# The frequency response of a defocused optical system

By H. H. Hopkins

*Department of Physics, Imperial College, London, S.W. 7*

(*Communicated by C. R. Burch, F.R.S.—Received* 17 *February* 1955)

The response of a defocused aberration-free optical system to line-frequencies in the object is studied analytically. Curves are given showing the response as a function of line-frequency for a range of values of defect of focus. A comparison is made with the results to be expected from geometrical optics. A tolerance for defect of focus is obtained, which accords well with published experimental results. Both circular and rectangular apertures are considered.

## 1. General formulae

It will be convenient to recall first the notation of an earlier communication (Hopkins 1953). Let a ray from the axial point, $O$, of the object plane of an optical system (figure 1) be inclined at an angle $\alpha$ to the optical axis, and let it intersect the reference sphere in the object space at a height $h$. The reference sphere has its centre at $O$ and is of such radius that it passes through $E$, the axial point of the entrance pupil. This ray emerges from the system and proceeds to the axial image point $O'$, making an angle $\alpha'$ with the optical axis and meeting the reference sphere in the image space at a height $h'$. If any other ray has rectangular co-ordinates $(a, b)$, $(a', b')$ at the object and image reference spheres respectively, the fractional co-ordinates

$$x = \frac{a}{h} = \frac{a'}{h'}, \quad y = \frac{b}{h} = \frac{b'}{h'} \tag{1}$$

denote the point of the pupil traversed by the ray. In a system with a circular aperture, it is convenient to choose the ray $h$ to pass through a point at the edge of the diaphragm, so that $x^2 + y^2 = 1$ denotes the limiting aperture. Points in the object plane are then denoted by the rectangular co-ordinates

$$u = k(n \sin \alpha)\,\xi, \quad v = k(n \sin \alpha)\,\eta, \tag{2}$$

where $k = 2\pi/\lambda$, $n$ is the refractive index, and $(\xi, \eta)$ the geometrical co-ordinate distances of the point. Similar expressions, with appropriate primes, define rectangular co-ordinates for the image plane.

Let that wavefront, associated with a disturbance originating at $O$, which lies in the object reference sphere have unit amplitude and zero phase. The part of the wavefront that lies outside the circle $x^2 + y^2 = 1$ is not transmitted by the optical system; moreover, for each point within the circle, there will be, in general, some loss of light and aberrational effects. The pupil function $f(x, y)$ takes account of both these factors by specifying the disturbance on the reference sphere in the image space. The complex amplitude in the image of a point source, situated at $(0, 0)$ in the object plane, is then given by the Fourier transform

$$F(u', v') = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} f(x, y) \exp \mathrm{i}(u'x + v'y)\, \mathrm{d}x\, \mathrm{d}y, \tag{3}$$

in which the factor $2\pi$ implies a convenient choice of amplitude. Infinite limits may be employed since $f(x, y)$ is zero outside the region $x^2 + y^2 = 1$. The intensity in the diffraction image of the point source is given by

$$G(u', v') = | F(u', v') |^2, \tag{4}$$

but it is, of course, rarely known explicitly.

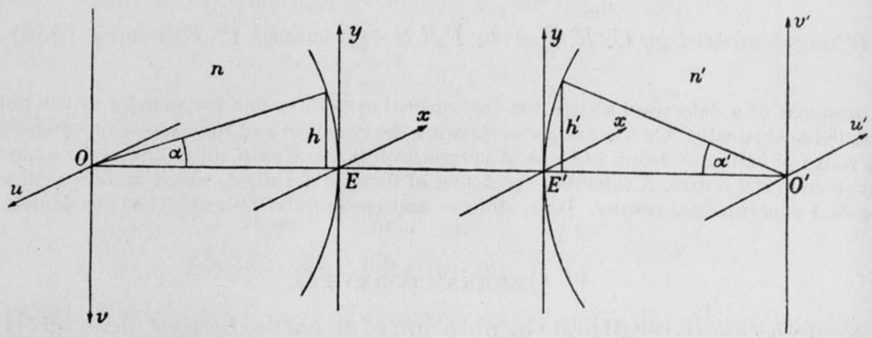


FIGURE 1. $O, O'$, object and image panels; $E, E'$, entrance and exit pupils.

We shall assume that the pupil is isoplanatic, which requires that the diffraction image of a point source situated at the point $(u, v)$ of the object plane be given by $G(u' - u, v' - v)$. In practice, this requires that the aberration of the optical system shall be constant to a small fraction of a wavelength for all points in a region of the geometrical image that is large compared with the extent of the diffraction image of a point source formed by the system. This is clearly satisfied by a system which is free from aberration and merely suffers from defective focusing. This is the case to be considered here.

In incoherent light the distribution of intensity in the image plane is found by integrating the intensity distributions in the diffraction images associated with each point in the object. Thus, if $B(u, v)$ is the intensity at $(u, v)$ in the object plane, the intensity at the point $(u', v')$ in the image is obtained from the formula

$$B'(u', v') = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} B(u, v)\, G(u' - u, v' - v)\, \mathrm{d}u\, \mathrm{d}v, \tag{5}$$

the factor $1/2\pi$ being inserted for convenience. Following Duffieux (1946), we may apply the convolution theorem (Titchmarsh 1937) to (5) and obtain

$$b'(s, t) = g(s, t)\, b(s, t), \tag{6}$$

lower-case letters denoting the inverse Fourier transforms of the functions denoted by the corresponding capital letters. A constant intensity in the object plane has only the zero frequency $b(0, 0)$ in its Fourier spectrum, and this gives rise to a constant intensity in the image. It is for this reason convenient to replace $g(s, t)$ by the normalized transmission factor (response) defined by

$$D(s, t) = \frac{g(s, t)}{g(0, 0)}, \tag{7}$$

the constant $g(0, 0)$ implying merely a change in photometric units. With this normalization, $D(0, 0) = 1$.

If one knows the response function $D(s,t)$ the image distribution corresponding to any object may be calculated, although numerical methods have usually to be employed. If the transparency of the pupil is uniform, and the wavefront aberration of the optical system is denoted by $W(x,y)$, the pupil function has the form

$$\left.\begin{aligned} f(x,y) &= \exp\{ikW(x,y)\} \quad (x^2+y^2 \leqslant 1), \\ &= 0 \qquad\qquad\qquad\quad (x^2+y^2>1), \end{aligned}\right\} \tag{8}$$

because, $W(x,y)$ being the optical distance between the reference sphere and the emergent wavefront, $kW(x,y)$ measures the phase advance at the point $(x,y)$ of the reference sphere relative to that at the origin $(0,0)$. Now the inverse transform $g(s,t)$ is defined by

$$g(s,t) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} |F(u',v')|^2 \exp\{-i(u's+v't)\}\,du'dv'$$

or, applying Parseval's theorem (Titchmarsh 1937) and using * to denote a complex conjugate,

$$g(s,t) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} f(x,y)\,f^*(x-s,y-t)\,dx\,dy,$$

and therefore, by a shift of origin,

$$g(s,t) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} f(x+\tfrac{1}{2}s,y+\tfrac{1}{2}t)\,f^*(x-\tfrac{1}{2}s,y-\tfrac{1}{2}t)\,dx\,dy. \tag{9}$$

It is this integral, with $f(x,y)$ of the form (8), which has to be evaluated to find the response function of an optical system suffering from aberration or a defect of focus.

Let the suffix zero be used to denote all variables and functions in (9) before a rotation of axes. If we now consider two unit circles (figure 2) in the $(x_0,y_0)$ plane centred respectively on the points $(\pm\tfrac{1}{2}s_0, \pm\tfrac{1}{2}t_0)$, the integrand in (9) vanishes outside the region common to these circles, because $f(x_0,y_0) = f^*(x_0,y_0) = 0$ for $x_0^2+y_0^2>1$. If the co-ordinate axes $(x_0,y_0)$ are replaced by axes $(x,y)$, in which the $x$-axis passes through the points $(\pm\tfrac{1}{2}s, \pm\tfrac{1}{2}t)$, the origin being at the mid-point of their join, the formula (9) transforms to

$$g(s,0) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} f(x+\tfrac{1}{2}s,y)f^*(x-\tfrac{1}{2}s,y)\,dx\,dy, \tag{10}$$

where $s = \sqrt{(s_0^2+t_0^2)}$, and $f(x,y)$ derives from $f_0(x_0,y_0)$ by the substitutions

$$\left.\begin{aligned} x &= x_0\cos\psi + y_0\sin\psi, \\ y &= y_0\cos\psi - x_0\sin\psi, \end{aligned}\right\} \tag{11}$$

the angle $\psi$ being $\tan^{-1}(t_0/s_0)$. There is, in consequence, no loss of generality if we confine attention to objects in the form of line structures, that is, objects in which the intensity is constant along the lines $v$ = constant and varies only with $u$. For the above considerations show that the transmission factor (response) of the optical system for the frequency pair $(s_0,t_0)$ is precisely the same as that for the single frequency $(s,0)$, providing the pupil is rotated through an angle $\psi$. Alternatively, we may interpret the result as identifying the response for the frequency pair

$(s_0, t_0)$ with that for the frequency $(s)$ of a line structure whose direction is inclined at an angle $\psi$ to the meridian section of the pupil. This result, which applies to any form of pupil function, not only simplifies the algebra in the analytical evaluation of the response function in any given case, it also implies the important corollary that all the information about the image-forming properties of any system may

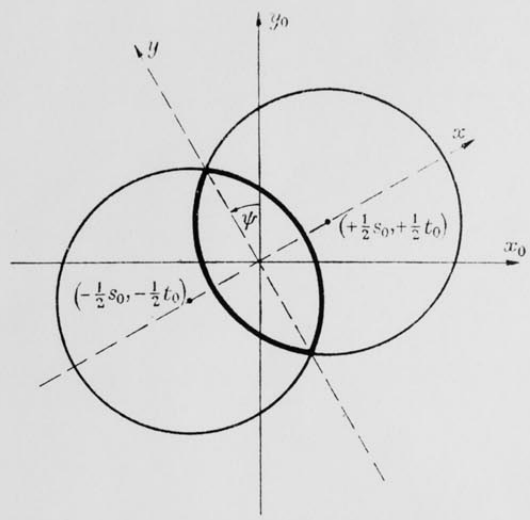


FIGURE 2. Region of integration for the frequency pair $(s, t)$.

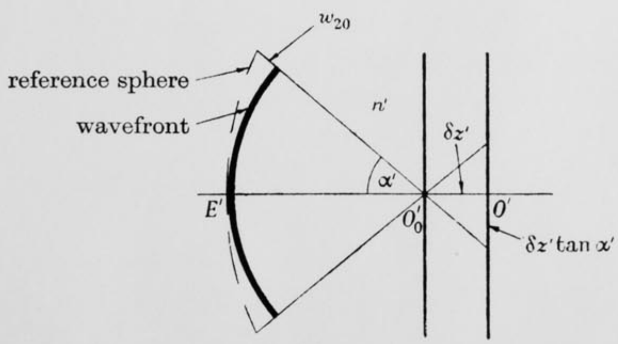


FIGURE 3. The measure of defect of focus.

be obtained in practice using only line structures and scanning slits. The photometric advantages deriving from this fact are obvious. Moreover, the result allows one to think in terms of the more easily visualized case of a unidimensional object.

To study the problem of aberration-free image formation in the presence of a defect of focus, we use a pupil function of the form

$$\left.\begin{aligned} f(x,y) &= \exp\{ikw_{20}(x^2+y^2)\} \quad &(x^2+y^2 \leqslant 1), \\ &= 0 &(x^2+y^2 > 1), \end{aligned}\right\} \tag{12}$$

the coefficient $w_{20}$ measuring the defect of focus by the optical path length of the intercept between the emergent wavefront and a reference sphere centred on the axial point $O'$ of the out-of-focus image plane (figure 3). If $\delta z' = O_0'O'$, where $O_0'$ is in the true focal plane, the defect of focus coefficient is given by

$$w_{20} = \tfrac{1}{2}n' \sin^2 \alpha' \delta z', \tag{13}$$

which is the customary formula for a longitudinal focal shift (Hopkins 1950).

Because of rotational symmetry in the present problem, we may simply treat the case of a line structure with direction parallel to the $v$-axis, and ignore the rotation of the pupil function, because $x_0^2 + y_0^2$ transforms into $x^2 + y^2$. $B(u)$, $B'(u')$ will denote the object and image functions, their inverse transforms being defined by

$$b(s) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{+\infty} B(u) \exp(-\mathrm{i}us)\,\mathrm{d}u, \tag{14}$$

and the corresponding formula with primes. The normalized response is then denoted by

$$D(s) = \frac{g(s,0)}{g(0,0)}, \tag{15}$$

and the distribution of intensity in the image plane is expressed by the formula

$$B'(u') = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{+\infty} D(s)\, b(s) \exp(\mathrm{i}u's)\,\mathrm{d}s, \tag{16}$$

which is correctly normalized, since $B(u) = 1$ gives an image $B'(u') = 1$, the inverse transform in this case being $b(s) = \sqrt{(2\pi)}\,\delta(s)$, where $\delta(s)$ is a delta function.

To relate the frequency variable $s$ to the corresponding number of lines per unit length, we note that the length of one period $u_0$ is such that $u_0 s = 2\pi$. If $R$, $R'$ are the number of lines per unit length measured in the object and image respectively (so that $R = 1/\xi_0$, $R' = 1/\xi_0'$, $\xi_0$, $\xi_0'$ being the pattern sizes in these planes), the definition (2) gives

$$\frac{2\pi}{s} = u_0 = \frac{2\pi}{\lambda}(n \sin \alpha)\frac{1}{R} = \frac{2\pi}{\lambda}(n' \sin \alpha')\frac{1}{R'}.$$

The connexion between $s$ and $R$, $R'$ is thus

$$s = \frac{\lambda}{n \sin \alpha} R = \frac{\lambda}{n' \sin \alpha'} R', \tag{17}$$

and, if $\lambda$ is measured in millimetres (say), $R$ and $R'$ denote the number of lines per millimetre, as measured in the object and image planes respectively.

## 2. The evaluation of the response function $D(s)$

The integrand in (10) is zero outside the region common to the two unit circles centred on the points $(\pm \tfrac{1}{2}s, 0)$. Within this region, which is indicated in figure 4, the integrand has the form (according to (12)),

$$\exp(\mathrm{i}kw_{20})\{[(x+\tfrac{1}{2}s)^2 + y^2] - [(x-\tfrac{1}{2}s)^2 + y^2]\} = \exp(\mathrm{i}ax) \tag{18}$$

for the case of a defect of focus, where $a = 2kw_{20}\,|\,s\,| = 4\pi w_{20}\,|\,s\,|/\lambda$. It is convenient to use the modulus of $s$ here, since $g(-s,0) = g(s,0)$, the region of integration of (18) being symmetrical about the $y$-axis. When $s = 0$ this region is simply the unit circle centred on the origin, so that $g(0, 0) = \pi$. The normalized response function (15) is thus given by the integral

$$D(s) = \frac{1}{\pi}\iint_s \exp(\mathrm{i}ax)\,\mathrm{d}x\,\mathrm{d}y, \tag{19}$$

the symbol $s$ being used to denote the region of integration. Because of the symmetry of this region, the integral reduces to

$$D(s) = \frac{4}{\pi a}\int_0^{\surd\{1-(\frac{1}{2}s)^2\}} \sin a\{\surd(1-y^2) - \tfrac{1}{2}\,|\,s\,|\}\,\mathrm{d}y.$$

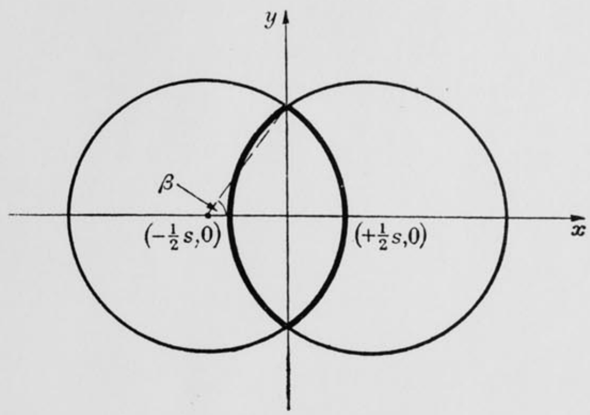


FIGURE 4. Region of integration for the frequency $s$.

If the substitution $y = \sin\theta$ is now made, the integrand may be expanded to give

$$D(s) = \frac{4}{\pi a}\cos\tfrac{1}{2}a\,|\,s\,|\int_0^\beta \sin(a\cos\theta)\cos\theta\,\mathrm{d}\theta - \frac{4}{\pi a}\sin\tfrac{1}{2}a\,|\,s\,|\int_0^\beta \cos(a\cos\theta)\cos\theta\,\mathrm{d}\theta.$$

The geometrical significance of the limit $\beta = \cos^{-1}\frac{1}{2}\,|\,s\,|$ is indicated in the diagram. Using the expansions of $\sin(a\cos\theta)$ and $\cos(a\cos\theta)$ in terms of Bessel functions, the above two integrals are easily evaluated, and give the result

$$\begin{aligned} D(s) = {} & \frac{4}{\pi a}\cos\tfrac{1}{2}a\,|\,s\,|\,\{\beta J_1(a) + \tfrac{1}{2}\sin 2\beta(J_1(a) - J_3(a)) - \tfrac{1}{4}\sin 4\beta(J_3(a) - J_5(a)) + \ldots\} \\ & - \frac{4}{\pi a}\sin\tfrac{1}{2}a\,|\,s\,|\,\{\sin\beta(J_0(a) - J_2(a)) - \tfrac{1}{3}\sin 3\beta(J_2(a) - J_4(a)) \\ & + \tfrac{1}{5}\sin 5\beta(J_4(a) - J_6(a)) - \ldots\} \quad \left(a = \frac{4\pi}{\lambda}w_{20}\,|\,s\,|,\ \beta = \cos^{-1}\tfrac{1}{2}\,|\,s\,|\right), \end{aligned} \tag{20}$$

with an obvious grouping of terms. These series are convergent, and are in a convenient form for numerical evaluation.

If we let $w_{20} \to 0$, the response function is easily shown to tend to the form

$$D_0(s) = \frac{1}{\pi}(2\beta - \sin 2\beta), \tag{21}$$

which is the known result for the in-focus image. It is of interest to note that

$$\left|\frac{\partial}{\partial s} D(s)\right|_{s=0} = \left|\frac{\partial}{\partial s} D_0(s)\right|_{s=0} = -\frac{2}{\pi} \tag{22}$$

and

$$\left|\frac{\partial}{\partial s} D(s)\right|_{s=2} = \left|\frac{\partial}{\partial s} D_0(s)\right|_{s=2} = 0, \tag{23}$$

which are particular cases of a more general result. These results are most easily obtained by considering the series expansions of (20) about the points $s = 0$, $s = 2$. Moreover, $D(0) = 1$ and $D(s) = 0$ for $s > 2$, for in the latter case the region of integration vanishes. From (20) it may be seen that $D(s)$ is an even function of $w_{20}$. Hence the response of the lens is identical in planes at equal distances on the two sides of the focus, and the images of any object formed by an aberration-free lens are therefore symmetrical about the true focal plane.

The intensity $B(u)$ at any point in the object plane is necessarily real. In consequence the transform (14) satisfies the condition $b(-s) = b^*(s)$. Moreover, an immediate consequence of the definitions (10) and (15) is that $D(-s) = D^*(s)$. To find the frequency component $b'(s)$ of the image intensity $B'(u')$, we multiply the frequency component $b(s)$ of the object by its transmission factor $D(s)$, so that $b'(s) = D(s).b(s)$. It follows that $b'(-s) = b^*(s)$, and hence $B'(u')$ is always real—as is, of course, necessary. If we write $b(s)$ in terms of a modulus and argument, $b(s) = \beta(s) \exp\{i\phi(s)\}$, and similarly

$$D(s) = T(s) \exp\{i\theta(s)\}, \tag{24}$$

the object intensity function may be written

$$B(u) = \sqrt{\frac{2}{\pi}} \int_0^\infty \beta(s) \cos\{us + \phi(s)\} \, ds,$$

and the image is then described by

$$B'(u') = \sqrt{\frac{2}{\pi}} \int_0^\infty T(s)\, \beta(s) \cos\{u's + \phi(s) + \theta(s)\} \, ds.$$

Comparison of these expressions shows that the component of frequency $s$ has its amplitude modulated by a factor $T(s)$, and is phase-shifted relative to the geometrical image by an amount $u' = \theta(s)/s$. For a symmetrical pupil function $f(x, y)$, $D(s)$ is wholly real. There is then no phase-shift, but negative values of $D(s)$ denote a reversal of phase, which can, of course, be regarded as a phase shift equal to one-half the pattern size corresponding to the frequency $s$.

The response curves have been calculated for defects of focus corresponding to $w_{20} = (n/\pi)\lambda$, $n$ having values between 0 and 60. The largest value, $n = 60$, denotes a defect of focus $w_{20} = \pm 19{\cdot}1\lambda$. For a system of numerical aperture $\sin\alpha' = 0{\cdot}10$, with $n' = 1$, this corresponds to $\delta z' = \pm 1{\cdot}9$ mm, if $\lambda = 0{\cdot}5\mu$.

A marked feature of these curves, shown in figure 5, is the very rapid deterioration of the response of the lens for higher frequencies with the introduction of small amounts of defect of focus in excess of $\lambda/\pi$. This is well illustrated in figure 6, in which the limiting frequency for which $D(s) > 0$ is shown as a function of the defect of focus. Beyond the point $w_{20} = 3\lambda$ the effect of increasing defect of focus on the bandwidth transmitted by the lens is, by comparison, very slow. For this value, $w_{20} = 3\lambda$, the limiting value of $s$ is equal to 0·10. If $\sin\alpha' = 0{\cdot}10$, $\lambda = 0{\cdot}5\mu$, $n' = 1{\cdot}0$

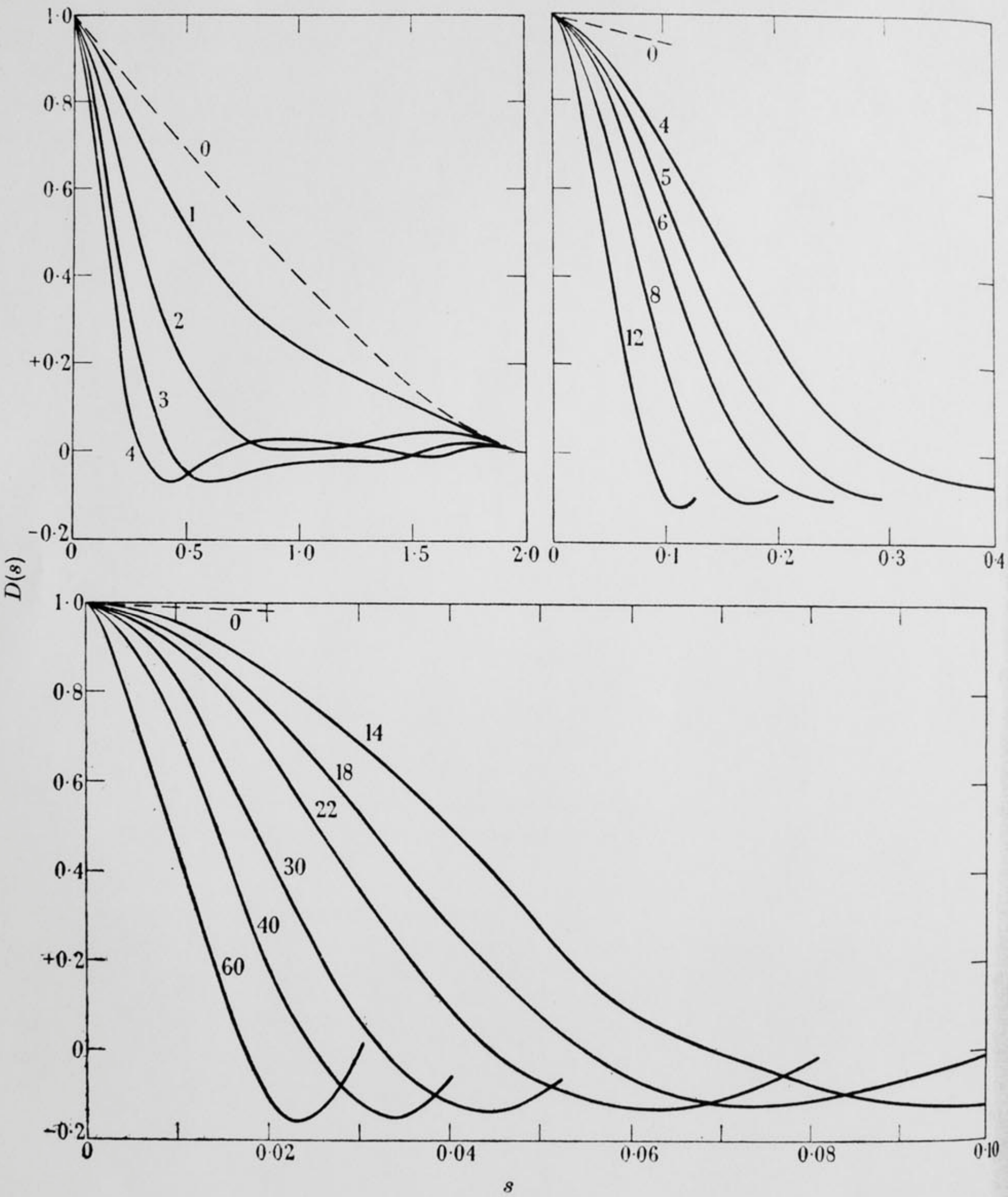


FIGURE 5. Out-of-focus response curves. The curve numbers relate to a defect of focus $w_{20} = n\lambda/\pi$.

(that is an $F/5$ objective) the number of lines per millimetre $R'$ corresponding to any line frequency $s$ is $R' = 200s$, as may be seen from (17). Thus $s = 0{\cdot}10$ corresponds in this case to 20 lines/mm in the image.

Beyond the transmitted bandwidth, for larger values of $w_{20}$, there are side bands showing reversal of contrast, for which the maximum contrast transmission factor $T(s)$ increases numerically from 7 %, for $w_{20} = (4/\pi)\,\lambda$, up to 16 % for $w_{20} = (60/\pi)\,\lambda$. This accounts for the known spurious resolution which is obtained with a well-corrected lens used with a large defect of focus. There are successive sidebands beyond the first, showing alternately correct and reversed contrast. It is shown below that, according to geometrical optics, the maximum numerical values of $T(s)$ are equal to the turning values of $2J_1(a)/a$, $a$ being the quantity defined in (20).

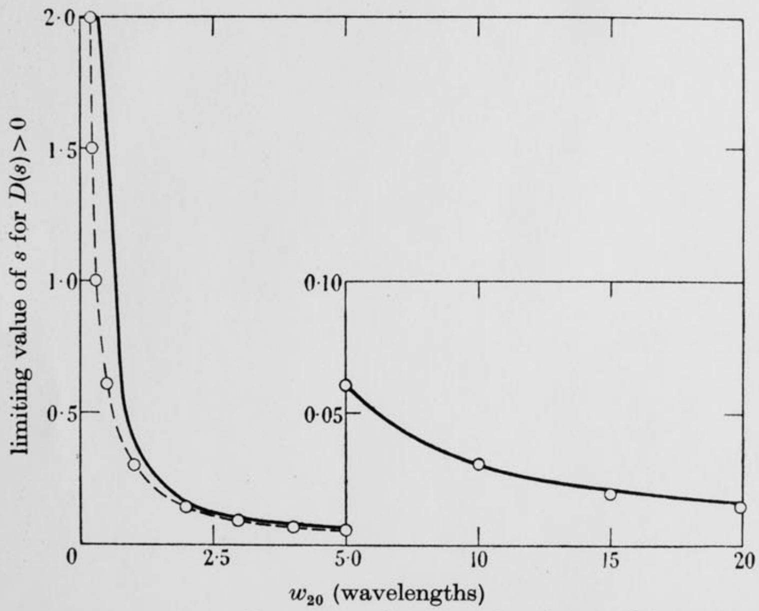


FIGURE 6. Bandwidth as a function of defect of focus. The points ○ are calculated on the basis of geometrical optics.

### 3. Comparison with geometrical optics

If the object consists of a single point source situated on the axis, we may write $B(u, v) = 2\pi\delta(u, v)$, $\delta(u, v)$ being a delta function. This has a Fourier spectrum $b(s, t) = 1$. Denoting by $D(s, t)$ the transmission factor for any chosen plane of focus, the intensity in the point source diffraction image in that plane is expressed by

$$B'(u', v') = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} D(s, t)\exp\{\mathrm{i}(u's + v't)\}\,\mathrm{d}s\,\mathrm{d}t,$$

inversion of which gives

$$D(s, t) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} B'(u', v')\exp\{-\mathrm{i}(u's + v't)\}\,\mathrm{d}u'\mathrm{d}v', \qquad (25)$$

showing that $D(s, t)$ is the inverse Fourier transform of the intensity distribution in the point source image. If this latter is calculated on the basis of geometrical optics, we are able to use (25) to find the response function to be expected according to geometrical optics.

For an aberration-free lens all rays from an axial object point pass through $O_0'$, the conjugate image point. In a plane, distance $\delta z'$ from the true focal plane, the image consists of a uniform circular patch of light of radius $\delta z' \tan \alpha'$, as shown in figure 3. If we introduce polar co-ordinates $p' = \sqrt{(u'^2 + v'^2)}$, $\psi' = \tan^{-1}(u'/v')$ $B'(u', v')$ is equal to a constant for $p' < p_0'$, where

$$p_0' = \frac{2\pi}{\lambda}(n' \sin \alpha')\, \delta z' \tan \alpha' = \frac{4\pi}{\lambda} w_{20} \sec \alpha'. \tag{26}$$

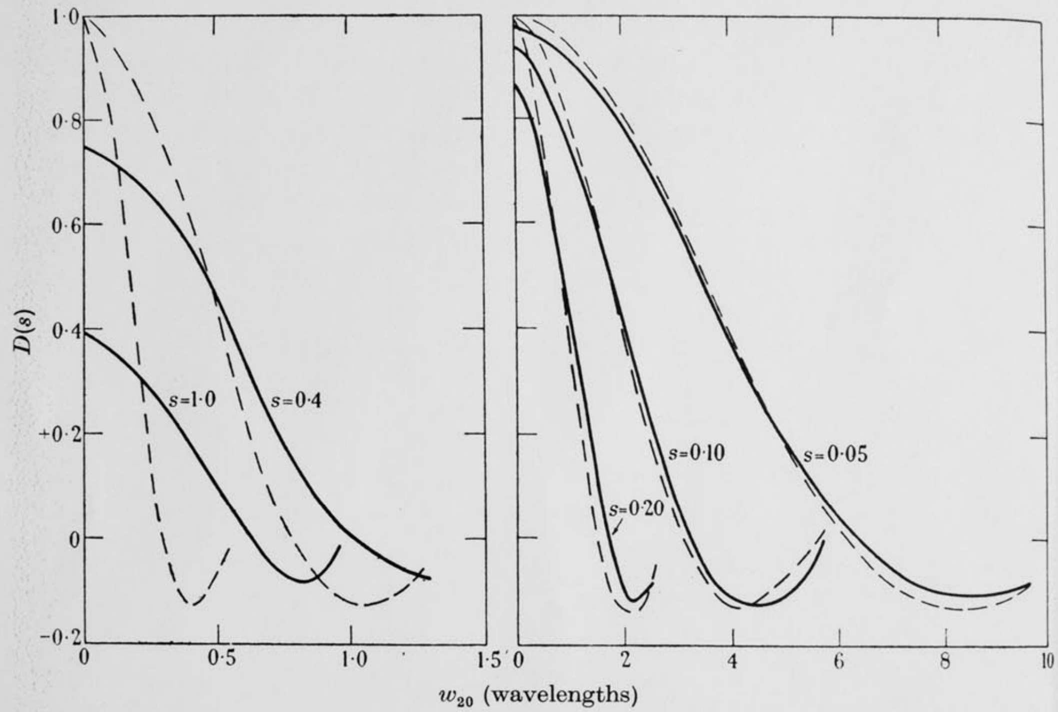


FIGURE 7. Response as a function of defect of focus. Full lines, according to diffraction theory; broken lines, according to geometrical optics.

Since the response to a frequency pair $(s_0, t_0)$ is the same as that for the single frequency $(s, 0)$ where $s = \sqrt{(s_0^2 + t_0^2)}$, we may put $t = 0$ in (25), and obtain

$$D(s) = \frac{1}{2\pi} \int_0^{p_0'} \int_0^{2\pi} \exp(-ip's \sin \psi')\, p'\, dp'\, d\psi',$$

the intensity in the image patch being taken equal to unity. This gives the formula

$$D(s) = \frac{2J_1(a)}{a} \quad \left(a = \frac{4\pi}{\lambda} w_{20} \,|\, s \,|\right), \tag{27}$$

when normalized to make $D(0) = 1$. The use of $a = 4\pi/\lambda w_{20} \,|\, s\,|$ implies that $\sec \alpha' = 1$, which is valid to the order of accuracy contemplated.

The first zero of (27) occurs when $a = 3{\cdot}83$, and this gives the limiting value of $s$ for which $D(s) > 0$. The bandwidth is therefore found from the formula

$$s = 0{\cdot}30\lambda/w_{20}. \tag{28}$$

Points calculated from this formula are shown in figure 6. From these results it appears that (28) is near enough valid for $w_{20} > 2\lambda$. It should be remembered, however, that this result applies only to the transmitted bandwidth, and not necessarily to the response for all frequencies within the bandwidth.

It is often suggested that geometrical optics holds for large aberrations and large defects of focus, and the above result supports this contention. What is more important, however, is to study how far geometrical optics is valid for different frequencies, irrespective of the magnitude of the aberration or defect of focus. To this end the curves in figure 7 have been drawn. Each curve shows the variation of $D(s)$ with increasing defect of focus. The full lines correspond to calculations based on diffraction theory, and the broken lines indicate what one would expect from geometrical optics.

A striking conclusion at once emerges, namely that geometrical optics gives results accurate to a small percentage for $|s| < 0{\cdot}10$, the maximum error reaching only about 12 % for $s = 0{\cdot}20$. For an $F/5$ objective $s = 0{\cdot}10$, $0{\cdot}20$ correspond in lines/mm to 20 and 40 respectively. This suggests that geometrical optics might be near enough valid for the treatment of image defects in photographic lenses, even though the aberrations are small, for resolution greater than $s = 0{\cdot}10$ will not often be in question.

## 4. A tolerance formula for defect of focus

Since the derivative of $D(s)$ at $s = 0$ is equal to $(-2/\pi)$, and is independent of aberration and defect of focus, the first approximations to $D_0(s)$, $D(s)$ will be of the form

$$D_0(s) = 1 - \frac{2}{\pi}|s| + O(s^3),$$

$$D(s) = 1 - \frac{2}{\pi}|s| + O(s^2).$$

and it is therefore in the coefficient of $s^2$ that one first finds the influence of aberration or defect of focus.

With this in mind (20) may be expanded to give

$$D(s) = 1 - \frac{2}{\pi}|s| - \frac{2\pi^2}{\lambda^2} w_{20}^2 s^2. \tag{29}$$

The ratio of the response of the defocused system to that obtaining in the true focal plane is the additional modulation arising from the defect of focus. Denoting this factor by $M(s)$,

$$M(s) = \frac{D(s)}{D_0(s)} = 1 - \frac{2\pi^2}{\lambda^2} w_{20}^2 s^2, \tag{30}$$

if quantities $O(s^3)$ are ignored. Let the acceptable defect of focus be such that the contrast does not fall below 80 % of that in the true focal plane, that is $M \geqslant 0{\cdot}80$. Using now the formula (13) and (17), for $w_{20}$ and $s$ respectively, gives the tolerance

$$\delta z' = \pm \frac{0{\cdot}20}{R' \sin\alpha'}, \tag{31}$$

with the notation used earlier. Measurements of the contrast in images in different planes of focus have been described by MacDonald (1951). The above expression shows good agreement with these results. In each case the focal spread for which $M(s) \geqslant 0{\cdot}80$ is slightly greater than that indicated by the formula (30). This is in accordance with the known increased depth of focus in the presence of aberration.

It is significant that the tolerance on $\delta z'$ according to (31) is independent of the wavelength, the refractive index of the image space, and involves $\sin\alpha'$ rather than the square of this quantity. This is understandable if one recalls that the scale of the point source diffraction image relative to the pattern size for a given $R'$ is proportional to $\lambda/n'\sin\alpha'$.

Comparison of the values of $D(s)$ calculated using the approximate formula (29) with those obtained from the full expression (20), shows that the error in (29) does not exceed 2 % providing $D(s) \geqslant 0{\cdot}80$. For the smaller values of $s$, less than about 0·2, the expression (30), and in consequence the tolerance formula (31), is also valid to the same order of accuracy.

## 5. Rectangular apertures

The form of the image of a line structure parallel to one axis of a rectangular aperture is of importance in the theory of the spectrograph. Let the angular width of the side of the aperture perpendicular to the direction of the line structure be $\alpha'$, as seen from the axial image point. Co-ordinates $(x, y)$, $(u', v')$ are defined as in (1) and (2) above. The rectangular aperture is then of half-width $x = 1$. Let $y = y_0$ be the half-length of the slit. The area of the aperture is then $g(0,0) = 4y_0$, and in place of (19), we write

$$D(s) = \frac{1}{4y_0}\iint_s \exp(\mathrm{i}as)\,\mathrm{d}x\,\mathrm{d}y,$$

the symbol $s$ now denoting the area common to two rectangles with their centres at the points $(\pm\frac{1}{2}s, 0)$. The limits of $x$ are thus $\pm(1-\frac{1}{2}|s|)$, and those of $y$ are $\pm y_0$. The transmission function is therefore

$$\left.\begin{aligned} D(s) &= \frac{\sin\{a(1-\frac{1}{2}|s|)\}}{a} \quad (|s| \leqslant 2),\\ &= 0 \quad (|s| > 2),\end{aligned}\right\} \tag{32}$$

which reduces to the known result for the in-focus image

$$\left.\begin{aligned} D_0(s) &= 1-\tfrac{1}{2}|s| \quad (|s| \leqslant 2),\\ &= 0 \quad (|s| > 2),\end{aligned}\right\} \tag{33}$$

as $w_{20} \to 0$. Recalling that $a = 4\pi/\lambda w_{20}|s|$, the defect of focus being now measured along the line $x = +1$, the bandwidth is found from the value of $s$ giving the first zero of $D(s)$ (32), that is as the appropriate root of the equation

$$s^2 - 2s + \frac{\lambda}{2w_{20}} = 0, \tag{34}$$

or, for larger values of $w_{20}$,

$$s = 0{\cdot}25\lambda/w_{20}, \tag{35}$$

which, we shall see, is precisely the result obtained on the basis of geometrical optics.

The geometrical point-source image in a plane distance $\delta z'$ from the true focal plane comprises a uniformly illuminated rectangular patch of half-width

$$u_0' = 4\pi/\lambda w_{20} \sec \alpha',$$

and half-length $= y_0 u_0'$. The formula (25) now becomes

$$D(s,0) = \int_{-u_0'}^{+u_0'} \int_{-y_0 u_0'}^{+y_0 u_0'} \exp(-\mathrm{i} s u')\,\mathrm{d}u',$$

apart from a normalizing factor to make $D(0,0) = 1$. According to geometrical optics, therefore, the response function is

$$D(s) = \sin a/a, \tag{36}$$

in which $\sec \alpha'$ has again been put equal to unity. The bandwidth is thus given by (35). In this case it is easy to see how the transition to geometrical optics occurs with increasing defect of focus. For, on the one hand, the appropriate root of (34) becomes progressively less different from the value (35), and (32) behaves increasingly like (36), the significant range of values of $s$ being small. Moreover, if $s < 0{\cdot}10$ (say) the transmission factor (32) does not differ from that calculated from (36) by more than a small percentage. This confirms the result obtained numerically for circular apertures, namely that geometrical optics can be expected to give results of reasonable accuracy for structures for which $s < 0{\cdot}10$.

The response function found above for a rectangular aperture is qualitatively similar to that for a circular aperture. Moreover, being in closed form, it is a simple matter to obtain numerical values. For these reasons numerical results have not been given in the text.

I am indebted to Miss J. M. Drewitt both for checking the analysis and for the numerical computations.

## References

Duffieux, P. M. 1946 *L'intégrale de Fourier et ses applications à l'optique.* Bescançon: privately printed.

Hopkins, H. H. 1950 *Wave theory of aberrations.* Oxford: Clarendon Press.

Hopkins, H. H. 1953 *Proc. Roy. Soc.* A, **217**, 408.

MacDonald, D. E. 1951 *Symposium 'Optical image evaluation'.* U.S. Bureau of Standards, Circular 526, 1954, p. 62.

Titchmarsh, E. C. 1937 *Theory of Fourier integrals.* Oxford University Press.

Notes by Rik Littlefield, rj.littlefield@computer.org, 6/27/2014:

1. The variable *s* is spatial frequency, in line pairs per unit length, scaled such that *s* = 2 is cutoff based on the Airy disk.

2. To further make sense of Equation 31, it may be helpful to compare it to the standard formula for quarter-lambda wavefront error. Then we have:

$$\delta z' = \pm \frac{0{\cdot}20}{R' \sin \alpha'} \qquad (31)$$

$\delta z' = \pm \frac{\lambda}{2NA^2} = \pm \frac{\lambda}{2\,(\sin \alpha')^2}$ (ignoring refractive index)

When are these two values equal?

$$\pm \frac{0.20}{R' \sin \alpha'} = \pm \frac{\lambda}{2NA^2} = \pm \frac{\lambda}{2\,(\sin \alpha')^2}$$

$$\frac{0.20}{R' \sin \alpha'} = \frac{\lambda}{2\,(\sin \alpha')^2}$$

$$R' = 0.2 \frac{2 \sin \alpha'}{\lambda}$$

But $\frac{2 \sin \alpha'}{\lambda}$ is recognizable as the cutoff frequency.

So, one special point for Hopkins' formula occurs when we run at quarter-lambda wavefront error and focus our attention on the frequency R' where D(*s*) = 0.80, roughly out to 0.2 of the cutoff frequency for that lambda.

If we now consider what happens when lambda changes, we find that the MTF curve as a whole sags by the change squared, but our position along that curve moves left (or right) by 1/change, and because the effect of defocus is quadratic near *s*=0, everything cancels out so that the loss of contrast does not depend on lambda.

Dependence on sin *a*' (instead of its square) and lack of dependence on lambda are also characteristics of the ray optics model, so equation 31 is further substantiation of Hopkins' viewpoint that ray optics is close enough for many purposes.

Paraphrasing, diffraction only matters when you're concerned with frequencies that are a substantial fraction of cutoff.