolution at $r_{\bar{\psi}} = 2\%$.
Round spot	$\tau = 1$	0.50 N_e**		
Round spot	$\tau = \cos^2 r$	.45 N_e	See figure 17.4	
Exponential spot	$\tau = \epsilon^{-r/r_0}$	.38 N_e	See figure 17.5	
Exponential spot	$\tau = \epsilon^{-(r/r_0)^2}$	.244 $N_{cr}*$		
Theoretical lens		.222 N_e	See figure 17.6	
		.20 N_e	See figure 17.7	

$N_e \simeq 0.63/\delta_{0.5}$ is an approximation computed from the flux distribution occurring in practical star images. The factor 0.63 is a compromise value depending on the percentage of flux in the haze surrounding the nucleus of the star image.

The line number for known round apertures is expressed in relative units N/N_δ that refer to the aperture diameter $\delta = l/N_\delta$, where l is the unit of length ($l=1$ mm, or $l=V=$ vertical picture dimension). Relative to the equivalent passband N_e, the diameter of these apertures is expressed by the relations given in table 17.5.

An equivalent aperture or point image of specified characteristics can thus be obtained for a system element by the insertion of its N_e-value into the relations given in table 17.5.

The equivalent passband N_e (television lines) and the equivalent aperture sizes of a number of system elements used in photographic processes are summarized in table 17.6.

Equivalent Passband and Aperture Diameter of Processes Containing a Number of Elements in Cascade

The sine-wave response characteristic of a number of system elements in cascade, including the eye if desired, can be computed accurately by forming the products of the response factors $r_{\tilde{\psi}_1}$ $r_{\tilde{\psi}_2}$, . . . $r_{\tilde{\psi}_n}$ of actual response characteristics at corresponding line numbers. The equivalent passband $N_{e(p)}$ of the process is thus given accurately by the integral

$$N_{e(p)} = \int_0^\infty \left(r_{\tilde{\psi}_1} r_{\tilde{\psi}_2} \ . \ . \ . \ r_{\tilde{\psi}_n} \right)^2 dN \tag{14}$$

Because of the nature of the response characteristics of lenses, films, and television tubes it has been found that the equivalent sampling area of a combination of such "apertures" can be evaluated with usually less than 5-percent error by simply adding the equivalent aperture areas of the components or, as expressed in terms of equivalent aperture diameters:

$$\delta_{(p)} \simeq (\delta_1^2 + \delta_2^2 + \ . \ . \ . \ + \delta_n^2)^{\frac{1}{2}} \tag{15a}$$

also

$$1/N_{e(p)} \simeq (1/N_{e_1}^2 + 1/N_{e_2}^2 + \ . \ . \ . \ + 1/N_{e_n}^2)^{\frac{1}{2}} \tag{15b}$$

Thus, it becomes a simple matter to compute the equivalent passband $N_{e(p)}$ and the aperture diameter of photographic systems by the use of eq 15 in conjunction with table 17.6. Equation 15 is exact for exponential apertures $\tau = \epsilon^{-(r/r_0)^2}$ because the response characteristic figure 17.6 has the form $r\tilde{\psi}_{(N)} = \epsilon\text{-}KN^2$. The response characteristic of a system of two-dimensional apertures tends to approach this form (fig. 17.6), which may therefore be used as an equivalent response characteristic with a line-number scale $N_\delta = N_e/1.6$.

The simplified method will lead to larger errors and should not be used when electrical components of a television system such as amplifiers or filters with sharp cutoff or a rising frequency characteristic

are included. Although equivalent pass-bands (N_e) for such components have a significance, they cannot be treated as normal optical apertures.

Tests for Visual Equivalence

Now that an objective measure for the equivalence of imaging devices or systems has been established, it remains to be shown that the equivalence holds under visual examination. It is obvious that an equivalence indicated by equal measures N_e will be satisfactory when obtained from sine-wave response characteristics that are very similar in shape. A repetition of aperture processes always tends to approach an ϵ^{-kN^2} shape. This observation can be proven when apertures of widely different form such as the combination of lenses and film shown in figure 17.16 are cascaded. In combination with the eye (fig. 17.17), the curve shape becomes even more normalized, indicating that the measure N_e agrees with visual observations. A most critical test for visual equivalence is a comparison of single imaging processes with widely different characteristics but equal measures N_e. This test can be made by comparing pictures made with a complex aperture (lens) such as shown in figure 17.10 with pictures made by an equivalent round aperture of constant transmittance or with a round aperture with $\cos^2$ transmittance. The response characteristic of the latter is shown by broken lines in figure 17.10, differing materially in resolving power from the actual lens. From this point of view the equivalence may appear rather inade-
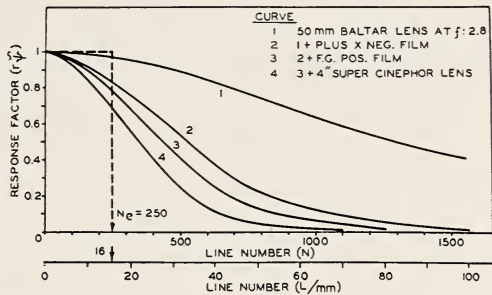

FIGURE 17.16. *Response characteristics in a 35-millimeter motion-picture system.*
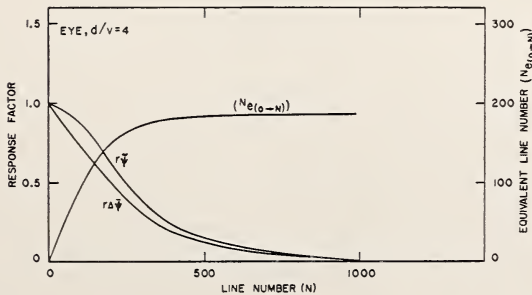

FIGURE 17.17. *Response characteristics of the eye at a viewing distance of four times the vertical picture dimension.*
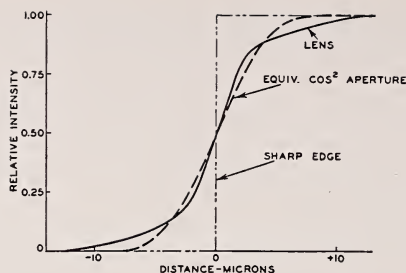
FIGURE 17.18. *Edge transitions of equivalent apertures shown in figure 17.10.*

quate; but when pictures are scanned or produced with these equivalent spots, the transitions at sharp edges do not differ nearly as much as shown by figure 17.18. To permit a visual comparison of images, different aperture shapes and intensity distributions have been synthesized accurately by out-of-focus projections with a precision enlarger. When a pinhole is imaged out of focus it assumes the shape of the enlarger lens diaphragm. In this position, variation of the lens-stop diameter causes a proportional variation in diameter of the out-of-focus point image. The f numbers corresponding to the lens aperture components in figure 17.11 are indicated in the lower portion of the insert as well as the relative exposures (1/speed) for obtaining respective flux values.

Photographs were produced in this manner by multiple exposures, the flux values being measured by a photoelectric device. To be able to show any difference more pronouncedly, the pictures were made with aperture sizes three times larger than that of the lens (respective to the picture frame for which the lens is designed) and thus represent a three times magnification of the image. The N_e-value in the originals was $N_e=210$, whereas that of the lens is $N_e=630$ in a frame 7 mm high. At a normal viewing distance these magnifications are found to have substantially equal sharpness, the equivalent round aperture with constant transmittance appearing slightly sharper. Upon close study the actual lens image exhibits higher resolution but slightly softer edges. Similar observations and tests made by comparing television images and motion pictures of equal N_e ratings have shown that a visual evaluation of sharpness is in good agreement with the objective equivalent N_e.

Conclusions

A study of the factors determining the sharpness in images produced by optical, photographic, and electrical image systems has shown that the relative performance of an imaging device with respect to detail rendition and edge sharpness is accurately specified by its sine-wave response characteristic. Given the characteristics of the components, the performance of any combination forming a multi-element imaging system can be computed accurately. It has thus been possible to solve many problems arising from a combination of optical, photographic, and television processes. Measurement of a

248

considerable number of lenses and television components has shown clearly that the resolving power has little significance as a measure of quality being perhaps the least important point on a response characteristic. The objective measure N_e specifying the equivalent optical passband of the imaging device or system is ideally suited for a universal system of rating image quality, its use permitting a great simplification of many problems. The sine-wave response factors and the measure N_e state performance in relative units. Absolute values of contrast require one additional specification—the large area contrast in the image plane, which can be determined by well known methods. A specification of lens quality must, therefore, include data on its large area contrast as well as a family of response characteristics (or N_e-numbers) as a function of f:number and angle. In view of the inadequacy of the present method of rating image quality by the resolving power of the device, and because of the lack of precise or useful information on detail rendition by lenses intended for pictorial purposes, the author recommends that the above system of rating be studied for adoption as an objective standard. A lens bench for photoelectric measurement of the response characteristic of lenses (including microscopes) has been developed by RCA for the Office of Naval Research under a study contract requiring an accurate evaluation of image quality. It employs moving sine-wave or square-wave test patterns photographed on the sound track of motion-picture film and analyzes the image formed by the lens with a slit aperture, which, referred to the image plane may be made as small as 0.01 micron. Direct readings of the measure N_e can be obtained with a calibrated "noise" film permitting, for example, an experimental and rapid determination of optimum spacings of lens elements. The proposed system of rating image-forming devices has been applied successfully to solve many problems in television systems that require a combination of optical, electrical, and photographic elements.

O. H. Schade, Electro-optical characteristics of television systems, RCA Rev. **9** (1948).

O. H. Schade, Image gradation, graininess and sharpness in television motion-picture systems, J. Soc. Mot. Pict. Televis. Engrs. **56**, No. 2 (Feb. 1951) and **58** No. 3 (March 1952).

O. H. Schade, A new system of measuring and specifying image definition, (unpublished paper given at the 69th Semiannual Convention of the Society of Motion-Picture and Television Engineers in New York City on May 3, 1951).

Discussion

Dr. R. C. Spencer,[1] Air Force Cambridge Research Laboratories, Cambridge, Mass.: The repeated accounts over the past 2 days of the subtle differences between resolving power on close-packed parallel lines and overall picture quality, coupled with Otto Shade's excellent treatment of the subject using communication theory, prompt the following remarks.

[1] Presented in written form following the Symposium as a summary of Dr. Spencer's oral discussion.

Most optical tests involve the high-frequency cut-off of the optical instrument, i. e., its ability to resolve two points (such as stars), two parallel lines, or to follow an abrupt change in illumination (step function). Let us consider now a more general target or input function $G_0(x,y)$ containing lower frequencies that can be represented in the neighborhood of any point by a Taylor's series. Assume that if G_0 were a point (impulse) the optical instrument would spread it out into a system function $F(x,y)$. In general the output function $G_1(x,y)$, when $G_0(x,y)$ is the input function, is the convolution (Faltung) of G_0 and F, expressed as $G_1 = G_0 \times F$. I have shown [2] that for a one-dimensional case

$$G_1(x) = \left(\sum \frac{i^n/y_n}{n!} \, D^n \right) G_0(x) = P \, G_0$$

where P is an operator. Thus, if the nth moments of $F(x)$, defined by $\mu_n = \int x^n F(x) dx$ exist, the output function is a series of products of moments of $F(x)$ and derivatives of $G_0(x)$. Thus, for an even function with μ_0 normalized to unity

$$G_1(x) = G_0(x) - \tfrac{1}{2}\mu_2 D^2 G_0(x) + \ \ldots \ .$$

This form is ideal for the correction or estimation of error of, say 10 percent, introduced into an input function by the limited resolving power of the instrument. Note that for positive moments the output curve G_1 lies always on the concave side of the $G_0(x)$ curve, thus cutting the corners. The case of the second moment for estimating distortion due to television scanning spots was substantiated by Mertz and Gran [3] in 1934. I later generalized the term "instrument" to include circuits, recording galvanometers, etc., in which cases the μ_2 term can be made zero or negative by proper underdamping.

It is noteworthy that the series of corrections is a differential operator P; that a series of instruments can be used in tandem and that the overall operator P is the product $P_1 P_2 P_3 \ldots$; that if each $F(x)$ is shifted to its own center of gravity, thus making each $\mu_1 = 0$, and all μ_0's are normalized to 1, the second moments μ_2 are additive; also the μ_3's.

These results are in agreement with the diffusion processes, and laws of combination of random errors; in particular, the *root-mean-square width* of the apparatus function is the *square root* of the *sum* of the *second moments* of the individual system functions.

The simple theory breaks down in the case of the second moment of the slit diffraction pattern $[(\sin \phi)/\phi]^2$, which is infinite. This was solved by Norbert Wiener who had independently realized the importance of the second moment and by 1941 had derived the best filter for an aperture, such that the second moment of its intensity diffraction pattern, and hence its rms width, would be a minimum. According to Wiener the amplitude transmission function over a slit aperture should be $\cos x$ with the edges of the aperture at the first zeros; also for a circular aperture the amplitude transmission

[2] R. C. Spencer, Phys. Rev. **38**, 618–629 (1931); [A]**46**, 337 (1934); [A] **48**, 4 3 (1935); **52**, 761 (1937), of which eq 4 should read $\mu_n = K^n \int_{-\infty}^{\infty} \psi^{2n} f(\psi) d\psi$; [A] **55**, 239 (1939); [A] **60**, 172 (1941). J. Appl. Phys. **20**, 413–414 (1949).

[3] P. Mertz and F. Gray, Bell Systems Tech. J. 8, 464–515 (1934).

should be $J_0(r)$ with the edge coincident with the first null circle.

In conclusion, let me say that with most of the effort going into the study of the limit of resolution using sharp targets, we should begin to pay more attention to faithfulness of reproduction. The discussion above presents the low-frequency extreme. It does reiterate Shade's contention that the moment of inertia is important, especially for faithfulness of reproduction.

A complete solution, if successful, would have to incorporate considerable portions of the theory of information and in particular Wiener's theory of filtering which not only encompasses the complete spectrum of frequencies but enables the engineer to specify the optimum filter for maximum faithfulness of response of a typical class of input functions in the presence of unwanted noise. During the present sessions we have seen repeated indications of variety of: (a) *input functions* such as star points, lines, sine waves, step functions, square or round areas and typical shapes for letters, trucks, ships, etc.; (b) *noise* from atmospheric haze and temperature variations, photographic haze and particle size and resolution of the human eye. There is still lack of agreement on what is considered to be maximum faithfulness of reproduction for any one set of conditions but there is agreement on the fact that defocusing and other aberrations change the filtering characteristics of the instrument in a measurable manner, these characteristics being resolution and contrast.

Dr. D. S. Grey, Polaroid Corp., Cambridge, Mass.: I would like to comment on the papers by Selwyn and Schade. It has been some 3 years now since Selwyn and Schade have answered for us two questions that have received considerable discussion at this symposium. One question is, what sort of resolution target should we use? The second question is, how, from the energy distribution of the image of a point source, can the lens designer determine what results his lens will achieve under any particular test system with particular targets?

Now, the answer that Selwyn and Schade have given to the first question is that it is entirely immaterial what type is used. It is a question of how you interpret your results in connection with the type of target you use. If you use a line target it does not matter what the contrast in your lens is. If you interpret your results properly you can get the factors that Selwyn and Schade have shown to be pertinent and apply their analysis, and get the test results you would receive with any other type of test target.

The other question that is of interest particularly to the lens designer, is simply that the lens designer may obtain a spectrum of his point image and then he can predict just what his lens will do under a particular specified resolution of test.

I would like to point out one limitation to certain general aspects of the Selwyn-Schade method of analysis, which is the dependence on the object being illuminated by incoherent radiation. If the radiation is coherent, many of the linear properties do not hold, and since I have been working mainly in microscopy since Selwyn and Schade's paper appeared, I am, therefore, excused for not having climbed on the bandwagon sooner.

Now, we can go even further than Selwyn and Schade have indicated directly this morning. I am sure they had it in mind, however. If we analyze the image by their method we can predict what the results will be under test by any particular method. In particular, we can

predict what we are going to get in the photographing of a real object, if we know what the object is.

We can go even further than that by using certain very general theorems that have been developed in communication theory. We need not know exactly what our object is before we specify how our lens is going to perform, that is, to permit us to distinguish certain detail and say, "Is it there?" or "Isn't it there?" We can use the communication theorems and not concern ourselves with precisely what objects we are trying to image, but with the certain general properties of certain general classes of objects that we would like to examine. All the theorems of communication theory can go right over into optics just by a trivial change from one-dimensional notation to two-dimensional notation. We just draw a wiggly line under certain variables indicating that they are vectors and that certain multiplications are dot products of certain vectors.

I have been quite surprised that there has been so much discussion about just what type of target we should use so long after Schade and Selwyn first published their work.

CHAIRMAN B. O'BRIEN, University of Rochester, Rochester, N. Y.: Thank you, Mr. Grey. Is there any further discussion?

DR. G. TORALDO DI FRANCIA, Instituto Nazionale di Ottica, Florence, Italy: I should like to remark upon Dr. Selwyn's paper and Dr. Schade's paper.

Dr. Selwyn measured the increasing contrast that shows up in the curve representing the contrast versus the spacing. He attributed it to the defects of the eye. I think that one can very well explain the phenomenon if one keeps in mind the fact that the resolution by the eye is not a static thing but a kinetic thing. When the spacing between the lines you have to resolve is larger, you need a larger contrast because the eye needs a longer time to scan it.

In this connection, may I remind you of a very fine experiment which I think was known to Helmholtz; looking through a pinhole in a cardboard at the sky and moving the pinhole before one's eye, one may see the blood vessels of the eye very clearly. But as soon as the cardboard is held stationary the blood vessels can no longer be seen. This indicates that the veins in the eye cannot be seen because they are not moving with respect to the retina.

Dr. Schade mentioned that the resolving power depends not on the amplitude of the illumination but on the flux, and I agree with him. We have made many experiments at the Optical Institute in Florence that confirm that the resolving power does not depend on the amplitude but on the energy. We have called it the energetic theory of resolving power. Our experiments are in agreement.

CHAIRMAN O'BRIEN: Thank you, Dr. Toraldo. Now, at the risk of being informal, and with Dr. Gardner's permission, I would like to add a comment to the discussion. What I have to say is that one can resolve quite successfully complex patterns where the total time of illumination is a microsecond or so and no possible scanning of the eye can take place. In order for the visual mechanism to sense the detail, it is necessary that the scanning be in space or in time. It is not essential that it be a scanning in space. Either is adequate. This does not mean that Dr. Selwyn's results are to be disagreed with. As a matter of fact, the agreement of published data is even better than he said. This is, no doubt, modesty on his part. Even

in Cobb's figures you will find perfect agreement with the curve sketched by Dr. Selwyn. Moreover, the fact that the angular separation of lines as it broadens requires a higher, not a lower, contrast of threshold, is in keeping not only with the physiology but the anatomy of the structures.

I shall now call time on myself and ask for more discussion.

DR. R. E. STEPHENS, Optical Instruments Section, National Bureau of Standards, Washington, D. C.: Dr. O'Brien, there is another thing to add to that discussion, I think. The eye seems to be responsive to gradients and also gradients of gradients. There is an experiment I remember in which one produces a split field, one side of which has uniform brightness, the other having a constant gradient where the brightness at the boundary is the same as that of the uniform field. On opposite sides of the boundary there is an infinitesimal difference in brightness nevertheless, a definite line is visible. For sinewaves with the constant contrast, not only the gradient but the gradient of the gradient is smaller for large spacing than for short spacing.

CHAIRMAN O'BRIEN: I know you will agree that this is perhaps a more elegant way of stating the same quantity. Data have been published as a function of the breadth of that. All of these tie neatly together. Is there any further discussion?

MR. J. M. NAISH, Royal Aircraft Establishment, South Farnborough, Hants, England: I would like to make a remark upon the question of standardization of resolving power by measurements. This may not be a very appropriate moment to mention it in the course of the symposium but I feel that sooner or later some remark should be made upon this question. We believe that progress will be accelerated by standardization of resolving-power measurements.

From the symposium so far the importance of resolving power as a means of specifying the performance of the lens in terms of picture quality must be amply illustrated. In spite of the many defects in resolving power as a means of measurement, I think it seems that there remains a thing that will be measured and Dr. Washer's description of the measurement he makes, makes it very clear that it is not a simple matter to carry out a large number of measurements on a given day. The importance of other criteria for picture quality cannot be called in question and, indeed, we have done a certain amount of work on the relation between knife-edge test-object gradients and interferometer quality values with the corresponding resolving-power measurements in the same focal plane. I would think that is an important correlation, especially if the results may be weighted for the variation of wavelength, because of the importance of particularly transverse chromatic aberrations, but nevertheless we feel that we shall all be called back to the question of making these resolving-power measurements in the standard fashion.

I would like to invite attention to the importance of standardizing this procedure. Difficult though it is, it remains very important. We have a British standard. I am not sure whether there is an American standard for measuring resolving power, but I think that eventually we must arrive at an international standard.

The difficulties envolved in accomplishing this are enormous but may I briefly invite attention to the necessity for at least keeping

this goal in mind in the future?

CHAIRMAN O'BRIEN: Thank you. That is a noble goal. I hope some day we do achieve it. It certainly is worth keeping in mind.

Is there any further discussion of these papers?

PROF. F. ZERNIKE, Natuurkundig Laboratorium, Groningen, Netherlands: One question to Mr. Feder. In considering the calculation of intensity in the aberration image, he has said that he is troubled somewhat by the infinities that may occur there and he tries to get away from them by looking at the energy on a certain small surface.

I would like to know how he determines the size of that surface because everything will depend on that. Of course, in reality the finite wavelength of light takes care of this infinity and it will never be infinity at all.

CHAIRMAN O'BRIEN: Is Mr. Feder here?

MR. D. P. FEDER, National Bureau of Standards, Washington, D. C.: I think my point was that with respect to geometrical optics alone, in which the question of energy density has no meaning. Do you agree with this?

PROF. ZERNIKE: Yes.

MR. FEDER: And that one might try something of this nature if the image is poor with respect to the Airy disk? One could, for instance, divide the entrance pupil of the lens into a series of apertures that cover the entire entrance pupil—say you divide it up in squares—and consider where the rays from the corners of the squares go. Assume that the energy then goes into the spot formed by these rays.

PROF. ZERNIKE: How many squares do you need?

MR. FEDER: I don't know the answer to that, but I would like to hear from Dr. Herzberger. Perhaps he will have an answer.

DR. M. HERZBERGER, Eastman Kodak Co., Rochester, N. Y.: Thank you very much. I would like to give an answer to that. In the case of very well corrected optical systems, microscopes or telescopes, it is, of course, necessary to calculate the diffraction image. This is the only way to get the light intensity; but you can get a kind of geometric optical pattern for a photographic lens in the following way. Make a grating in the entrance pupil by dividing the entrance pupil into equal parts. One can do that by a large number of equilateral triangles. Then each ray represents an equal amount of light. If one has a large number of rays and intersects them with the image plane, one gets a distribution that is equivalent, or practically equivalent, to the light distribution in the image. In that manner one can get, by calculating the mumber of points within a given area, the amount of light falling within that area.

PROF. ZERNIKE: Well, my impression is that in order to make, as I would say, an improvised method of taking into account the wavelengths of light—

DR. HERZBERGER: No, the wavelengths of light do not come into this pattern because it would not be very different for different wavelengths—at least as long as the aberrations of the system are large with respect to the wavelengths of light.

PROF. ZERNIKE: I do not agree because Mr. Feder said expressly that the intensity would be infinite.

DR. HERZBERGER: Yes. You see, if one looks at these pictures one gets an immediate impression of the light distribution because the points are so dense. In geometrical optics there would be an

infinite density at the caustic. W. R. Hamilton already has investigated this problem and has tried to compare the intensity at different parts of the caustic. One compares equal areas of the caustic and thus finds a comparative measurement of the intensity at the caustic. The spot diagrams, on the other hand, do not give a correct but an approximate picture of the intensity distribution. They do, however, seem to provide a more graphic method of showing the intensity differences in the image than drawing lines of equal intensity.

CHAIRMAN O'BRIEN: Thank you, Dr. Herzberger. This subject could easily be the topic of another symposium as extensive as this one, and Dr. Gardner has asked that we try to undertake the afternoon program very soon. However, before I spoke, at least Dr. Lucy had waved and perhaps others, so rather than cut off too abruptly, may we have your remarks?

DR. F. A. LUCY, University of California, Los Alamos Scientific Laboratory, Los Alamos, N. Mex.: I have a brief comment. I don't think we have to worry too much about the product of an infinite density over a zero area. The exact limit of the caustic here being really zero in area, we have an indeterminate form that can have a finite product.

DR. G. H. CONANT, JR., Harvard College Observatory, Cambridge, Mass.: May I say in defense of Mr. Feder, that the problem that he was discussing was one of computing with automatic computing machinery and your machine cannot evaluate these indeterminate things. The machine, itself, can only handle a certain finite range of magnitude.

CHAIRMAN O'BRIEN: Mr. Feder, would you like to cap your climax, so to speak?

MR. FEDER: No, I was thinking about Dr. Zernike's question. It seems to me that it is essentially a matter of how closely you examine the image. In other words, if you are going to look at it with a very high powered microscope, then I don't think this method will work. But if you are going to examine it with something that has farily low sensitivity, that is if you take a fairly large area of the image, and if the image is bad with respect to the Rayleigh limit, it is a question of selecting the triangles or squares or what have you in the entrance pupil, so that the energy from one triangle or one square gets into the portion of the area you examine in the image.

CHAIRMAN O'BRIEN: Dr. Hopkins?

DR. R. E. HOPKINS, University of Rochester, Rochester, N. Y.: It seems to me that it is easy to see what the problem is if you think of the interference patterns that I showed. When you have a round, circular area free from fringe, you have to tell from physical optics where the rays within that area are going.

CHAIRMAN O'BRIEN: It is refreshing to see you in such complete agreement.

PROF. A. MARÉCHAL, Institut d'Optique, Paris, France: I have to mention that Prof. Durand in France obtained a better approximation in taking account of the different phases of the vibrations corresponding to the traced rays. It is simpler than complete calculation of diffraction patterns and nevertheless gives a good approximation.

In order to know the contrast of images, or the resolving power of an instrument, information is needed on the diffraction pattern. Those quantities determining the optical quality of the instrument

can be computed by various integrations over the diffraction image. Fortunately, Prof. Duffieux [4] and his co-workers have shown that those quantities can also be computed by convenient integrations performed on the exit pupil. The example of the contrast of images of lines has been studied in a previous communication.

Mr. W. A. Allen,[5] U. S. Naval Ordnance Test Station, Inyokern, China Lake, Calif.: In reference to the paper presented by Mr. Feder, it might be mentioned that many other laboratories have been using automatic computing machinery for the purpose of ray tracing. Specifically, the Los Alamos Scientific Laboratory [6] and the U. S. Naval Ordnance Test Station [7] have investigated and designed many lenses by automatic computing-machine methods. Recent work at the Naval Ordnance Test Station has been based on generalized equations used to trace rays through uncentered-spherical interfaces, prisms, and aspheric surfaces. All surfaces may be considered either refracting or reflecting. Our equations for an aspheric surface appear to converge more rapidly than those used by the National Bureau of Standards.

We have heard much discussion here about the desirability of considering the problem of lens design from the standpoint of physical optics. It should be mentioned that ray tracing in the notation used by the National Bureau of Standards and the Naval Ordnance Test Station provides a calculation of the optical pathlength from object to image. Machine performance is such that this distance is accurate to the wavelength of light. Thus, the data from geometrical ray tracing can be translated by straightforward methods into the language of physical optics. Aspheric surfaces have been designed by the methods of equalizing all optical pathlengths from object to image.[5]

In conclusion, I would like to suggest that there exists a great need for a program involving the systematic evaluation of existing lens prescriptions in terms of resolving power. Such a program would entail the adoption of some reasonable arbitrary standard, such as the criterion suggested by Dr. Hopkins, relating the results of ray tracing to resolving power. Some agency should undertake the analysis and publication of results pertaining to all lenses described in the patent literature. At the present time, any lens design can be investigated exhaustively and rapidly with respect to available glasses, contemplated magnification, and configuration of the image surface, by means of automatic computing machinery. Some resolving-power data, however, brief and inadequate, should be known in advance of such a thorough analysis in order to screen the large group of all known lenses down to a few that might meet prescribed minimum-performance standards. As Dr. Hopkins has remarked, many electronic calculators have been built at the public expense; ray-tracing data obtained by their use should become public property. In my opinion, publication of such information would eliminate much present duplication of effort.

[4] P. M. Duffieux, L'intégrale de Fourier et ses applications à l'optique, (Faculte des Sciences de Besançon, 1947).
[5] Received after the close of the symposium.
[6] W. A. Allen and R. H. Stark, J. Opt. Soc. Am. *41*, 636 (1951).
[7] W. A. Allen and J. R. Snyder, J. Opt. Soc. Am. (in press).

In 1944, Dr. James G. Baker used the IBM Automatic Sequence Controlled Calculator (Mark I) at Harvard for tracing a series of skew rays through an optical system. The equations used and an example of recorded data will be found in one of the National Defense Research publications [1]. This is believed to be the earliest use of automatic programmed computing machines for ray tracing.

In 1945 Grosch [2] reported on the tracing of meridional and skew rays by means of International Business Machines Corporation equipment. The equipment available at that time did not have large storage capacity or provision for sequence programing. Consequently the tracing of a ray through a single surface required the application and removal of 28 plug-boards. Such a procedure really constituted a *tour de force* and could be only used when a large number of rays, say 100 or more, were to be traced through a given optical system. With such a requirement it is possible and desirable to perform a given operation on all the rays before the plug-board is changed for the next operation. In this method the rays are traced *in parallel* instead of in series.

At the fall meeting of the Optical Society of America, October 1949, Grosch [3] reported on the tracing of skew rays by the Selective Sequence Electronic Calculator (SSEC) of the International Business Machines Corporation. This is a programmed machine with sufficient storage capacity to permit ray tracing *in series* through a series of surfaces. In other words, a given ray is traced through the entire optical system before a computation of a second ray is begun. On March 9, 1950, at the time of the meeting of the Optical Society of America in New York there was a demonstration of ray tracing by this machine for the members of the society. At this same meeting Epstein [4] described a method for calculating the third-order aberrations of an optical system by means of some of the less complex IBM equipment.

In March 1950 at a meeting of the Association for Computing Machinery, Donald P. Feder and Benjamin F. Handy, Jr., of the National Bureau of Standards presented a paper titled "Optical Ray Tracing Problems and The Card Programmed Calculator" (CPC). This paper described the use of the CPC in tracing skew rays through a system of spherical surfaces. Also in March of 1950 the Standards Eastern Automatic Computer (SEAC) was first used to trace rays. The use of both of these machines was further discussed by Feder [5] in a talk given a year later at the March 1951 meeting of the Optical Society. At this same meeting Wooters [6] described a method of ray tracing by means of the IBM 604 unit.

In September 1951 Feder [7] gave the detailed formulas that had been used with automatic computing machinery for tracing skew rays through any rotationally symmetric optical system and also formulas for calculating third-order image errors for such systems. At the October 1951 meeting of the Optical Society papers on the use of automatic digital computers for use in geometrical optics were given by Cox and Ledda [8]; Jacobs, May, and Scholnick [9]; Wooters [10]; and Woodson [11].

The foregoing is believed to be a complete summary of the papers appearing in the Journal of the Optical Society of America as presented at its meetings during the past 7 or 8 years. Other applications of programmed computing machinery to optical computations have perhaps been made by other firms and institutions in this country but the information is not available for a complete listing. It may be mentioned, however, that at the National Bureau of Standards, fifth-order aberration equations derived from the equations of Wachendorf [12] are being used for computations on the SEAC machine.

---

[1] Design and development of an automatically focusing 40-inch f/5.0 distortionless telephoto and related lenses for high-altitude aerial reconnaissance, NDRC, Section 16.1, Optical Instruments.

[2] H. R. J. Grosch, Ray tracing with punched-card equipment, Abstract 27, J. Opt. Soc. Am. **35**, 803 (1945).

[3] H. R. J. Grosch, Ray tracing with the selective sequence electronic calculator, Abstract 42, J. Opt. Soc. Am. **39**, 1059 (1949).

[4] L. Ivan Epstein, Calculation of third-order aberrations with the aid of IBM machine, Abstract 9, J. Opt. Soc. Am. **40**, 255 (1950).

[5] Donald P. Feder, Ray tracing with automatic computing machinery, Abstract 45, J. Opt. Soc. Am. **41,** 289 (1951).

[6] Glenn Wooters, Ray tracing with the IBM Electronic Calculating Punch Type 604, Abstract 44, J. Opt. Soc. Am. **41,** 289 (1951).

[7] Donald P. Feder, Optical calculations with automatic computing Machinery, J. Opt. Soc. Am. **41,** 630 (1951).

[8] Arthur Cox and Catherine E. Ledda, IBM automatic equipment in optical design, Abstract 48, J. Opt. Soc. Am. **41,** 874 (1951).

[9] Donald H. Jacobs, Michael May, and Seymour Scholnick, A compact ultra-high Speed digital ray-tracing machine, Abstract 49, J. Opt. Soc. Am. **41,** 874 (1951).

[10] Glenn Wooters, Computing effects of lens variations with electronic calculator, Abstract 50, J. Opt. Soc. Am. **41,** 874 (1951).

[11] Robert A. Woodson, An analysis of H. D. Taylor's f:2 photo lens using a card programmed electronic calculator, Abstract 51, J. Opt. Soc. Am. **41,** 874 (1951).

[12] F. Wachendorf, Bestimmung der Bildfehler 5. Ordnung in zentrierten optischen Systemen, Optik. **5,** 80 (1949).

18. Position of Best Focus of a Lens in the Presence of Spherical Aberration

By R. Kingslake [1]

Introduction

The relation between the image formed by a lens and the amount and kind of spherical aberration present, is of perennial interest to optical designers. Since some residual spherical aberration is usually unavoidable, the designer is concerned with knowing how best to subdivide this residual between marginal overcorrection and zonal undercorrection. The properties that must be considered in this connection are (a) resolving power; (b) depth of focus; (c) the position of the best image, and hence the "Focus shift" (i. e., a longitudinal displacement of the plane of best definition as the lens is stopped down to a smaller aperture); (d) contrast (aberrational haze causes a loss of contrast in the image; this is, of course, entirely distinct from the loss of contrast due to flare light from the lens).

No attempt was made in the present work to measure image contrast, although that would be necessary for a complete study of the problem.

Historical Outline

For very small amounts of primary spherical aberration, in which the sum of the maximum positive and negative departures of the light wave from a perfect sphere does not exceed about a quarter of a wavelength (the Rayleigh limit), it is well known [1][2] that the best-focus plane falls midway between the marginal and paraxial images, the light distribution in the elementary star image being then very similar to that of the ideal Airy disk. From integrations made by Lommel and others, Picht [2] was able to plot a longitudinal section of the light distribution in that simple case.

In 1925, H. G. Conrady [3, 4] studied the properties of a lens having up to 20 times the Rayleigh limit of primary spherical aberration, and found that neither the position of the best focus, the resolving power, nor the depth of focus then agreed with simple theoretical predictions. Her graphs indicate that as the amount of aberration is gradually increased from zero, the plane of best resolution first falls midway between the paraxial and marginal foci, up to the Rayleigh limit, but after that it remains approximately stationary relative to the paraxial focus while the marginal focus moves progressively further away. Thus, neither the simple optical-path-difference theory nor the circle-of-least-confusion theory derived from ray paths represents the facts when large amounts of primary aberration are present.

The more general case of a lens having a zonal residual of undercorrection with some marginal overcorrection has been attacked by

[2] Figures in brackets indicate the literature references on p. 266.

various workers, an outstanding paper being that by Flügge [5], but without any general results being established. In view of the importance of the problem, it seemed worth while to make some direct determinations of the properties of a lens having a variable amount of zonal aberration, in the hope of finding out some definite laws.

Experimental Procedure

A convenient source of a variable amount of spherical aberration was found to be an f/4.5 photographic lens of the Tessar type, of 4 inches focal length, focused by moving the front element. At each position of the front element, the spherical aberration curve was plotted by a modified form of the well-known Hartmann Test, taking five zones of the lens in succession, using "Super-X" panchromatic 16-mm film, D–76 developer, and the band of green light transmitted by a No. 57 filter. The lens was mounted on a nodal-slide lens testing bench, with a 25-mm photographic lens used as a microscope objective to magnify the aerial image 14 times before it was photographed. The film was housed in a Ciné-Special 16-mm camera, without lens, this camera being particularly suitable for the purpose as it is equipped with a direct-view reflex finder and a single-frame crank. The longitudinal position of the viewing microscope was indicated by a dial gauge on the bench, graduated in thousandths of an inch (.025 mm). The test object was either a point source or a resolution chart about 70 feet distant, the chart lines being spaced at intervals of $\sqrt[3]{2}$ from 24 to 300 lines per mm in the aerial image.

For each lens position, a strip of film was obtained, showing on successive 16-mm frames: five pairs of dots from the five lens zones at the inner Hartmann plane just within the best focus; then five pairs of dots from the same zones of the lens at the outer Hartmann plane 0.080 inch (2 mm) beyond the inner Hartmann plane; and finally a succession of photographs of the resolution test object at one-thousandth-inch steps, passing through the best focus region. The exposure time for the Hartmann dot patterns was about 40 seconds, and for the resolution charts an 8-second exposure was satisfactory. In some cases, the lens was subsequently stopped down to a known reduced aperture by an iris diaphragm mounted behind the lens in the same plane as the Hartmann diaphragm, the exposure time being increased to compensate for the reduced lens aperture.

The spacing of the successive pairs of dots in the Hartmann images was measured by projection on a distant screen, the exact projector magnification being immaterial since the position of each zonal focus between the two Hartmann planes was found by simple proportion. Four of the zonal foci were used to determine the four coefficients a, b, c, and d in a power series of the type

$$F = a + b Y^2 + c Y^4 + d Y^6, \tag{1}$$

the coefficient a representing, of course, the position of the paraxial image-point. The fifth zonal focus was then used as a check, the observed and calculated positions never differing by more than 0.001 inch (.025 mm).

The series of resolution-chart images on the film were first examined macroscopically with the unaided eye, to locate the cleanest image where the contrast is a maximum. They were then further studied

260

under a 10X loupe to read the actual resolution achieved at each focus position. The resolution figures were plotted on the same graph as the measured spherical aberration curve for easy comparison.

Experimental Results

The effects of unscrewing the front element of the lens used for these experiments, are shown in figures 18.1 to 18.6. The front element of the lens was mounted on a triple-40 thread, hence it moved 1.905 mm/turn.

In these diagrams, each aberration curve is plotted against semi-aperture, and above it is added a curve of the resolving-power data read directly from the film. The shaded circle on each chart represents the position of the image of greatest contrast and clearness, determined from the film by direct view without a magnifier. It is probable that this would represent the preferred image position for
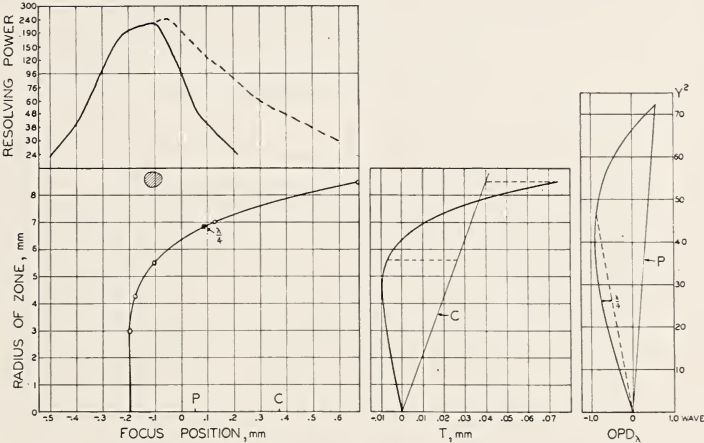
FIGURE 18.1 *Front element screwed in fully.*
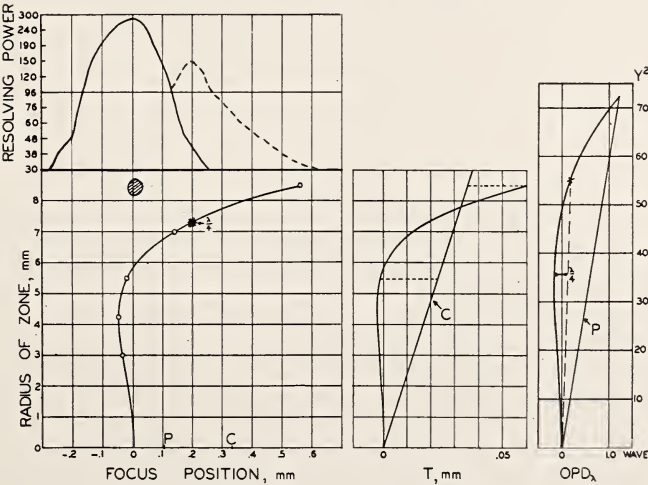
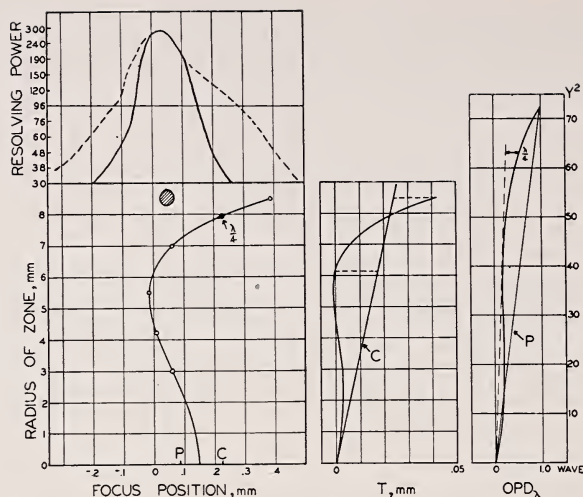FIGURE 18.2. *Front element unscrewed 1/8 turn (.24 mm).*

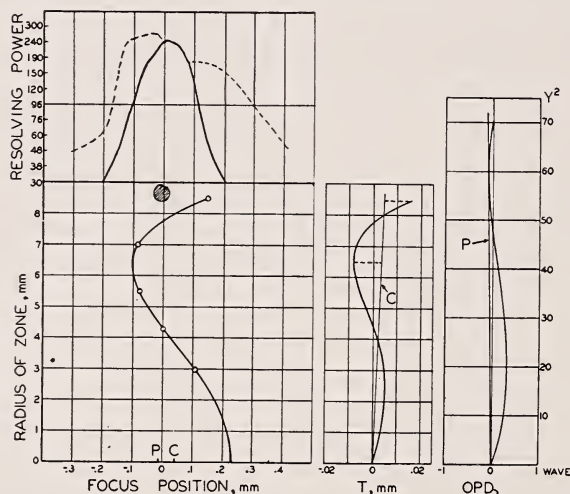FIGURE 18.3. *Front element unscrewed 1/4 turn (.48 mm).*

FIGURE 18.4. *Front element unscrewed 3/8 turn (.71 mm).*

ordinary photography, although the focal position corresponding to the peak of the resolution curve might be preferred if fine detail of high contrast is being studied. An excellent illustration of this phenomenon is given in [6].

The resolving-power graphs consist of two branches. The heavy curve represents genuine resolution in which the separate lines in the target are clearly imaged with sharp edges. However, in many cases after genuine resolution had ceased or almost ceased, the *finer* patterns would become clear again down to a very small pattern, after which resolution would vanish a second time (fig. 18.7). The limits of this extra-fine pseudo resolution are marked on the various graphs by dotted lines. It must be regarded as a form of spurious resolution,
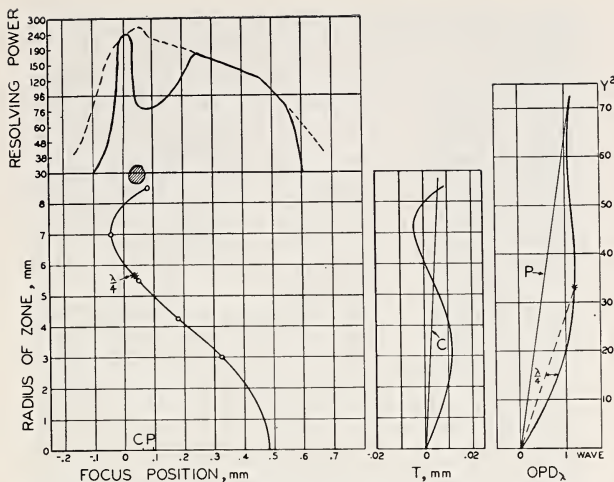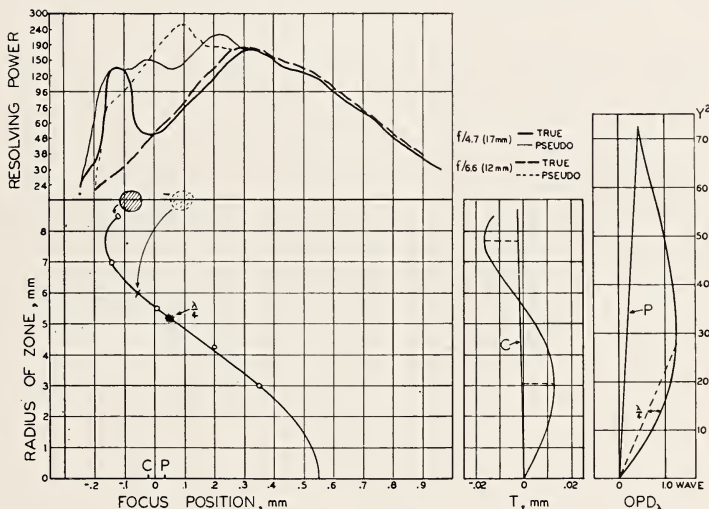
FIGURE 18.5. *Front element unscrewed 1/2 turn (.95 mm).*

FIGURE 18.6. *Front element unscrewed 5/8 turn (1.19 mm).*

even though each pattern shows the correct number of lines (in this case four) and there is no reversal of black and white. This case has been recently recognized by workers at the National Bureau of Standards [7].

The curves for extreme undercorrection, figures 18.5 and 18.6, show two distinct peaks of genuine resolution, one a broad peak representing good resolution but a rather hazy image, and the other an abrupt peak at the point of maximum contrast where the image is very clear and clean. The latter resolution peak is evidently caused by the bend in the aberration curve, for it vanishes when the lens is stopped down (fig. 18.6). It should be noted that the peak resolution in lines per millimeter is not greatly different in any of the charts, showing

FIGURE 18.7. *An enlargement from a typical 16-mm frame showing pseudo resolutions.*

that even 1 mm of spherical aberration is not serious in this respect. The depth of focus increases greatly in the presence of undercorrection, but it is very little effected by overcorrection.

The resolution figures obtained in this test were very high, being substantially the visual resolution of the lens for monochromatic light. It is probable that the pseudo resolution observed in these experiments would not have been found had the aerial image been photographed directly without magnification.

By far the worst practical objection to spherical undercorrection is the shift of focus with aperture. The small shaded circle on each chart represents the position of the "best" image for ordinary photography, and it is clear that if the lens were stopped down to a very small aperture, the plane of "best" focus would move to the paraxial image position. To illustrate this for the case shown in figure 18.6 the series of test-chart photographs was repeated after stopping the lenses down from 17- to 12-mm aperture, and the "best" focus was found to have moved to the position marked by the dotted shaded circle on that chart. However, the peak of the *resolution* curve (ignoring the psuedo resolution) was unchanged in both position and height by stopping the lens down in this way. This bears out the conclusion of Conrady [3, 4] for lenses having pure primary aberration, that the best-resolution plane remains approximately fixed relative to the paraxial focus, even though the marginal focus is moved. This phenomenon of focus shift did not appear when the lens was overcorrected.

Theoretical Attempts to Explain the Observed Phenomena

Two standard modes of approach have been used in the past in attempting to predict the position of the best focus when the form of a spherical aberration curve is known: (a) to determine the position of

the geometrical circle of least confusion, and (b) to locate the center of the sphere that best fits the emerging wave-front.

These procedures are well known, and they have been applied to each of the six cases considered in this investigation. The geometrical approach consists in plotting the curve of transverse aberration, T, against semiaperture, Y, computed by $T = F(Y/l')$. Here F is the—longitudinal focus position relative to an arbitrary zero, and l' is the distance from the Hartmann diaphragm to the image, in this case about 80 mm. The position of the geometrical circle of least confusion marked C on the graphs, is found by drawing such a sloping straight line through the origin that the maximum positive and negative departures of the T curve from the sloping line are equal to each other. Taking any point on this line, the coordinates of which are Y and T, the longitudinal position of the circle of least confusion on the F scale is found at $\delta F = T(l'/Y)$. It is seen from the charts that only in the presence of an undercorrected zone did the geometrical circle of least confusion agree closely with the observed position of best contrast.

The second approach, namely, to plot the shape of the emerging wavefront and try to fit a sphere to it as closely as possible [8, 9], is more elaborate as it requires some additonal calculation. The optical path difference, OPD, is equal to the integral of the angular spherical aberration with respect to the aperture-height Y. Hence,

$$OPD_Y = \frac{1}{l'}\int_0^Y T\, dY = \frac{1}{l'^2}\int_0^Y FY \cdot dY = \frac{1}{l'^2}\int_0^Y (aY + bY^3 + cY^5 + dY^7)\, dY.$$

$$\therefore OPD_Y = \frac{aY^2}{2l'^2} + \frac{bY^4}{4l'^2} + \frac{cY^6}{6l'^2} + \frac{dY^8}{8l'^2}. \tag{2}$$

(The most convenient unit of OPD is the wavelength.) After substituting in this formula the values of a, b, c, and d found for each case, the OPD curve was plotted against Y^2 as ordinate, and a straight line added such that the sum of the greatest plus and minus departures of the curve from the line is a minimum. In most of the cases considered here, such a line merely joined the two ends of the curve. By reading the Y^2 and OPD values of a point on this line, and substituting them in the formula

$$\delta F = 2l'^2 \cdot OPD/Y^2,$$

the position of the center of curvature, P, of the best-fitting sphere could be found. The maximum residual OPD between the wavefront and the straight line was also read off the graph, giving a measure of the aberration in multiples of the Rayleigh limit ($\lambda/4$). By adding a second straight line to the wavefront graph, at a maximum distance $\lambda/4$ from it, it was possible to ascertain by how much the lens must be stopped down to reduce the aberration to the Rayleigh limit; this point is shown on the aberration curves by an asterisk.

The position of physical best focus, P, agreed with the focus for maximum contrast remarkably well, except in the case of extreme overcorrection.

The numerical results of this investigation may be summarizd in tabular form as follows. The "depth of focus", in column 7, represents the longitudinal distance between the planes of 30-line genuine resolution on the two sides of the peak. The lens was, in all cases, used at its full f/4.5 aperture.

Movement of front lens	Marginal spherical aberration	Spherical aberration in multiples of Rayleigh limit	Focus shift, i. e., distance from paraxial focus to:		Peak resolution	Depth of focus (30 to 30 line)
			Maximum-contrast image	Maximum-resolution image		
mm	mm		mm	mm	Lines/mm	mm
0	+0.85	5.0	+0.10	+0.10	230	0.62
0.24	+0.57	3.4	0	−0.02	280	0.51
0.48	+0.24	1.5	−0.10	−0.10	270	0.46
0.71	−0.08	1.2	−0.24	−0.20	240	0.40
0.95	−0.40	2.8	−0.45	{−0.25 {−0.45	{180 {240	0.70
1.19	−0.68	4.2	−0.63	{−0.23 {−0.67	{180 {130	1.17

Conclusions

The best distribution of undercorrection and overcorrection is probably that shown in either figure 18.2 or figure 18.3, in which the marginal overcorrection is several times as great as the zonal undercorrection. In these cases the peak resolution is very high, while the best-contrast image and the peak of resolution fall close to the paraxial focus. Such a lens will show negligible focus shift when it is stopped down. The actual position of the best focus can be calculated by finding the center of the sphere that best fits the calculated wave front, but in making this determination the extreme overcorrected margin of the lens must be ignored as the overcorrected aberration tends to move the calculated point, P, more than it moves the observed best focus. It must always be remembered that any significant zonal undercorrection will lead at once to a focus shift.

The experiments described in this paper refer, of course, to a lens of one particular focal length and aperture, and to one amount of spherical aberration. It may well happen that in a much larger or much smaller lens, different conclusions as to the best distribution of marginal overcorrection and zonal undercorrection would be reached. Also, if accurate measurements on image contrast had been included, the advantages of a large amount of overcorrection might well be offset by the excessive haze accompanying it.

References

[1] A. E. Conrady, Optics of the microscope, Glazebrook's Dictionary of App. Phys. **4**, 217–218 (Macmillan, London, 1923).
[2] J. Picht, Über den Schwingungsvorgang, Ann. Physik [4] **77**, 217–218 (Macmillan, London, 1923).
[3] H. G. Conrady, An experimental study of the effects of varying amounts of primary spherical aberration on the location and quality of optical images, Phot. J. **66**, 9–25 (1926).
[4] H. G. Conrady, Effects of primary spherical aberration on optical images, Proc. Opt. Conv. p. 830–838 (1926).
[5] J. Flügge, Über die Verundeutlichung des Bildes photographischer Systeme durch die Sphärische Aberration, Z. Instrumentenk. **46**, 333–354, 389–415 (1926).
[6] M. Koomen, R. Scolnik, and R. Tousey, Night Myopia, J. Opt. Soc. Am. **41**, 89 (1951).
[7] R. N. Hotchkiss, F. E. Washer, and F. W. Rosberry, spurious resolution of photographic lenses, J. Opt. Soc. Am. **41**, 600–60; (1951).
[8] R. Richter, Zur beugungstheoretischen Untersuchung optischer Systeme, Z. Instrumentenk. **45**, 1–15 (1925).
[9] Wang Ta-Hang, Note on the best focus in the presence of spherical aberration, Proc. Phys. Soc. **53**, 157–169 (1941).

19. Image Evaluation by Edge Gradients

By Arthur Cox [1]

The elementary fact to be faced in connection with any lens is that it yields a light patch of finite size in the focal plane, and that the quality of image it yields depends on the shape and size of this light patch. A practical problem of importance is to bridge the gap between knowing that the image quality is correlated to the form of light patch, and establishing the correlation in a simple and quantitative way.

No correlation can be attempted until an adequate definition or series of definitions of image quality, are established. Setting up a suitable and useful criterion of image quality is a major problem, and a number have been proposed. Two criteria of quality will be considered in this paper, namely the resolving power of a lens as measured with a target having equal lines and spaces, and the ability of a lens to detect a fine dark strip against a bright background. The way in which these may be correlated with the form of light patch, either measured or computed, will be examined. The correlation between other criteria of quality and the form of light patch may be established, in some cases, by an extension of the method proposed.

Suppose that the object field in front of a lens comprises two semi-infinite areas, one bright, the other perfectly dark, with an abrupt brightness transition across the boundary between them. If the light patch were infinitesimally small the intensity distribution in the focal plane would be represented by the line $ABB'C$ in figure 19.1. Because of the finite size of the light patch the actual intensity is represented by the continuous curve of figure 19.1.

Suppose next that the object field comprises a bright strip of finite width on a dark background. In the ideal case the boundaries of the image of this strip would lie along BB' and DD'. With a semi-infinite area terminating on BB' the intensity at a point P would be given by PX in figure 19.1. With a semifinite area terminating on DD' the intensity at P would be given by PX'. The difference between PX and PX' is the intensity at P due to the strip of finite width. The intensity curves in figure 19.1, both the continuous curve corresponding to BB' and the broken curve corresponding to DD', are identical and parallel to one another. Hence PX' is equal to QY where $DB=PQ$, and the intensity at P is the difference between PX and QY. This procedure for obtaining the intensity due to a bright strip may be extended to the case where the object field contains lines and spaces of equal width, the spaces being completely dark. Thus with the edge intensity distribution and the object field pattern shown in figure 19.2 the intensity at P is given by $PX-P_1X_1+P_2X_2 \ldots P_nX_n-P_{-1}X_{-1}+P_{-2}X_{-2}-P_{-3}X_{-3} \ldots +P_{-m}X_{-m}$. The summation of the terms with negative subscripts is over an even
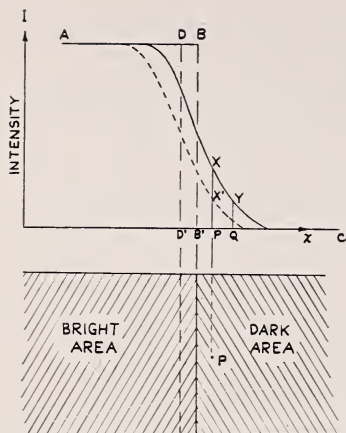
FIGURE 19.1. *Variation of light intensity across a single boundary.*
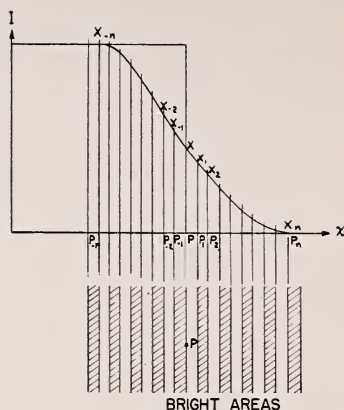

FIGURE 19.2. *Variation of light intensity due to a repeated pattern.*

number of terms ending on the flat part of the intensity curve.

In this way the intensity at any point, due to a repeated pattern of bright lines and dark spaces, may be readily determined. In particular, the intensity at the center of a dark space may be compared with the intensity at the center of a bright line. With a coarse line pattern there is a considerable difference between these intensities. As the pattern becomes finer the difference becomes less, goes through zero and reverses its sign. By setting a suitable ratio of these intensities, usually less than unity, as a criterion of resolution we can readily determine the resolving power of the lens.

The above treatment is appropriate to the case where there is infinite contrast between lines and spaces. When the contrast is finite we can consider a pattern of infinite contrast superimposed on a background intensity, and so derive the resolving power for finite contrast.

Thus if we have as light patch a uniform disk of diameter D, and if X is the relative width of line or space, the intensities I_1 and I_2 at the centers of the lines and spaces are given by table 19.1. Assuming a criterion of resolution as $I_2 \geq .8I_1$ we have $x = .45$. For a finite contrast we can take a base intensity I_0 and $(I_0+I)/I_0 = \alpha$ (α is the contrast ratio). Then we have

$$I'_1 = I_0 + II_1 = I_0\left(1 + \overline{\alpha-1}I_1\right)$$

$$I'_2 = I_0 + II_2 = I_0\left(1 + \overline{\alpha-1}I_2\right)$$

For $\alpha = 2$ and $I'_2 \geq .8I'_2$ we have $I_2 = -.2 + .8I_1$, and $X = .55$. The reduction in contrast has reduced the resolving power by about 20 percent.

The same basic technique used to determine resolving power can be used to determine the visibility of a dark object on a bright background. Thus if the object field consists of a strip of width d separating two semi-infinite areas, as shown in figure 19.3, the intensity dip at the center of the strip is given by the difference between PX

TABLE 19.1.

x	I_1	I_2	x	I_1	I_2
1.0	1.000	0.000	0.35	0.440	0.560
0.9	0.963	.037	.30	.414	.586
.8	.896	.104	.25	.462	.538
.7	.812	.188	.20	.538	.462
.6	.715	.285	.15	.486	.514
.5	.609	.391	.10	.511	.489
.4	.495	.505	.05	.508	.492
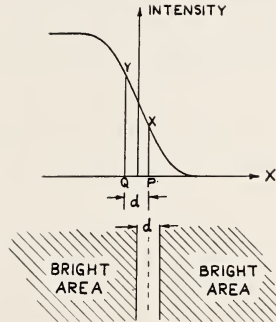
FIGURE 19.3.   *Light intensity in a narrow gap.*

and QY. With a suitable criterion for a perceptible dip in intensity it is a simple matter to determine the least angular width of a dark strip that may be detected against a bright background.

The intensity gradient across an image boundary may be found quite readily from the computed lens data, or determined experimentally. As far as the computed intensity gradient is concerned, it may be evaluated purely on the basis of geometrical optics, or on the basis of diffraction theory and physical optics. These methods of evaluation will be considered in turn.

At any point in the focal plane of the lens the deviations from the principal rays, of rays through other points in the entrance pupil, may be evaluated by standard ray tracing. A considerable economy in ray tracing is effected if an aberration function of suitable form is adopted to represent these deviations. A small number of rays then furnish aberration constants from which the points of intersection with the focal plane of a large number of rays may be readily determined. A convenient way in which to represent the deviation components X and Y is the following:

$$\Delta X = \xi a_0 + a_1 \xi \eta + \xi(a_2 \eta^2 + b_0 \xi^2) + \xi\eta(a_3\eta^2 + b_1\xi^2) + \xi(a_4\eta^4 + c\xi^3) + b_2\xi^2\eta^2$$

$$\Delta - 1 = \eta\alpha_1 + (\alpha_2\eta^2 + \beta_0\xi^2) + \eta(\alpha_3\eta^2 + \beta_1\xi^2) + (\alpha_4\eta^4 + \beta_2\eta^2\xi^2 + \gamma_0\xi^4)$$
$$+ \eta(a_5\eta^4 + \beta_3\eta^3\xi^3 + \eta a_5\eta^4 + \beta_3\eta^2\xi^2 + \gamma_1\xi^4)$$

This is the most general formula taking in terms up to the fifth order of the aperture, and is in effect a determination of the characteristic function relative to a principal ray. The coefficients are functions of

269

the position of the image in the field of the lens. To determine them at any field position 16 rays have to be traced. This is no great problem with automatic computing machinery. Using these coefficients we can determine the points of intersection with the focal plane of rays that come through a large number of pupil points. We have prepared tables for a less general aberration function, and calculated the intersection points of 64 rays in each half of the exit pupil, for a total of 128 intersection points, the calculation being carried out with a Marchant desk calculator. We estimate that with the more general aberration function, using an IBM 604 Electronic Calculating Punch, we can calculate the intersection points of 200 rays from each half of the pupil, giving a total of 400 intersection points, in about 15 minutes. Prepunched decks of cards are used on which the aberration coefficients are subsequently punched.

Given these intersection points we can prepare spot diagrams of the type used by Herzberger and by Linfoot and Hawkins.

For present purposes they can be used to determine edge gradients and then resolving power, using the technique already described. Thus if we have the spot diagram shown in figure 19.4 for the points of intersection of rays from an object point, and if we have a semi-infinite bright area terminating on AA', then the intensity at P is given by the fraction of image points lying to the right of the line $P'P''$ through P parallel to AA'. By varying the position of P, and so of $P'P''$, and counting the number of points to the right of $P'P''$ we get the intensity gradient across AA'. The counting can be carried out quickly on an IBM Sorter.

To summarize the method of evaluating resolving power we have the following steps:

1. Trace sufficient rays to determine the coefficients in an aberration function.

2. Determine a sufficient number of points in the focal plane, using the aberration function.

3. Determine the edge gradient across the boundary between a light and dark area by counting the points in a focal plane to the right of a line parallel to the boundary.

4. Determine the intensity, due to a repeated pattern of lines and spaces, by erecting ordinates to the edge gradient curve with the spacing of the test pattern and summing these alternatively positive and negative.

5. Carry out the process described in step 4 for points corresponding to the centers of lines and spaces. The resolving limit is achieved when the ratio of these intensities surpasses a prescribed value. The procedure described in step 4 must be modified, of course, to take into account finite contrast, in the way previously described.

The above treatment has been described for the case where the ray theory of geometrical optics is applicable and gives a sufficiently close account of image formation. It may readily be extended, in theory, to the case where the image intensity is determined by diffraction theory. Thus if the intensity is $I(x, y)$ at the point (x, y) in the neighborhood of the Gaussian image X of an object point,—then the intensity of I (P), at a point P near the boundary of a bright area and a dark area, as shown in figure 19.4 is given by

$$I = \int I d\sigma,$$

where $d\sigma$ is an element of area, as shown, and the integration is carried out over the shaded area shown in figure 19.4. This integration may be carried out numerically if $I(x, y)$ is known at a sufficient number of points within this area. Evaluating $I(x, y)$ at a large number of points is a tedious process even with automatic computing machinery commercially available. With the IBM 604 Electronic Calculating Punch, for example, we estimate about 4 minutes per point in the image plane, the integration over the pupil being replaced by summation over 400 points in the pupil. We have no estimates of how long the process would take on SEAC or any like computer. The time taken by Maréchal's analogue computer is 8 minutes per point in the image plane, and is restricted to a narrow class of aberrations.

The equation for $I(P)$ given above leads immediately to an experimental method of determining the edge gradient, namely by using the knife-edge test. If a knife edge is traversed from right to left across the image of a fine object point, then the intensity at a point such as P due to the area bounded by AA', is the light transmitted past the knife edge when it has moved from $S'S''$ to $P'P''$.

Under suitable circumstances the same basic technique can be applied to determine the resolving power of a film-lens combination.

Thus if we have an evenly illuminated area of film, sharply bounded by the line AA', the film density across the line AA' will not show an abrupt change, but will be represented by the curve shown in figure 19.5, a. Suppose that we can consider this curve as resulting from the operation of two factors, without specifying in detail the physical basis underlying them. As a result of the first factor, which may be connected for example with scattering in the emulsion, points near the boundary AA' receive an effective exposure represented by the curve shown in figure 19.5, b. Exposure here is measured in terms of incident light energy, not in terms of its logarithm. The second factor
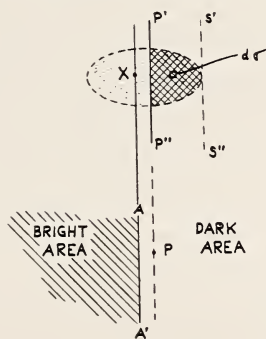

FIGURE 19.4. *Relation of light intensity to image patch.*

FIGURE 19.5. *Combination of factors into film-density curve.*

conditions the response of the emulsion to the effective exposure, and may be of the general form of the H and D curve of the emulsion, showing an initial toe and a region of saturation, as shown in figure 19.5, c.

We have to determine the effective exposure at a point due a given pattern of radiation incident on the emulsion. Again we may proceed by determining first the effective exposure at a point due to a semi-infinite bright area bounded by a line AA', taking into account the intensity curve $I(x)$ due to the finite size of the light patch, and the effective exposure curve $E(x)$ due to the scattering another factor previously defined. The effective exposure $E(y)$ at a point distant y from the boundary AA' is given by

$$E(y) = \int I(x) \cdot \frac{dE(y-x)}{dx} \cdot dx$$

taken over all values of x for which the integrand does not vanish.

The geometrical interpretation of this equation is that if we graph $I(x)$ as ordinate against $E(y-x)$ as abscissa, then $E(y)$ is the area under the graph. To evaluate this geometrically for a range of values of y we proceed as follows. With the same abscissa x draw the graphs of $I(x)$ and $E(-x)$. Erect ordinates at a suitable number of points $x_1, x_2 \ldots x_w$ and read off values of $I(x)$ and $E(-x)$ to serve as ordinates and abscissae in the $I(x)$ versus $E(-x)$ graph. By erecting ordinates at a distance y, positive or negative, from the original ordinates, and by reading $I(x)$ from the original ordinates and $E(-x)$ from the new ordinates we get the appropriate values of $I(x)$ and $E(y-x)$ to insert in a graph such as that shown in figure 19.6. In one extreme position the graph becomes the straight line MN; in intermediate positions it takes the form shown by the dotted lines; and in the other extreme position the graph degenerates to the origin O. The parameter y defines a family of curves, and the area under the curve defined by a particular value of y represents the effective exposure at a point distant y from the boundary of an illuminated area.

Given the effective exposure due to a semi-infinite area we can proceed as before to find the effective exposure due to a repeated pattern of lines and spaces. Using the response curve, which relates the image density to the effective exposure, we can determine the image densities due to the repeated pattern of lines and spaces. By adopting a criterion of resolution in terms of image density we can proceed in this way to determine the resolving power of the lens-film combination.

It is readily seen that the effective exposure curve obtained in this way extends over a distance $x_1 + x_2$, where $I(x)$ and $E(x)$ extend over distances x_1 and x_2. This corresponds to the rough rule given by Katz that the reciprocal of the resolving power of a combination is the sum
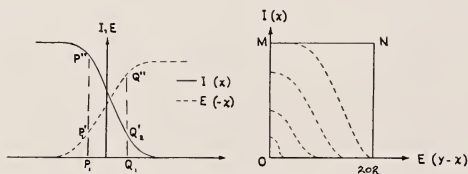

FIGURE 19.6. *Density distribution on film of finite resolving power.*

272

of the reciprocals of the components. The exact relation between component and combination resolving powers depends on the detailed forms of $I(x)$ and $E(x)$. Under the conditions stated it is not a matter of particular difficulty to establish this in any given case when the forms of $I(x)$, $E(x)$ and the response curve are known. Numerical and graphical methods have to be used, since the determination of the area under the $I(x)$, $E(y-x)$ curve does not as a rule lend itself to simple analytical treatment.

The technique described enables us to predict the resolving power of a lens when its aberrational characteristics are known. When the lens performance is described adequately by geometric optics the determination of resolving power may be carried out quickly using an IBM 604 Electronic Calculator. When the lens performance is governed by diffraction theory, and if the diffraction pattern is known, a simple adaptation of the method enables us to predict resolving power in this case also. If we make reasonable assumptions about the photographic emulsion we can obtain the resolving power of the lens-film combination.

20. A Proposed Approach to Image Evaluation

By R. V. Shack [1]

Introduction

The basic element in the formation of any general image is the image of a point. The description of a point image in terms of the aberrations affecting it evaluates the imaging properties of the system. This is the information with which the lens designer is most concerned, but the user obtains from it only a rough qualitative idea of the imaging properties as far as general usage is concerned. That is, he knows that if the aberrations are small, the complex image he obtains is better than if they are large. However, he finds it difficult to establish a good psychophysical correlation between this type of evaluation and the quality of the complex image he generally observes.

For quite some time the user has had his own way of evaluating an image forming system. He determines its ability to resolve the components of a repetitive object, the most common type now in use consisting of alternate black and white lines of equal width. A resolving-power test has several virtues. It is simpler and more rapid to perform than any other existing test, and it gives the user an indication of the size of the finest image detail he can obtain.

However, such a test has its disadvantages also. It is very difficult for a designer to evaluate his design in terms of resolving power. The resolving power obtained in a given test depends not only on the imaging properties of the system, but also on the contrast of the object and the amount of superimposed, nonimage-forming light falling on the image plane. The fact that it is a threshold phenomenon results in a low degree of precision in its determination and variability from observer to observer. It is in a sense an artificial test, even to the user, because he is in general not interested in the ability to separate similar objects but in the ability to image random detail. It has been found recently that the correlation between resolving power and the subjective feeling of sharpness in the image is not as good as it had been assumed to be.

Consequently, considerable work has been done in analyzing the imaging process in order to improve testing procedures. The following approach is based on the image of a straight edge, the basic method of image derivation being suggested by Cox.[2] This was chosen because its properties are two-dimensional, thereby simplifying the analysis, and because its use makes the presentation of the concepts involved much simpler than they otherwise would be. The information obtained from a straight-edge image, or edge gradient, also has the advantage of lying midway between the type of information contained

---

in a point image, of interest to the designer, and the type of information that the user is interested in. The light distribution in the edge gradient can be obtained directly from the light distribution in a point image, and yet the nature of the edge gradient is of direct interest to the user because much of the structure of his general image consists of edges.

The Edge Gradient Concept

In figure 20.1, a, the upper drawing represents the light distribution in a straight-edge object, where the light passing the straight edge is assumed to be superimposed on a general background light. The lower drawing represents the light distribution in the image. This curve is the edge gradient. In figure 20.1, b, the straight edge is assumed to block out light from the general background. The broken curve in the lower drawing might be called an edge gradient attenuation curve, because it represents the amount of light to be subtracted from the background to yield the edge gradient, which is shown as a solid curve.

Single Line Image

1. *The formation of a line image.* A line can be considered to be a combination of two straight edges. In figure 20.2, a, the solid curve in the center drawing represents the edge gradient for the left edge of the line, as if the right edge were indefinitely far away. This light
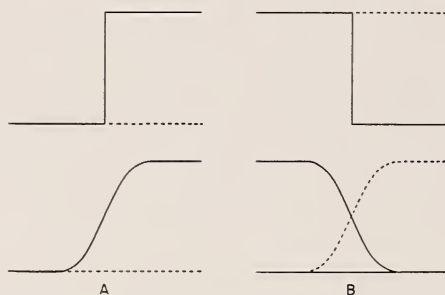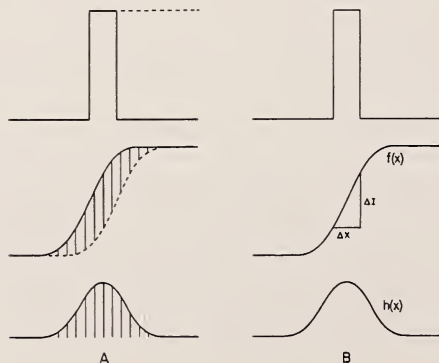

FIGURE 20.1. *The edge gradient.*

FIGURE 20.2. *Derivation of a line image.*

distribution is attenuated according to the dashed edge gradient attenuation curve for the right edge of the line. The remaining light is the light actually existent in the image of the line. The point by point ordinate difference between the two curves yields the image curve for the line, as shown in the bottom drawing.

Now a little reflection will show that the dashed curve is identical in shape to the solid curve. The two curves are separated by the width of the line, so the ordinate differences could just as easily have been obtained by using one curve and obtaining ordinate differences between points on the curve whose abscissa difference is equal to the line width. This is illustrated in figure 20.2, b.

The ordinate difference ΔI is equivalent to the illumination at the corresponding point in the image; Δx is the width of the idealized geometrical image. These two methods for obtaining the image curve are equivalent.

2. *Variation in the line image with line width.* The edge gradient can be represented by a function $f(x)$. The line-image curve can be represented by a function $h(x)$, which will, in general, be different for each Δx, diminishing as Δx approaches zero. This is illustrated by the curves in figure 20.3, a. The marks on the curves indicate the magnitude of the Δx associated with each curve. If each curve is divided point by point by its associated Δx, the set of curves in figure 20.3, b, will be obtained (actually, the curves as shown were divided by $2\Delta x$ for convenience in presentation). The result of this process is that the area under each curve is equal to the area under any other curve in the set. The heavy curve at the bottom corresponds to the distribution of the illumination in the image of an infinitesimally fine line.

3. *The relation between the fine-line image and the edge gradient.* This last curve is the limit as Δx approaches zero of $h(x)/\Delta x$. But, as obtained, $h(x)=f(x+\Delta x)-f(x)$. Therefore, by the definition of a derivative, this limit is the first derivative of $f(x)$. That is, the curve
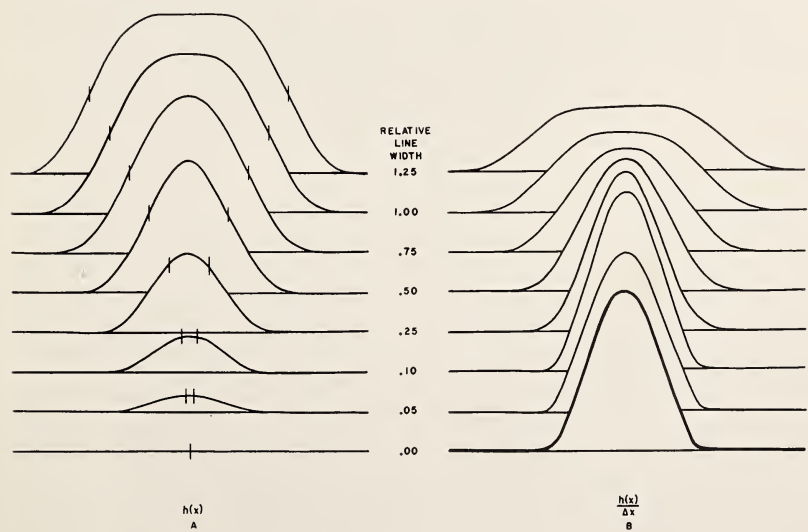

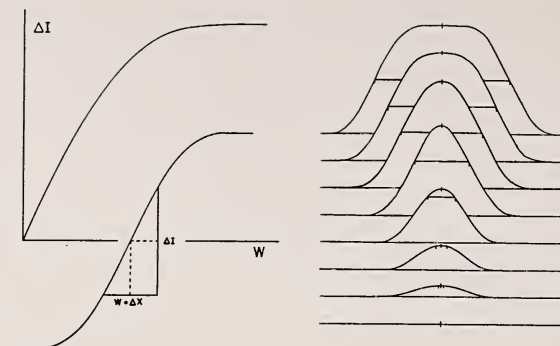FIGURE 20.3. *Line-image variation with changing object line widths.*

Figure 20.4. *Central illumination increment versus ideal image line width.*

of the derivative of the edge gradient has the same shape as the distribution of the illumination in a fine-line image.

4. *The relation between the illumination at the center of a finite line image and the edge gradient.* If the illumination at the centers of the line images in figure 20.3a are plotted against each corresponding Δx, a curve is obtained such as that in figure 20.4. Hereafter, Δx will be called W, the width of the ideal image line.

It will be noticed that, assuming the edge gradient to have point symmetry about its center, each half of the edge gradient may be obtained by plotting $\frac{1}{2}\Delta I$ against $\frac{1}{2}W$. If the edge gradient is not symmetrical, the plot of $\frac{1}{2}\Delta I$ against $\frac{1}{2}W$ will yield a curve of average ordinates between the two halves, assuming that the lower half is rotated 180° about the center to bring it into near coincidence with the upper half. This curve is identical to the curve obtained by plotting the line centers against the widths except that it is half the size.

It is convenient to normalize the ordinate in order to make the result independent of the brightness of the object line. The ordinate then becomes $K = \Delta I/(\Delta I)_{\text{max}}$, which varies from 0 to 1.

5. *Summary—correlations.* This curve of K versus W then is a simple curve, characteristic of the imaging process, and independent of variations in the contrast of the object, or, for that matter, any amount of nonimage-forming light. This is true because K is a ratio of the differences between two pairs of illumination levels and so is independent of change of illumination, which acts on all the levels either proportionally or additively. The curve is first of all a plot of the variation in the relative illumination at the center of a line image versus the ideal width of the line. By section 4, its shape is the same as that of the edge gradient, or very nearly so, depending on the symmetry. And by section 3, it is a plot of the relative amount of flux contained within the limits of W in the fine-line image. These relations are illustrated in figure 20.5. For example, at a point corresponding to a K of .5, the W obtained is the width of the ideal line image where the corresponding actual line image has only half the relative peak illumination. The same W is the width in the fine-line image, which contains 50 percent of the total flux. And if there were some good method of obtaining a numerical value of the edge gradient, it could also be applied to this K versus W curve.
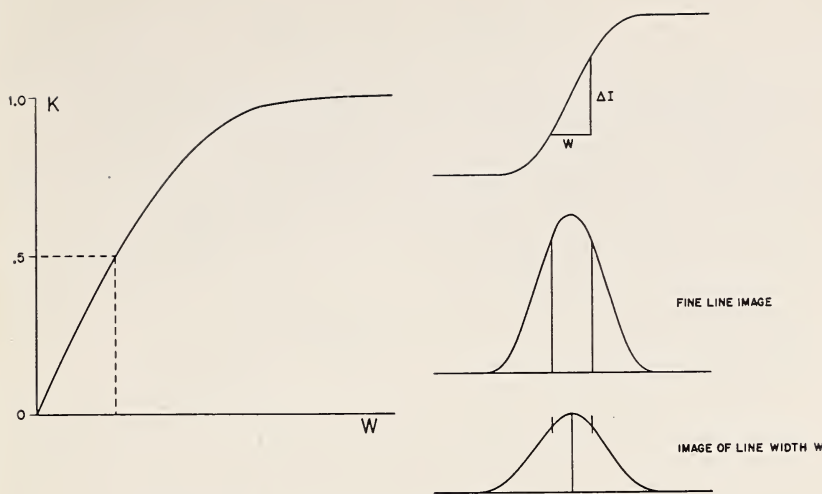
278

FIGURE 20.5. *Line image, edge gradient relationships.*

Therefore, such a curve is suggested as the basis for a new approach to image evaluation. The lens designer can plot the curve for a design if he knows the light distribution in the point image. This means that he can predict the performance in terms that convey considerable information to the user. This information is, of course, the nature of the image of an edge, and the ability to reproduce isolated fine detail.

6. *Image-object contrast relations.* Both the eye and a photographic emulsion can be assumed to be linearly sensitive to the logarithm of the illumination in the image. A convenient definition of the contrast between two illumination levels is the difference between their logarithms, or the logarithm of their ratio, which amounts to the same thing. Photographically, this is equivalent to the density difference between the corresponding parts of the image divided by γ.

The relation between the object brightness ratio R_o and the image illumination ratio R_i is given by the following equation,

$$K = (1+\alpha)\frac{(R_i-1)}{(R_o-1)}, \tag{1}$$

where K is the image evaluation function obtained above; α is a measure of the nonimage-forming light present; $\log R_i$ is the image contrast; and $\log R_o$ is the object contrast.

This equation, derived in appendix 1, applies to dark lines in a bright field as well as bright lines in a dark field.

All the quantities considered in eq 1 are ratios, and so anything affecting illumination proportionally, such as transmission loss in the lens, does not enter into the situation. In other words, such an effect does not alter the contrast.

The quantity $(R-1)$ is what is usually considered the optical contrast. $\log R$ has the advantage of a better psychophysical correlation.

In dealing with photographic images, an emulsion contrast might be defined as $\gamma \log R$. This would be simply the density difference in the emulsion.

Multiple Line Image

1.   *The formation of a multiple-line image.*   In this section, only patterns having equal line and space widths will be considered.

The approach that was used to obtain the image of a single line can be extended to images of a series of parallel lines.   As illustrated in figure 20.6, if the center of the edge gradient is located at the center of one of the lines, then the illumination in the image at that point is given by the sum of the components produced by each line within the effective range of the edge gradient.   An equivalent illumination is obtained at the center of a space.   Curves for both the line response and the space response may be plotted as K_l' and K_s' versus W, where the K's are analogous to the K obtained for a single line, the prime being used to distinguish them.

If the lines and spaces are equal in width, then $K_l'+K_s'=1$, and the curves are symmetrical about $K'=.5$, as shown in figure 20.6.   The point where the curves first meet as W diminishes is the physical limit of resolution, for it is there that both line and space images have the same illumination.   Regions where the curves have crossed each other show line patterns, but the resolution here is spurious.   This occurs because of the repetitive nature of the object.

2.   *Image-object contrast relations.*   The contrast relations for a multiple-line image derived in appendix 2, are a little more complex than they are for the single-line image.   They are as follows.

$$\rho = (1+\beta)\left(\frac{M_i}{M_o}\right). \tag{2}$$

ρ is a contrast reduction function defined in the same way as Schade's amplitude response factor.[3]   It is equal to $2K_l'-1$. β is a measure of the nonimage-forming light present. M_i is an image contrast factor, where $M_i=(R_i-1)/(R_i+1)$, log R_i being the image contrast. M_o is the same type of factor for the object contrast.

That this equation is valid is indicated by figure 20.7.   The three graphs on the left are three plots of the same experimental data.   They were plotted this way so as to indicate the nature of the function involved.   The image contrast is the contrast in the emulsion.   The three graphs on the right, plotted from eq 2, are based on entirely theoretical material, an analytical function being chosen for ρ as a function of line frequency, which is similar to the ρ of the experimental
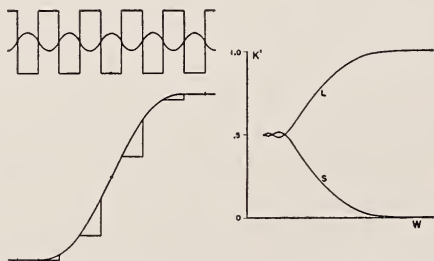

FIGURE 20.6.   *Derivation of a multiple line image.*
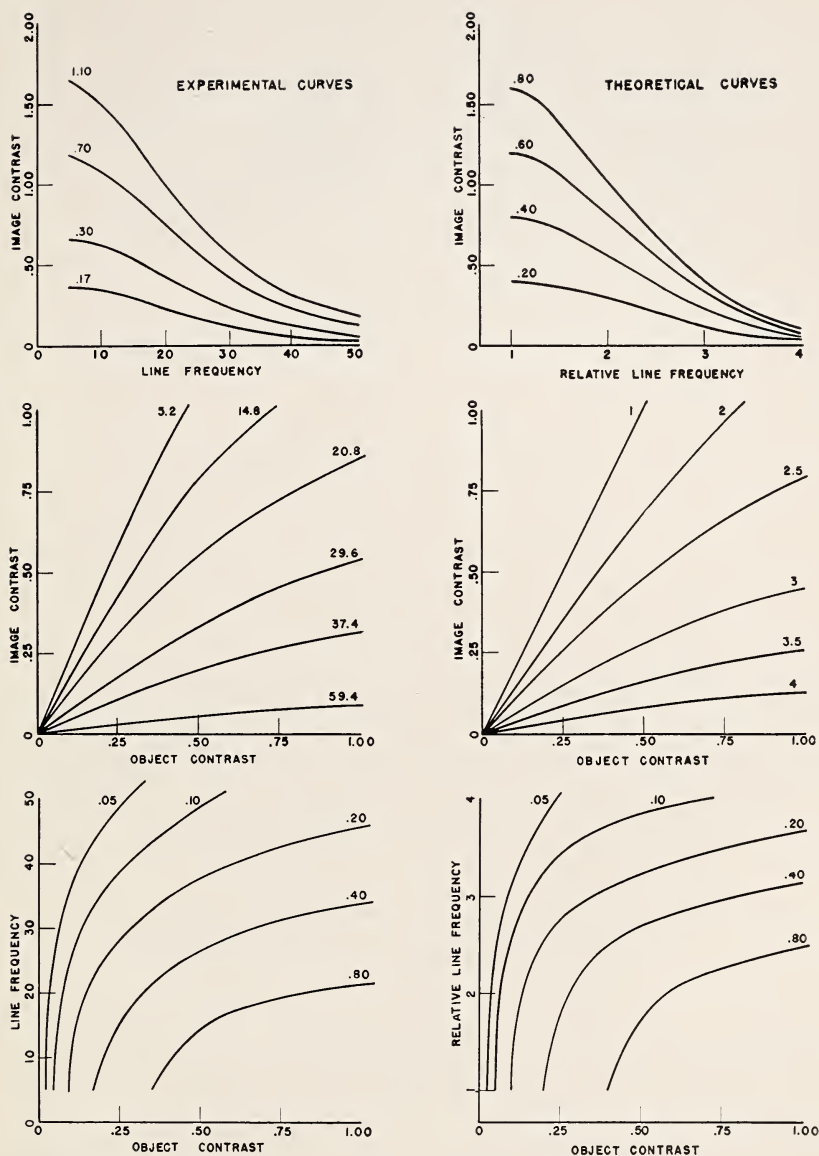
O. H. Schade, RCA Rev. June, 1948.

FIGURE 20.7. *Comparison of experimental and theoretical contrast relationships.*
The experimental data is for a 127-millimeter Ektar lens on V-G emulsion.

curves. Also, for convenience, a γ of 2 was assumed. The nature of the theoretical curves seem to agree quite well with the nature of the experimental curves.

3. *Correlation between the multiple-line image and the one-line image.* Figure 20.8 shows the relationship between the K-curve for the one-line image, plotted as a solid line, and the K'-curves for the
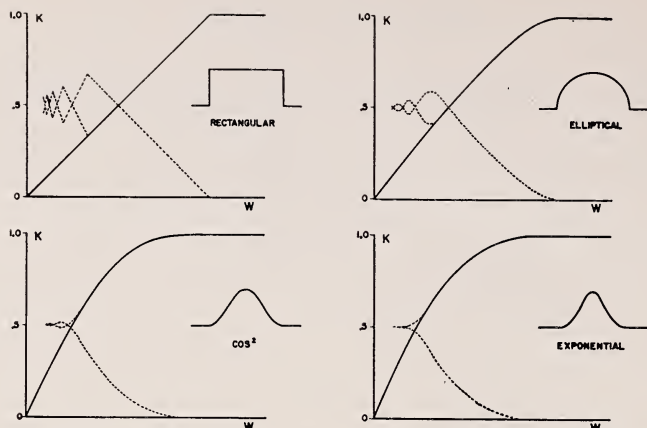
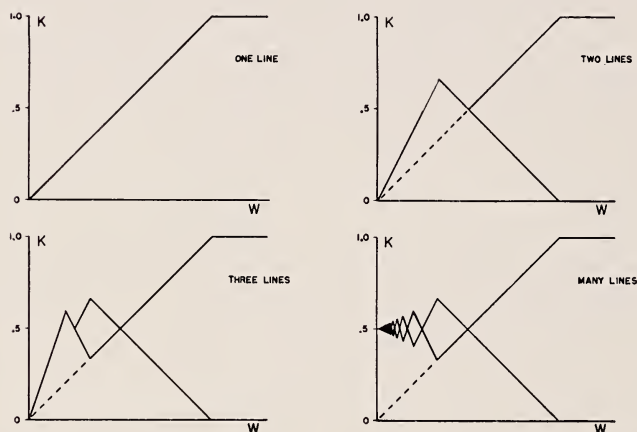FIGURE 20.8. *Relation between single and multiple line images.*

FIGURE 20.9. *Comparison of two- and three-line images with the multiple line image.*

multiple-line image, plotted as dotted lines. K_l' is identical to K except for a small region of W smaller than its value for a point near the limit of resolution.   The inserts are profiles of the fine-line images associated with each graph.

4.   *Two- and three-line images.*   Figure 20.9 indicates the properties of two- and three-line images.   In general they are identical with those of the multiple-line image until the pattern is so fine that the width of the edge gradient (or fine-line image) overlaps the entire pattern of two or three lines.   For smaller patterns the effect is approximately as if the entire pattern were replaced by one broad line.

Figure 20.10 illustrates the nature of the image of a three-line pattern as the line width in the pattern relative to the width of the edge gradient diminishes.   All the curves are drawn with the abscissae expanded so that the apparent widths of the object lines are the same for all the patterns.
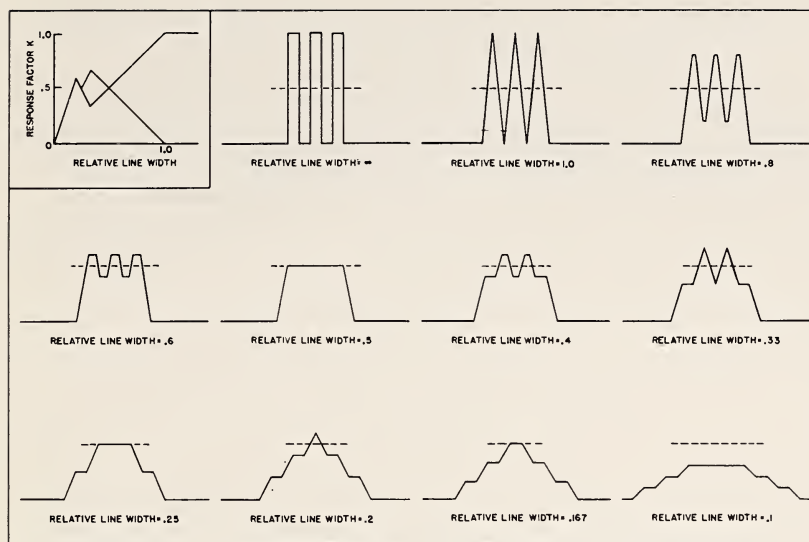
282

FIGURE 20.10. *Variation in appearance of three-line image with changing pattern size.*

Summary

This analysis accomplishes two things. First, it suggests a new approach to image evaluation, which conveys considerable information to the user and yet is close enough to the designer's approach for him to predict the performance of his design in terms of the new approach. Second, it provides an explanation for the phenomena involved in a resolution test, thereby linking the new approach to the older resolution test. Experimental work is going on at present to investigate the validity of the assumptions and conclusions obtained.

Appendix 1. Derivation of Image-Object Contrast Relations for an Isolated Single Line

Let B_1 be the background brightness and B_2 the line brightness in the object. The line may be either brighter or darker than the background. The corresponding illuminations in the image are $I_1 = kB_1$ and $I_2 = kB_2$, assuming no image degradation. Let I_l be the illumination in the center of the actual line image, and i be the illumination in the image plane produced by nonimage-forming light.

Furthermore, let the following symbols be defined as indicated.

$$K \equiv \frac{I_l - I_1}{I_2 - I_1} \tag{1}$$

$$\alpha \equiv \frac{i}{I_1} \tag{2}$$

$$R_o \equiv \frac{B_2}{B_1} = \frac{I_2}{I_1} \tag{3}$$

$$R_i \equiv \frac{I_l + i}{I_1 + i}. \tag{4}$$

Then

$$K = \frac{(I_l + i) - (I_1 + i)}{I_2 - I_1} = \left(\frac{I_1 + i}{I_1}\right)\left(\frac{R_i - 1}{R_o - 1}\right) = (1 + \alpha)\left(\frac{R_i - 1}{R_o - 1}\right). \tag{5}$$

283

Appendix 2. Derivation of image-object contrast relations for a multiple line object

Let B_1 be the space brightness and B_2 the line brightness in the object. The corresponding illuminations in the image are $I_1 = kB_1$ and $I_2 = kB_2$, assuming no image degradation. Let I_l be the illumination in the center of a line in the actual image and I_s be the illumination in the center of a space. Let i be the illumination in the image plane produced by nonimage-forming light.

Furthermore, let the following symbols be defined as indicated.

$$\rho \equiv \frac{I_l - I_s}{I_2 - I_1} \tag{1}$$

$$\beta \equiv \frac{i}{(I_2 + I_1)/2} \tag{2}$$

$$R_o \equiv \frac{B_2}{B_1} = \frac{I_2}{I_1} \tag{3}$$

$$M_o \equiv \frac{R_o - 1}{R_o + 1} = \frac{I_2 - I_1}{I_2 + I_1} \tag{4}$$

$$R_i \equiv \frac{I_l + i}{I_s + i} \tag{5}$$

$$M_i \equiv \frac{R_i - 1}{R_i + 1}. \tag{6}$$

ow

$$M_i = \frac{(I_l + i) - (I_s + i)}{(I_l + i) + (I_s + i)} = \frac{I_l - I_s}{I_l + I_s + 2i}. \tag{7}$$

Because the lines and spaces are equally wide, the average illumination in the image is constant. This means that

$$\frac{(I_l + i) + (I_s + i)}{2} = \frac{(I_2 + i) + (I_1 + i)}{2}$$

or

$$I_l + I_s = I_2 + I_1. \tag{8}$$

Combining (2), (7), and (8) we obtain

$$M_i = \frac{I_l - I_s}{(I_2 + I_1) + \beta(I_2 + I_1)} = \frac{1}{1 + \beta} \left(\frac{I_l + I_s}{I_2 + I_1} \right). \tag{9}$$

Combining (1), (4), and (9) we obtain

$$\rho = (1 + \beta) \left(\frac{M_i}{M_o} \right). \tag{10}$$

Solving this, with (6), we would obtain for R_i

$$R_i = \frac{1 + \left(\dfrac{\rho}{1 + \beta} \right) M_o}{1 - \left(\dfrac{\rho}{1 + \beta} \right) M_o}. \tag{11}$$

Discussion

Dr. M. Herzberger, Eastman Kodak Research Laboratory, Rochester, N. Y.: I greatly enjoyed Dr. Cox' paper, which was very similar to my own ideas. I have some reservations on the last part with respect to graininess, but I have to disappoint him on one point. He said we are in disagreement on how to calculate the spot diagrams. I have tried at least 10 or 15 different ways and found that for one problem one method is preferable and for another problem, another.

Any method of calculating the light distribution is valuable. One advantage of the method that I happened to select for my paper is that this method gives the characteristic function, and thus a complete mathematical analysis of the optical system.

DR. D. S. GREY, Polaroid Corp., Cambridge, Mass.: Dr. Herzberger just touched in his last remark on what I wanted to say. His method is the equivalent of getting response characteristics on the lens whereas what you get is resolving power. If you admit in the beginning that the response characteristic is what you want, you can replace the integration by multiplication. The response-characteristics method is not intended to make work, it is intended to save work.

DR. R. E. HOPKINS, Institute of Optics, University of Rochester, Rochester, N. Y.: I asked Dr. Grey if he was going to start computing response factors for every lens he designed and he made the statement that he was not going to do it for every lens but he was going to do it for some of the lenses. He has already done some, but what he hopes is that this theory is going to give him rules of thumb for designing better lenses. I thought that was a wonderful way of putting it. I think that studies such as Cox is doing on the high-speed computing machine are going to give valuable information, but I think we must remember that not everybody is going to be able to use these machines. Since you are spending taxpayers' money to get this data—and I am a taxpayer—I think it is up to you to publish this data so I will be able to get this rule of thumb for making better lenses.

MR. A. COX, Bell & Howell Co., Chicago, Ill.: The equipment we are using is an IBM Commercial Unit, not the rather glorified equipment that the Bureau of Standards can afford.

CHAIRMAN B. O'BRIEN, University of Rochester, Rochester, N. Y.: Even so, it is your duty to publish it.

DR. G. H. CONANT, JR., Harvard College Observatory, Cambridge, Mass.: I would like to ask Mr. Cox a question, if I may. What digital accuracy do you find it necessary to carry in tracing the rays through the system to get the data that are later evaluated?

MR. COX: As a rule we planned on using machine 604. The work will be described in Chicago. We have used 60 program sometimes and for spherical surfaces, using the 604, we go to five-figure accuracy, which is quite sufficient for most photographic lenses. That takes about 16 seconds per surface. When we need to pick up the accuracy, then we do another run through four different boards, which provides us with an extra two-figure accuracy.

DR. K. V. PESTRECOV, Bausch & Lomb Optical Co., Rochester, N. Y.: My question applies to three or four speakers who talked about analysis of response factors, energy distribution and about predicting the resolution. However, perhaps I am presenting the question too soon. I am interested in why we did not have any tables here computed at a 25-degree resolution on double X. We actually obtained 13. Why is there no experimental confirmation or data? Maybe it is too early to have them. I would like to see them to know where we really stand in actual experimentation with these suggested methods for computing resolving power.

MR. COX: So far the results that have been computed have been on classified military instruments. We have found reasonably good agreement on a rather simple basis. With military instruments we would be working on or near the axis.

Dr. Pestrecov: It would be on or near the axis. I probably can design a lens that will resolve to the Rayleigh limit. I am a little lost as to what to expect from a lens of large field angle. I would like to know if you computed such lenses?

Mr. Cox: We computed an F.2 lens at 15 degrees and got close agreement.

Dr. Pestrecov: Thank you.

Dr. Grey: What else could have gone wrong? Are you asking whether we know our arithmetic?

Dr. Pestrecov: Well, on paper all of the theories sound very plausible. I don't know whether you have experimental confirmation but I would like to know how they work when you have extremely complicated aberrated images—I am still somewhat doubtful that your resolution measurements given in comparison will be correct. I am going to suggest a test at 25 degrees with an extremely complex image pattern and I would like to see somebody predict the resolution, at 15 degrees. Has anybody done anything like that?

Dr. Grey: I haven't had any complaints and I have used this. Even in tests 15 degrees off axis I would expect that the physical optics would enter in an order of magnitude that would be about the discrepancy you would expect. We don't know exactly what the response characteristic of the emulsion is so there are two sources of error. One is the emulsion and the other is the use of geometrical optics rather than physical optics.

Dr. J. G. Baker, Harvard College Observatory and Perkin-Elmer Corp., Norwalk, Conn.: I should like to add a few remarks. I have studied the effect of chromatic aberrations on resolving power because color addition is an important part of the problem. Also, in the presence of aberrations of large magnitude it is necessary to consider the particular resolving-power target rather than just the intensity distribution across an edge. One has to integrate over the pattern as a whole in order to determine what its photographic image will be.

There is another point I want to make. For a number of years I have been using at Harvard a method of testing related to the one Dr. Cox describes. I have called it the "Visual Hartmann Test." Instead of integrating across the image, in this particular form of test I have integrated across the exit pupil. The set-up requires a linear scale across the meridional diameter of the exit pupil and a micrometer controlled knife-edge movement in the vicinity of the mean image. Observing as in the Foucault test, one plots the positions of all bright-dark edges seen in the exit pupil according to scale readings against the micrometer reading. The knife-edge is moved in chosen small increments across the image. The resulting graph represents in effect the slopes of the wave-front. By integrating under the curve, one obtains in effect the shape of the wave-front error in fringes or in any other convenient units, and can correct to any other nearby focal setting. The test is particularly useful for aspheric figuring and I have used it during 1937 to 1939 perhaps several hundred times. I have used it also in the area figuring of a completely unsymmetrical optical system. The test is a useful one in the laboratory.

Excerpt From Letter From T. Smith [1]

Dr. Gardner: It was my sincere wish that Mr. T. Smith might be present at this symposium. He is so well known personally to all of you from across the Atlantic and so well known by his publications to those on this side that his presence would have added greatly to the meeting as a whole, and in particular his contributions to the discussions of the papers would have been most illuminating and helpful to all of us. I know that all here join me in regretting that Mr. Smith found it not possible to be with us. At my request he has prepared a letter for the symposium which I wish to read at this time.

"My hope is that on this occasion when you have present representatives of many countries, you may be able to present questions to them which they will like to consider with their colleagues when they have gone home, and that ultimately agreements may be reached which will be of service to all concerned with optical instruments.

"I will begin with a somewhat minor matter. I suppose that in these days almost all countries are interested in aerial surveys. We all desire the maps that are made to be highly accurate, and one of the factors necessary to this end is excellent correction in the photographic lenses used in the cameras. Perhaps each of the major countries has its own specification for the lenses used for this purpose. It would be interesting to know whether there is sufficient experience for all to agree on the relative importance of the various aberrations that must be kept within close limits. Since the focal lengths are greater than with most lenses intended to cover the same angular field with corresponding relative apertures, aberrations are more appreciable, and it becomes more difficult to keep the large number that may be significant within acceptable limits unless the construction becomes decidedly complex. Freedom from distortion seems to be an important property, and if this is achieved there seems to be no justification for demanding precise centering of the lens in the camera. The prints of the reference marks on the pressure plate enable every photograph to be placed in the same position relative to the lens, and this, rather than reference to the lens axis, is the essential requirement for the construction of accurate maps. It appears undesirable to insist on high accuracy in making an adjustment of some difficulty when this results in no improvement in the final product, and may possibly cause less perfect adjustment in ways that are of real importance, through the difficulty of securing close observance of many adjustments simultaneously.

"Another question that might be considered is whether tele-photo lenses should be used for taking photographs for maps. Unless the centre of the lens aperture coincides with a nodal point or its real image in part of the system the axis of the refracted cone of rays lies in a different direction from that of the incident cone, and this causes a displacement on the plate of the point representing a feature of the ground which lies out of the plane focused on the plate. The use of tele-photo lenses for purposes other than precise mapping seems free from objection. I should be greatly interested in knowing what opinions are held now on these questions—it is some years now since I was in touch with these matters.

"The second subject on which I would like to say a word is the graphical representation of aberrations. It was natural that the representations of aberrations used by von Rohr and others should have been adopted in the early days when new constructions were being evolved, but it is less clear that we should now continue to employ them rather than record our results in a different way. We ought, I think, to consider in the first place what the purpose of these records is. If we only mean to give some idea to the general public of the state of correction of lenses of different designs it may not matter much how these outstanding defects are recorded. But if we are concerned in giving information that will be of value

---

to technical workers the position is different.   For example, if longitudinal central aberration is recorded with the first power of the aperture zone as independent variable, the central part of the curve means nothing, and the outer parts which are of real importance are considerably squashed up; a diagram of this kind no doubt appeals to the makers of lenses that are less good than the best obtainable. But apart from this a diagram of this kind is not particularly useful to technical workers.   The prime consideration for work of high quality is that path differences should not exceed some definite limit—say a quarter of a wavelength of the light taken as a reference standard.   If the longitudinal aberration is plotted with the cosine of the angle made by the ray with the axis as ordinate, the position of focus where the differences of path are a minimum can be found by drawing a straight line parallel to the ordinate axis so as to cut off equal areas between this line and the aberration curve, the axis and the extreme aperture being the other limits of these areas.   In many cases the same use can be made of diagrams with the square of the aperture rather than the first power as the independent variable. Second-order ordinates should also be used to record coma.   Similar criticisms can be made of the usual representations of curvature and astigmatism as well as of distortion.   The central parts of these diagrams are useless, and the important parts made to look of little consequence.   If the square of the angular field, or better still the cosine of the inclination of the principal ray of a pencil to the axis, were taken as variables the diagrams would gain greatly in value.   For distortion the transverse displacement would be plotted as a fraction of the ideal distance from the axis.

"But there is another point to be considered bearing on the graphical representation of aberrations.   In all the early work it seems to have been assumed that it would suffice to record only those aberrations which are of types represented in the aberrations which I call the first order but are very frequently named third-order aberrations.   It happens that these can be exhibited as dependent on only one of the fundamental variables—some on the aperture and others on the field. Among the higher order aberrations are some which necessarily depend on both, and the way in which these aberrations—and they are becoming increasingly important in the development of modern instruments—can be represented graphically so as to be of value to lens designers is not clear.   For this reason I am inclined to think that it is of more importance to give aberrations by means of the values of coefficients than by means of diagrams.   The coefficients that occur in one of the Hamiltonian Characteristic functions—the directional function T has distinct advantages for this purpose—would meet all needs conveniently.   In this connection it is not superfluous to point out that the theory of these calculations has been fully worked out; the real bar to their use in the past for all but the best known aberrations has been the numerical computation; but with the advent of the very powerful and rapid computing engines developed in recent years the labor aspect of this suggestion seems no longer important.   It would, of course, become important to reach agreement on standard forms for the representation of those aberrations to which in the past so many of us have been willing to shut our eyes.

"I have taken the liberty of mentioning two or three optical topics that seem interesting to me, but I realize they may not fit at all into the program that you have in mind for the celebrations you are holding.   I should like to leave them in your hands to deal with exactly as you think fit.   I should in any case be most interested to hear your views on these and indeed on any other optical subjects that are discussed at your meetings."

PROF. F. ZERNIKE, Natuurkundig Laboratorium, Groningen, Netherlands: Yes. Mr. Chairman, and attendants of the symposium. As one of the foreign guests who have so generously been invited to come to this country and attend this wonderful symposium that has I think, and you will all agree, far surpassed our best expectations, I want to express our great gratitude to all who have given their time and efforts to the organization of this symposium as well as to the National Bureau of Standards and the other organizations who have backed it and altogether enabled us to come here.

Of course, we see all these personified in Dr. Gardner and therefore I expressly address Dr. Gardner and tell him how much we have

enjoyed being here. We have not only listened to so many, perhaps too many, scientific papers, all presented in a very congenial way I would say, but we have especially, also, met old friends and have made new ones, and I think this personal note is also of great importance to the progress of the science we represent. Thank you very much.

Dr. I. C. Gardner, National Bureau of Standards, Washington, D. C.: Dr. O'Brien, I would like to say that the success of this symposium has depended in the main on the fact that fortunately a very timely subject was chosen and those who participated have cooperated most wonderfully well.

○