AXIAL ASTIGMATISM OF THE ELECTRON MICROSCOPE OBJECTIVE

DISSERTATION

Presented in Partial Fulfillment of the Requirements
for the Degree Doctor of Philosophy in the
Graduate School of The Ohio State
University

By

ROBERT ARNOLD WATKINS, ScB, M.S.
The Ohio State University
1953

Approved by:

[Signature]
Adviser
To
Albert F. Prebus
ACKNOWLEDGMENTS

I could never give a full list of the persons who knowingly and unknowingly aided me in writing this dissertation, from my own family to the men I studied under at The Ohio State University.

In particular, I wish to express my gratitude to Dr. R. Adler, Assistant Director of Research and to Mr. D. Hoag, both of the Zenith Radio Corporation. I also wish to thank Dr. R. Shaw of Forest Park, Illinois who furnished the lenses used in the experimental verification of the calculations.

I sincerely thank Dr. A.F. Prebus, my adviser, who helped with services so far beyond those to be expected and who furnished the micrographs of Figure 12 as well as calibration data.
TABLE OF CONTENTS

I  INTRODUCTION
1. Statement of the Problem..............................1
2. Paraxial Trajectory Equations in an
    Asymmetric Magnetic Field..........................3
3. General Solution of the Trajectory Equations for
    Electron Motion in an Asymmetric Magnetic Field...8
4. Image Formation in the Glaser Field
    \[ H(z) = \frac{H_0}{1 + \left(\frac{z}{a}\right)^2} \text{ with Asymmetry } = \frac{Q H'}{I + \left(\frac{z}{a}\right)^2} \]..........................17

II  EVALUATION OF THE RESOLVING POWER OF AN
    ASYMMETRIC LENS
1. Determination of the Wave Surface Corresponding
    to a Converging Asymmetric Wave.....................23
2. The Evaluation of the Kirchhoff Solution of the
    Wave Equation for a Converging Asymmetric Wave...26
3. Numerical Results of the Theory.....................32
4. Light Optic Experimental Verification...............39

III  PROPOSED ELECTRIC CYLINDRICAL LENS FOR
    CORRECTION OF OBJECTIVE LENS ASYMMETRY
1. Magnetic Correction of Objective Lens Asymmetry...............43
2. Paraxial Electron Optics of the Correcting
    System.............................................44
3. Geometric Optics of the Correcting System...............49
4. Physical Construction of the Correcting Lens............52
5. Experimental Results................................56
6. Conclusions........................................59

IV  APPENDICES
A. Formal Analogy Between Light and Electron Optics...64
B. The Evaluation of the Integral........................67
C. Consideration of the Effect of Spherical
    Abberation on the Images Formed in the Light
    Optical Analogue...................................73

V  REFERENCES.........................................83

VI  AUTOBIOGRAPHY......................................84
ILLUSTRATIONS

Fig. 1 The trajectories and their projections for motion in an asymmetric field and the corresponding symmetric field.

Fig. 2 Image distortion due to axial astigmatism.

Fig. 3 Location of planes of interest in image space and the image of a point source formed on each.

Fig. 4 The Glaser field and definition of the new angle.

Fig. 5 Geometry and notation used in developing the asymmetric wave surface.

Fig. 6 Synthesis of essentially spherical wave surface.

Fig. 7 Relative intensity as a function of separation from the geometric image point for C=2 or F2 - F1 = .265 microns.

Fig. 8 Relative intensity as a function of separation from the geometric image point for C=3 or F2 - F1 = .48 microns.

Fig. 9 Equi-intensity contours for C=2.

Fig. 10 Equi-intensity contours for C=3.

Fig. 11 Equi-intensity contours for two self luminous points separated by s = 4 (6.8 μ for typical microscope lens) with the points lying along a focal line.

Fig. 12 Equi-intensity contours for two self luminous points separated by s = 4 (6.8 μ for typical microscope lens with the points lying on a line at 45° to a focal line.

Fig. 13 Diffraction patterns obtained for slightly astigmatic lenses.

Fig. 14 Diffraction patterns obtained with a range of values of astigmatism C=1.35 to C=13.5 in steps of 1.35.

Fig. 15 The effect of correction of an asymmetric lens with a cylindrical lens.

Fig. 16 Axial potential of the four slit lens.
Fig. 17 Geometry and notations used in developing the geometric optics of the correcting system.

Fig. 18 Schematic representation of the correcting electrostatic lens.

Fig. 19 Sectional drawing of the correcting electrostatic lens. Twice actual size.

Fig. 20 High voltage power supply.

Fig. 21 Fringe patterns observed in micrographs made with lenses having various amounts of astigmatism.

Fig. 22 Astigmatism introduced by the electrostatic cylindrical lens as a function of voltage.

Fig. 23 Geometry and notation used in developing the aspherical wave surface.

Fig. 24 Geometry and notation used in ray tracing.
I. INTRODUCTION

1. Statement of the Problem

The usefulness of the electron microscope as a tool for scientific research has been well demonstrated in the past fifteen years. Its usefulness arises from the very short effective wavelength of the source of illumination. As a result, the resolving power of this instrument as limited by diffraction effects is much greater than that of the conventional microscope. In particular, if the accelerating potential is 50,000 volts, and the large apertures used in the light microscope were possible, the effective wavelength of 0.05 Å would allow two objects separated by 0.03 Å to be resolved, that is, identified as separate objects. By comparison, using a high-power compound light microscope with ultraviolet light, the objects would have to be separated 1000 Å in order to be resolved.

Since image contrast is predominantly due to differential scattering, a corrected electron lens would need to be stopped down to small apertures with a diaphragm in order to have sufficient contrast. Further, the electric and/or magnetic fields that are used to form the lenses for the microscope have inherently a very large amount of spherical aberration. For these reasons, it is necessary to use very small apertures, in the range of 0.005 radians, which would indicate a resolution of 7 Å. The published
estimates of resolution are generally of the order of 20Å with uncompensated lenses. In all of these cases, extreme care had been taken in the operation of the microscope and hence some thought has been devoted to the discrepancy of a factor of 3.

The possibility of explaining this discrepancy in terms of an asymmetry of the objective pole piece field was first considered by Glaser. This asymmetry may be caused by either improper machining or inhomogeneities of the pole piece material. These field asymmetries lead to the formation of an astigmatic lens, that is, a lens which has different focal lengths in two perpendicular planes whose intersection is the optic axis. Glaser has shown that on the basis of geometric optics, a very small amount of field asymmetry can lead to a circle of confusion of the order of 50Å.

In a later paper, Hillier found essentially the same result for a different form of lens. In addition, Hillier indicated that the amount of astigmatism may be determined by the observation of the Fresnel fringes formed about the image of a hole in a thin membrane. He further gave a method whereby the correction of the astigmatism can be accomplished by the introduction of another asymmetric magnetic field. This method of correction, while quite effective, is time consuming and not permanent. Further, since this method also
corrects for the effect of any stray fields, the correction is valid for only one magnification. The latter defect is not ordinarily more than an inconvenience; the former precludes the possibility of correction for asymmetries that might be introduced by the specimen. Using this method, Hillier claimed the microscope was inherently capable of resolving $10^2$.

It is therefore the purpose of this work firstly, to evaluate the effect of axial astigmatism upon the diffraction pattern that is obtained as the image of a point source formed by an asymmetric lens; and secondly, to present a new method for correction of these asymmetries.

2. **The Paraxial Trajectory Equations in an Asymmetric Magnetic Field**

In order to discuss the image formation by electron lenses, it is necessary to determine the paths followed by the electrons under the influence of the existing static electric and magnetic fields. This motion takes place in a region free of magnetic materials, and it will be assumed that the effect of other electrons in the beam on the motion of the electron under consideration is negligible.

Under these conditions, it is possible to demonstrate that the electron paths as determined from the Newtonian equations of motion are identical with those determined from the variation problem:
or of the variation problem:

\[ (1.202) \int_{b}^{a} F'ds = 0; \quad F' = \sqrt{-2e m \phi} \cdot \frac{1}{\sqrt{1 + x'^2 + y'^2}} - \frac{e}{\varepsilon} \left( x'A_x + y'A_y + A_z \right) \]

(See reference 1 or appendix A)

In the above equations \( \phi \) is the electrostatic potential and \( \mathbf{A} \) is the magnetic vector potential of the imaging fields; \( e \) is the charge on the particles, \( m \) is their mass, and \( c \) is the velocity of light. The \( z \) axis will be taken to lie along the optic axis of the system.

This latter statement of the problem (1.201) is identical in form to a statement of Fermat's principle of minimum optical path length if \( F' \) is considered the index of refraction of the medium. That is, the trajectory of a charged particle in electric and magnetic fields is identical to that of a corresponding light ray which passes through a medium whose index of refraction varies according to \( F' \).

In the subsequent treatments, consideration will be given only to those paths which deviate but slightly from the optic axis, and whose slopes with the axis are likewise small. The paths so defined are termed the paraxial paths.

In determining the paraxial electron optics for the microscope objective, the case of a particular form of
magnetic field will be discussed. This field has two planes of symmetry, perpendicular to each other, and to the optic axis. This treatment will vary little from that of Glaser. The principal difference is in the assumption that the field is of finite axial length. This condition, which is met in practice, is particularly important since it shows the theoretical possibility of correction of an asymmetric magnetic lens field with an electrostatic cylindrical lens. If the X and Y axes are chosen to lie in the planes of symmetry, and the z axis to lie along the optic axis, it is then possible to expand the scalar magnetic potential in terms of its value along the optic axis, $\psi(z)$:

$$(1.203) \quad \psi(x, y, z) = \psi_0(z) + \psi_2(z) x^2 + \psi_2(z) y^2$$

where $\psi_0(z)$ and $\psi_2(z)$ are axial functions as yet to be determined. The asymmetry function $\Delta(z)$ is defined by the axial functions in the equation:

$$(1.204) \quad \Delta(z) = 2 \left( \psi_2(z) - \psi_2(z) \right)$$

The scalar potential must satisfy the Laplace equation. By substituting equation (1.204) in (1.203) and the resulting expansion for $\psi$ in the Laplace equation, we have:

$$(1.205) \quad \psi''(z) + 4 \psi_2(z) + \Delta(z) = 0$$
Equation (1.205) can be solved for \( \psi(z) \), and the resulting value substituted in equation (1.203) to give the following expansion to the second order in \( X \) and \( Y \) in terms of the value of the potential along the optic axis and its second derivative:

\[
(1.206) \quad \psi(x, y, z) = \psi_0(z) - \frac{\psi''(z) + \Delta(z)}{4} x^2 - \frac{\psi''(z) - \Delta(z)}{4} y^2
\]

The definition of the magnetic scalar potential is \( \mathcal{H}(x, y, z) = -\nabla \psi(x, y, z) \). Therefore on the optic axis

\[
(1.207) \quad \mathcal{H}(0, 0, z) = \mathcal{H}_0(z) = -\psi'(z)
\]

If (1.207) is substituted in (1.206), the expansion for \( \psi(x, y, z) \) becomes:

\[
(1.208) \quad \psi(x, y, z) = \psi_0(z) + \frac{\mathcal{H}_0' - \Delta}{4} x^2 + \frac{\mathcal{H}_0' + \Delta}{4} y^2
\]

The vector potential \( \mathbf{A} \) is defined by \( \mathbf{H} = \nabla \times \mathbf{A} \). If the components of \( \mathbf{H} \) determined from the definition of the magnetic scalar potential (using (1.208)) are equated to the like components of \( \nabla \times \mathbf{A} \), and the resulting three equations solved, it is found that the components of the magnetic vector potential are given by:

\[
(1.209) \quad A_x = -\frac{1}{2} \mathcal{H}_0'(z) y \quad A_y = \frac{1}{2} \mathcal{H}_0'(z) x \quad A_z = \frac{1}{2} \Delta(z) xy
\]

If the values from (1.209) are substituted in \( F(x, y, z) \) with \( \phi = \phi' = \text{constant} \):
where only terms to the second order in \( X, Y, X', Y' \), are retained. We now introduce a second coordinate system whose \( z \) axis is common to the original \( z \) axis, and whose \( x \) and \( y \) axes are rotated an angle \( \psi \) from the original \( X-Y \) axes. Then in terms of the new coordinates

\[
X = x \cos \psi - y \sin \psi ; \quad Y = x \sin \psi + y \cos \psi.
\]

If \( \delta \) is now chosen as

\[
\delta(z) = \int_{-\infty}^{z} H(z') \phi^{-h_0} dz',
\]

then

\[
\Phi(x, y, x', y', z) = \Phi_0^h + \frac{\Phi_0^h (x^2 + y^2)}{2} - \frac{q \phi_0^h}{\phi_0^h} x + \frac{q \phi_0^h}{\phi_0^h} y - \frac{q \phi_0^h \sin 2\psi (x', y')} - \frac{q \phi_0^h \cos 2\psi x y}{2}
\]

Having determined \( \Phi(x, y, x', y', z) \), we can immediately set down the paraxial trajectory equations for an electron in this magnetic field in a region of constant electric potential as the Euler-Lagrange equations:

\[
\frac{d}{dz} \left( \frac{\phi_0^h}{\phi_0^h} \right) + \frac{q \phi_0^h}{\phi_0^h} x = -\frac{1}{2} \eta \Delta \sin 2\psi x - \frac{1}{2} \eta \Delta \cos 2\psi y
\]

\[
\frac{d}{dz} \left( \frac{\phi_0^h}{\phi_0^h} \right) + \frac{q \phi_0^h}{\phi_0^h} y = +\frac{1}{2} \eta \Delta \sin 2\psi y - \frac{1}{2} \eta \Delta \cos 2\psi x
\]

In the case of a perfect, axially symmetric lens \( \Delta(z) \) is zero and the equations for \( x \) and \( y \) are identical in form and further, \( x \) and \( y \) are independent of each other. The image plane is rotated by an angle \( \psi \) from the object plane.

In the case of a pure electric field, a treatment similar to the above leads to equations that give \( x \) and \( y \)
as independent functions of $z$. In this case, the functions are no longer identical. A point object will be brought to focus in two lines, perpendicular to each other and to the optic axis, and with different intersections on the optic axis. There is no rotation of the image plane with respect to the object plane.

In the case being considered, there is a dependence between $x$ and $y$, and further, there exists the rotation $\gamma$.

3. General Solution of the Trajectory Equations for Electron Motion in an Asymmetric Magnetic Field

First it is assumed that the corresponding homogeneous equation

$$\frac{d}{dz} \left( z \sqrt{\eta} \right) + \frac{1}{\sqrt{\eta}} \left( \eta \frac{d}{dz} \right) = 0$$

has been solved, and has the two independent solutions, $g(z)$ and $h(z)$. It is further assumed that a sufficiently accurate result will be obtained for the general problem with the addition of the small perturbation functions, $s(z)$ and $t(z)$. The abbreviations

$$S_1 = \frac{1}{2} \eta \Delta \sin 2\gamma, \quad S_2 = \frac{1}{2} \eta \Delta \cos 2\gamma$$

are introduced, and the final assumption made that $tS_1$, $tS_2$, $sS_1$, and $sS_2$ may be neglected. We take as the particular solutions of the homogeneous equation (1.301) those with the properties:
Figure 1. The trajectories and their projections for motion in an asymmetric field and the corresponding symmetric field.

\[(1.303) \quad q(\alpha) = 1, \quad q'(\alpha) = 0, \quad h(\alpha) = 0, \quad h_o = h(z_o) = 1\]

From the form of the homogeneous equation, it can be seen that \( g \frac{d}{dz}(h'p^3) - h \frac{d}{dz}(g'p^3) = 0 \). After integration, this results in the equation:

\[(1.304) \quad q' h - q h = \text{const.} = h_o\]

where the evaluation of the constant is in accordance with the properties \((1.303)\).

Since any particular solution of the homogeneous equation can be written in the form \( x_o(z) = x_o g(z) + x_a h(z) \)
\( y_o(z) = y_o g(z) + y_a h(z) \), where \( x_o, y_o \) are the coordinates of the intersection of the trajectory and the object plane and \( x_a, y_a \) are the coordinates of the intersection of the trajectory and the aperture plane, (see figure 1) the solutions of the perturbed equations must be of the form:

\[(1.305) \quad x(z) = x_o g(z) + x_a h(z) + s(z) \quad y(z) = y_o g(z) + y_a h(z) + t(z)\]
where \( s(z_0) = s(z_a) = t(z_0) = t(z_a) = 0 \). If these forms are substituted in equations (1.212) and (1.213), and use is made of the above abbreviations and properties of \( g \) and \( h \), we obtain the following equations for \( s(z) \) and \( t(z) \):

\[
(1.306) \quad \frac{d}{dz} \left( s' g \right) + \frac{1}{\sqrt{\phi}} (\eta^* H') s = S_1 (x_0 g + x_w h) - S_2 (y_0 g + y_w h) = f_1
\]

\[
(1.307) \quad \frac{d}{dz} \left( t' g \right) + \frac{1}{\sqrt{\phi}} (\eta^* H') t = -S_1 (y_0 g + y_w h) - S_2 (x_0 g + x_w h) = f_2
\]

These last equations are conveniently solved by the method of variation of parameters. We therefore assume that \( s(z) \) can be written in the form

\[
(1.308) \quad s(z) = u(z) g(z) + v(z) h(z)
\]

and further that

\[
(1.309) \quad u'(z) g(z) + v'(z) h(z) = 0
\]

Making use of (1.309), we substitute (1.308) and derivatives determined from it in (1.306) to obtain

\[
(1.310) \quad u'(z) g'(z) + v'(z) h'(z) = \frac{f_1}{\sqrt{\phi}}
\]

Solving (1.309) and (1.310) algebraically for \( u'(z) \) and \( v'(z) \) provides the results:

\[
(1.311) \quad u' = \frac{h f_1}{\sqrt{\phi} (g h - h g)}, \quad v' = \frac{-g f_1}{\sqrt{\phi} (g h - h g)}
\]

Making use of (1.304) and integrating, we obtain values for \( u \) and \( v \) which may be substituted in (1.308) to give
In a similar fashion, $t(z)$ is determined as

$$t = \frac{1}{\eta^2 q^2} \left\{ -q \int_{z_0}^{z} h \, dz + h \int_{z_0}^{z} q \, dz \right\}$$  \hspace{1cm} (1.313)

It is to be noted that the lower limits of integration are different in the first and second terms. This is a result of the boundary conditions $s(z_0) = s(z_a) = t(z_0) = t(z_a) = 0$. Ordinarily, the aperture plane may be considered to lie in field free space. (Its position, as defined above, coincides with the image side focal point.)

Under these conditions, $f_1$ and $f_2$ are zero over the region of integration for the second terms, and hence these terms may be neglected. If the values of $f_1$ and $f_2$ as defined in (1.306) and (1.307) are substituted in (1.312) and (1.313), we obtain four basic integrals which have constant values in image space, and which we shall designate as:

$$A = \frac{1}{h^2 q^2} \int_{z_0}^{z} h^2 S_z \, dz \hspace{1cm} B = \frac{1}{h^2 q^2} \int_{z_0}^{z} h^2 S_z \, dz$$

$$C = \frac{1}{h^2 q^2} \int_{z_0}^{z} h^2 S_z \, dz \hspace{1cm} D = \frac{1}{h^2 q^2} \int_{z_0}^{z} h^2 S_z \, dz$$ \hspace{1cm} (1.314)

Using these abbreviations, and substituting the values of $s$ and $t$ as determined in (1.312) and (1.313) in (1.305) we have the equations of the paraxial trajectories in terms of the trajectories in the corresponding symmetric system and the intersections in the object and aperture.
Object Image

Figure 2. Image distortion due to axial astigmatism. Unity magnification, \( C = D = 0.25 \).

planes as:

\[
(1.315) \quad x(z) = x_o g(z)(1+c) - y_o g(z)D + x_o [h(z) + A g(z)] - y_o B g(z)
\]

\[
(1.316) \quad y(z) = -x_o g(z)D + y_o g(z)(1-c) - x_o B g(z) + y_o [h(z) - A g(z)]
\]

The first two terms give the location of the center of the aberation figure in a given plane in image space, and the last two terms are descriptive of the figure itself. It is obvious that even with infinitessimal apertures there still exists a distortion in the imaging due to the axial astigmatism. (See figure 2.) This distortion in itself causes no loss of resolution, the loss being due to the aberation figure determined by the last two terms. It is therefore these terms that are of primary interest.

To avoid having the first two terms present, we will
consider an axial object point, \( x_0 = y_0 = 0 \). In considering
the last terms, there are three values of \( z \), that is, three
planes in image space that are of most interest. (See figure 3.) The first of these is that for which \( h(z_1) = 0 \).
This corresponds to the Gaussian image plane for the
axially symmetric system. Under these conditions,

\[
(1.317) \quad x(z) = x_c = g(z)[A x_c - B y_c]; \quad y(z) = y_c = g(z)[-B x_c - A y_c]
\]

Introducing polar coordinates \( x = r \cos \phi, y = r \sin \phi \);
\( x_a = r \cos \phi, y_a = r \sin \phi \) in the image and aperture planes,
we obtain for the figure in the Gaussian image plane

\[
(1.318) \quad r_c = [x_c^2 + y_c^2]^{1/2} = g(z)[(A^2 + B^2)x_c^2 + y_c^2]^{1/2} = g(z)\sqrt{A^2 + B^2}; \quad \tan \phi = -\frac{A \cos \phi - B \sin \phi}{B \cos \phi + A \sin \phi}
\]

From these results, we see that the aberration figure has
dimensions differing from those of the aperture by a con­
stant factor, but which suffers an angular distortion. In
the case of a circular aperture, the aberration figure is
a circle.

The other planes of interest are those in which there
exists a line focus. The location of these planes is not
immediately apparent, but will be determined from the con­
ditions necessary for a line focus. We introduce a new set
of coordinates rotated through an angle \( \Theta \) from the \( x, y \)
axes. The new coordinates are designated as \( \xi, \eta \). In
terms of the new coordinates, \( x = \xi \cos \Theta - \eta \sin \Theta; y = \xi \sin \Theta + \eta \cos \Theta \).
Introducing these values of \( x \) and \( y \) in (1.315) and solving
for $\xi$ and $\eta$, we have:

\[(1.319) \quad \xi = x_a [(h+Ag)\cos \theta -Bg \sin \theta] + y_a [(h-Ag)\sin \theta - Bg \cos \theta] \]

\[(1.320) \quad \eta = x_a [-Bg \cos \theta -(h+Ag)\sin \theta] + y_a [(h-Ag)\cos \theta + Bg \sin \theta] \]

If $\theta$ is now chosen so that $\tan \theta = -(h - Ag)/Bg$, then the coefficient of $y_a$ in (1.320) is zero, and the coefficient of $x_a$ becomes

\[\frac{\cos \theta}{Bg} [h^2-g^2(A^2+B^2)] \]

This coefficient is zero if $h(z) = \pm \sqrt{A^2 + B^2} g(z)$. If this value of $h(z)$ is substituted in the expression for $\tan \theta$, we obtain:

\[(1.321) \quad \tan \theta_1 = \frac{A-\sqrt{A^2+B^2}}{B}, \quad \tan \theta_2 = \frac{A+\sqrt{A^2+B^2}}{B} \]

The product of the tangents in (1.321) is -1. Thus the two focal lines are perpendicular to each other, and of course lie in planes perpendicular to the optic axis.

If we further substitute the values of $h(z)$ and $\theta$ in the expression for $\xi$, we find

\[(1.322) \quad \xi_1 = \frac{2Bg \sqrt{A^2+B^2} x_a}{\sqrt{2(A^2+B^2)-2A^2+B^2}} - y_a \frac{2(A^2+B^2)-2A^2+B^2}{2(A^2+B^2)+2A^2+B^2} \]

\[\xi_2 = -\frac{2Bg \sqrt{A^2+B^2} x_a}{\sqrt{2(A^2+B^2)+2A^2+B^2}} - y_a \frac{2(A^2+B^2)+2A^2+B^2}{2(A^2+B^2)-2A^2+B^2} \]

and the maximum value of $\xi$ in either case is

\[(1.323) \quad \xi_{\text{max}} = 2Bg \sqrt{A^2+B^2} \]

Finally, we wish to determine the location of the
Figure 3. Location of planes of interest in image space and the image of a point source formed on each.
planes in which the focal lines lie. To do this, we make use of the fact that the region beyond the aperture plane is field free, in consequence of which the trajectories are straight lines. By definition, the trajectory \( h(z) \) passes through the aperture plane with a separation of one unit from the axis, and intersects the Gaussian image plane at the optic axis. The trajectory \( g(z) \) intersects the aperture plane at the optic axis, and the Gaussian image plane at a separation \(-M\) units from the axis where \( M \) is the magnification of the object by the lens. (See figure 3.) Thus we can write:

\[
(1.324) \quad h(z) = \frac{z_i - z}{L}; \quad g(z) = \frac{M(z_i - z)}{L} \quad \text{where} \quad L = z_i - z
\]

Substituting these values in \( h = \pm \sqrt{A^2 + B^2} \) \( g \), we get

\[
(1.325) \quad z_i = z_i \pm \frac{LM\sqrt{A^2 + B^2}}{1 + M\sqrt{A^2 + B^2}}
\]

The results of the geometric optical image formation by an asymmetric magnetic lens can be summarized as follows. There exist two planes in which the image is a straight line, and these line images are perpendicular to each other. Further, these lines are not in coincidence with the original \( x \) and \( y \) axes. On a third plane, lying between the planes of the focal lines and on the aperture side of the midpoint between them, the image is a circle. The separation of the planes containing the focal lines increases with increasing asymmetry. The length of the
focal lines and the diameter of the circular image increase both with increasing asymmetry and increasing aperture.

\[ \text{4. Image Formation in the Glaser Field, } H(z) = \frac{H_0}{1 + (\frac{z}{a})^2} \]

With Asymmetry \( \Delta(z) = \frac{Q H'}{1 + (\frac{z}{a})^2} \)

Several axial magnetic field distributions have been proposed which offer reasonable approximations to the distributions actually found in magnetic lenses, and which at the same time allow analytical solution of the paraxial ray equation. Of these, the so called Bell Shaped field proposed by Glaser is the most commonly used, although it is not as steep sided beyond the half width points as fields normally used. This field is

\[ (1.401) \quad H(r) = \frac{H_0}{1 + (\frac{r}{a})^2} \]

where \( H_0 \) is the maximum value of the field, and \( 2a \) is the half width of the field. (See figure 4.) If we substitute this field in the paraxial ray equation for an axially symmetric magnetic field (see(1.212)):

\[ (1.402) \quad \eta'' = -\frac{e H}{\gamma mc^2} \phi \]

and make use of the abbreviations: \( x = \frac{z}{a} \), \( y = \frac{r}{a} \), \( K^2 = \frac{eH_0 a^2}{\gamma mc^2 z} \)

we obtain

\[ (1.403) \quad y'' = -\frac{K^2 y}{(1+x)^2} \]
To solve equation (1.403), we first introduce the new variable $\phi$ by the relation $x = \text{ctn} \phi$. (See figure 4.)

Substituting, we obtain

$$(1.404) \quad \frac{d^2y}{d\phi^2} + 2 \text{ctn} \phi \frac{dy}{d\phi} + k^2 y = 0$$

We now make the further substitution, $y = A(\phi)f(\phi)$ in (1.404)

$$(1.405) \quad A \frac{d^2f}{d\phi^2} + \left(2 \frac{df}{d\phi} + 2A \text{ctn} \phi \frac{df}{d\phi} + \frac{dA}{d\phi} + 2 \frac{dA}{d\phi} \text{ctn} \phi + k^2 A \right)f = 0$$

and finally choose $A(\phi)$ so that the coefficient of $\frac{df}{d\phi}$ vanishes.

$$A = e^{-\int \text{ctn} \phi d\phi} = \frac{1}{\sin \phi}$$

$$(1.406) \quad \frac{d^2f}{d\phi^2} + (k^2+1)f = 0 \quad f = c_1 \sin \left( \frac{k \phi}{\sqrt{k^2+1}} \right) + c_2 \cos \left( \frac{k \phi}{\sqrt{k^2+1}} \right)$$

The general solution in terms of $\phi$, is therefore

$$(1.407) \quad y(\phi) = \frac{1}{\sin \phi} \left\{ c_1 \sin \left( \frac{k \phi}{\sqrt{k^2+1}} \right) + c_2 \cos \left( \frac{k \phi}{\sqrt{k^2+1}} \right) \right\}$$

By imposing the conditions $r(\phi) = 1$, $r'(\phi) = 0$
on the general solution, we obtain the particular solution,

$$g(\phi) = \frac{1}{\sqrt{k^2+1}} \left\{ \sin \phi \cos \left( \frac{k \phi}{\sqrt{k^2+1}} \right) + \cos \phi \sin \left( \frac{k \phi}{\sqrt{k^2+1}} \right) \right\} \quad \text{where} \quad \omega = \sqrt{k^2+1} \quad \text{and by}$$
setting \( g = 0 \), we determine the location of the aperture plane from the relation \( \omega \tan \phi = - \frac{\tan \omega (\varphi - \varphi_0)}{\omega} \). Using this last result, and the conditions \( r(\varphi) = 0 \), \( r(\varphi) = 1 \) in the general solution, we obtain the particular solution \( h(\varphi) \). If \( g(\varphi) \) is now rewritten to show explicit dependence on \( \varphi \), we have

\[
\begin{align*}
 h(\varphi) &= \frac{\sin \omega (\varphi - \varphi_0)}{\omega} \sin \omega (\varphi - \varphi_0) \quad \text{and} \quad \frac{\sin \omega (\varphi - \varphi_0)}{\omega} \\
 \omega \tan \phi &= - \frac{\sin \omega (\varphi - \varphi_0)}{\omega} \quad \theta = \sqrt{\kappa^2 + 1} \quad \varphi = \alpha \cot \varphi
\end{align*}
\]

(1.408)

for the particular solutions referred to in the last section.

In the selection of the function to represent the field asymmetry, we again adopt the choice of Glaser. Referring to (1.206), it can be seen that if the asymmetry is assumed to be of a mechanical nature, then its effect should increase in the same fashion as the axial function of the primary field which leads to image formation, and hence \( A(z) \) should be of the form

\[
\Delta(z) = -Q \frac{\psi'(\varphi)}{\psi(\varphi)} = Q \frac{H_0}{[1 + (1 + 1)]} \frac{\varphi}{\alpha} \quad \Delta(\varphi) = -2 \frac{H_0 Q}{\alpha} \cot \varphi \sin^3 \varphi
\]

(1.409)

Since this form does not admit of integration in closed form, the following assumption is made:

\[
\begin{align*}
 \Delta(z) &= - \frac{Q H_0}{1 + (1)} \frac{\varphi}{\alpha} \quad \Delta(\varphi) = -2 \frac{H_0 Q}{\alpha} \cot \varphi \sin^3 \varphi
\end{align*}
\]

(1.410)

It is to be noted that in the form chosen, there is an additional weighting of the effect in the region of strong
field. \( Q \) in these expressions is a numerical constant which is indicative of the quality of the lens, with a perfect lens having a \( Q \) of zero.

From the definition of \( \gamma(z) \), the rotation caused by the Bell Shaped field is found to be

\[
(1.11) \quad \gamma = \frac{e}{r \cdot m \cdot \phi_0} \int_{z_0}^{z} H_0(s) \, ds = \frac{e}{r \cdot m \cdot \phi_0} \int_{\phi_0}^{\phi} \sin^2 \left( \frac{\phi}{\phi_0} \right) = -K(\beta - \beta_0)
\]

Substituting these last results, along with the assumed field forms in the "basic" integrals (1.31):

\[
A = K_1 \int_{\phi_0}^{\phi} \sin^2 \omega(\beta - \beta_0) \sin 2\phi \, d\phi
\]

\[
B = -K_1 \int_{\phi_0}^{\phi} \sin \omega(\beta - \beta_0) \cos 2\phi \, d\phi
\]

\[
C = K_2 \int_{\phi_0}^{\phi} \sin \omega(\beta - \beta_0) \sin \omega(\beta - \beta_0) \sin 2\phi \, d\phi
\]

\[
D = -K_2 \int_{\phi_0}^{\phi} \sin \omega(\beta - \beta_0) \cos \omega(\beta - \beta_0) \cos 2\phi \, d\phi
\]

where \( K_1 = \frac{KQ}{\omega} \frac{\sin \beta_0}{\sin \omega(\beta - \beta_0)} \sin \beta_0 \); \( K_2 = \frac{KQ}{\omega} \frac{1}{\sin \omega(\beta - \beta_0)} \)

By use of the trigonometric relations for the sum and difference of two angles: \( \cos(A \pm B) = \cos A \cos B \mp \sin A \sin B \) and \( \sin(A \pm B) = \sin A \cos B \pm \cos A \sin B \), the integrals (1.41) can be brought to a sum of integrals of the forms \( \int \sin(nx + a) \, dx \) or \( \int \cos(nx + a) \, dx \) which can immediately be integrated. The resulting values are:

\[
A = \frac{K_1}{\phi} \left\{ \sin \left[ (1-K)\beta + 2K\phi \right] + \sin \left[ 2(1-K)\beta - 2K\phi \right] \right\} + \frac{\sin \left[ 2(1-K)\omega \beta - 2(1-K)\omega \phi \right]}{2(1-K)\omega} - \frac{\sin \left[ 2(1-K)\omega \beta + 2(1-K)\omega \phi \right]}{2(1-K)\omega}
\]

\[
B = -\frac{K_1}{\phi} \left\{ \sin \left[ (1-K)\beta + 2K\phi \right] + \sin \left[ 2(1-K)\beta - 2K\phi \right] \right\}
\]

\[
C = K_2 \left\{ \sin \left[ 2(1-K)\omega \beta - 2(1-K)\omega \phi \right] - \sin \left[ 2(1-K)\omega \beta + 2(1-K)\omega \phi \right] \right\}
\]

\[
D = -K_2 \left\{ \sin \left[ 2(1-K)\omega \beta - 2(1-K)\omega \phi \right] - \sin \left[ 2(1-K)\omega \beta + 2(1-K)\omega \phi \right] \right\}
\]
By solving the equation giving the location of the aperture plane for the particular case \( \phi = 0 \), that is, with the object at an infinite distance to the right, we obtain the location of the object side focal point as \( \phi_1 = \pi \). In the case of high magnification, the object may be considered to lie at approximately this focal point, and the image at infinity.

In order to determine the focal length of the lens without having first determined the location of the principal planes, use is still made of the properties of these planes. We therefore assume that there is an object
placed at the focal point, and slope of some ray leaving the object is known. Then the effect in image space is that the ray emerges from the lens system parallel to the optic axis at a separation \( f \tan \theta \), where \( f \) is the focal length and \( \theta \) is the angle at which the ray leaves the object. Thus the focal length is given by

\[
(1.414) \quad f = \frac{h(\theta-\omega)}{h'(\theta-\omega)} = \frac{\tan \omega}{\tan \theta} = \frac{a}{\tan \theta}
\]

If the accelerating potential is assumed to be 50\( \text{kV} \), the maximum magnetic field 5000 Gauss, and the half value width of the field 1 mm, the following numerical values are obtained: \( f = 3.6 \text{ mm} \), \( K = .667 \), \( \omega = 1.2 \), \( \theta_f = 150^\circ \), \( \theta_a = 34.8^\circ \), \( K_1 = .954 \), \( K_2 = .835 \), \( A = .430 \), \( B = -.446 \).

by further assuming that the object side aperture is .005 radians, the radius of the circle of confusion referred to object space becomes \( 1.1 \times 10^{-3} \text{mm} \). Thus a machining error of 23 parts per 100,000 would, according to the above assumptions lead to a 50\( \AA \) diameter circle of confusion.

As has been indicated earlier, the source of the asymmetry may be either machine errors or inhomogeneities of the pole piece material. Since the former effects may be reduced to a minimum by lapping and polishing techniques, the asymmetries can usually be attributed to the latter defect. Since a direct measure of the inhomogeneities is difficult, the treatment above has numerical meaning only in indicating the order of magnitude of resulting lens defects.
II EVALUATION OF THE RESOLVING POWER OF AN ASYMMETRIC LENS

The study of the diffraction pattern of asymmetric lenses has been limited previously to one of two cases: the case in which the amount of the aberration and the aperture is small enough so the path differences encountered are much less than one wavelength of the illumination used, and for rectangular apertures only without this restriction.

1. Determination of the Wave Surface Corresponding to a Converging Asymmetric Wave

In order to determine the aberration pattern due to the presence of primary astigmatism from the viewpoint of wave optics, it is first necessary to determine a surface of constant phase that corresponds to this case. Since the image space is assumed to be field free, the rays as determined in the geometric considerations must be normal to the wave surfaces.

To develop the surface in question, the origin of coordinates is assumed to lie in the plane in which the pattern is to be observed. The z axis lies along the optic axis, and the line foci are parallel to the x and y axes at separations $D_1$ and $D_2$ from the observation plane. (See figure 5.) Again, since these planes lie in field free space, the rays are straight lines. Therefore, the equations of the projection of a ray (intersecting
Figure 5. Geometry and notation used in developing the asymmetric wave surface.

the observation plane at \( x_0, y_0 \) in the X-Z and Y-Z planes can be written:

\[
(2.101) \quad x = \frac{x_s}{D_1} (D_2 - z), \quad y = \frac{y_s}{D_2} (D_2 - z)
\]

If these equations are rewritten in terms of the coordinates of some other point of the ray \( (x_s, y_s, z_s) \), then

\[
x = x_s \left( \frac{D_2 - z}{D_1 - z_s} \right), \quad y = y_s \left( \frac{D_2 - z}{D_2 - z_s} \right)
\]

or these equations may be rewritten as

\[
(2.102) \quad \frac{x - x_s}{D_1 - z_s} = \frac{y - y_s}{D_2 - z_s} = \frac{z - z_s}{D_2 - z_s}
\]

The equation of the normal of a surface \( F(x, y, z) = \text{const.} \) at the point \( (x_s, y_s, z_s) \) on the surface is given by:

\[
(2.103) \quad \frac{\partial F}{\partial x} \bigg|_{x_s, y_s, z_s} = \frac{\partial F}{\partial y} \bigg|_{x_s, y_s, z_s} = \frac{\partial F}{\partial z} \bigg|_{x_s, y_s, z_s}
\]

Thus, if the rays are to form the normals of this surface,

\[
(2.104) \quad \frac{\partial F}{\partial x} \bigg|_{x_s, y_s, z_s} = -\frac{x_s}{D_1 - z_s} f(x_s, y_s, z_s), \quad \frac{\partial F}{\partial y} \bigg|_{x_s, y_s, z_s} = -\frac{y_s}{D_2 - z_s} f(x_s, y_s, z_s), \quad \frac{\partial F}{\partial z} \bigg|_{x_s, y_s, z_s} = f(x_s, y_s, z_s)
\]
If the location of the intersection of the surface to be constructed and the z axis is restricted to being at a distance much greater than either $D_1$ or $D_2$, and if further, the maximum values of $x$ and $y$ to be considered are also much less than this distance, $-R$, it may be assumed with sufficient accuracy that the denominators of the first two partial derivatives are essentially constant and equal to $D_1 - R$ and $D_2 - R$ respectively. With this assumption, equations (2.104) can be integrated immediately to give the equation of the approximate surface as:

$$F(x, y, z) = \frac{x^2}{D_1 + R} + \frac{y^2}{D_2 + R} - 2z = 2R$$

where $z = -R$ is the intersection of the surface with the optic axis.

While the use of cartesian coordinates allows a quick development of the surface, in what follows there will be need for the separation of the surface from the origin of the observation plane in terms of the angles $\theta$ and $\phi$, where $\theta$ is the angle formed by the ray and the optic axis, and $\phi$ is the angle formed by the $X-Z$ plane and the plane containing the ray and the optic axis. By writing the direction cosines of the line joining the origin with a point $(x, y, z)$ on the surface equal to the direction cosines of the normal to the surface at that point:

$$\alpha = \sin \theta \cos \phi = \frac{x}{\sqrt{R^2 + k}} ; \beta = \sin \theta \sin \phi = \frac{y}{\sqrt{R^2 + k}} ; \gamma = \cos \theta = \frac{z}{\sqrt{R^2 + k}}$$
These equations are valid only if the surface is an infinite distance from the observation plane. Expressions for \( x \) and \( y \) in terms of \( \Theta \) and \( \Phi \) can be determined immediately from these equations. If these results are then substituted in equation (2.105) \( z \) is similarly given as a function of \( \Theta \) and \( \Phi \).

\[
(2.107) \quad x = -(R+D) \tan \Theta \cos \Phi; \quad y = -(R+D) \tan \Theta \sin \Phi; \quad z = -R(1 + \frac{1}{2} \tan^2 \Theta + \sec \Theta \tan \Phi + 2 \sin \Phi)
\]

Finally, if the expressions (2.107) are substituted in the equation for the separation of \((x, y, z)\) from the origin,

\[
r = x \sin \Theta \cos \Phi + y \sin \Theta \sin \Phi + z \cos \Theta,
\]

\[
(2.108) \quad r(\Theta, \Phi) = -R - \frac{1}{2} \sin^2 \Theta \left[D_1 \cos^2 \Phi + D_2 \sin^2 \Phi\right] + O_4
\]

In (2.108), \( O_4 \) represents fourth and higher order powers of \( \sin \Theta \).

2. The Evaluation of the Kirchhoff Solution of the Wave Equation for the Case of a Converging Asymmetric Wave.

The intensity distribution which is established by any electron lens must be determined from a solution of the Schrödinger equation,

\[
(2.201) \quad \nabla^2 \psi + \frac{\hbar^2}{2m} (E - V(x, y, z)) \psi = 0
\]

However, since the image plane lies in field free space, the potential function must be a constant, and the equation reduces to the wave equation in a region of constant
index of refraction. In this case, the wave length is the De Broglie wavelength. The Schrödinger equation reduces to the form:

\[(2.202) \nabla^2 \psi + k^2 \psi = 0 \quad \text{where} \quad k = \frac{2\pi}{\lambda} \quad \text{and} \quad \lambda = \frac{k}{\sqrt{2meq}}\]

The Kirchhoff solution for the amplitude function at a point \((x,y,z)\) in terms of the known function \(u\) on the bounding surface \(S\) is:

\[
\psi = -\frac{1}{4\pi} \int \left\{ \nabla^2 \left( \frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial u}{\partial n} \right\} dS
\]

where \(r\) is the distance from the surface element \(dS\) to the point \((x,y,z)\). For the purpose of this problem, it is possible to synthesize the function \(u\) on an essentially spherical surface of very large radius about the origin of the image plane. To construct the surface, it is merely necessary to add the same radial distance \(R\) at every point of the original surface, as illustrated in Figure 6.

It is apparent that the resulting surface will differ imperceptibly from a sphere if \(R\) is very great, but that the distribution in phase or path length as a function of the angles is still preserved. If this new surface is taken as the surface of integration, then \(U\) must be of the form obtained with a spherical wave: \(U = \frac{e^{ikr_0}}{r_0}\), and

\[(2.203) \psi = -\frac{1}{4\pi} \int \left\{ \frac{e^{ikr_0}}{r_0} \frac{\partial}{\partial n} \left( \frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial}{\partial n} \left( \frac{e^{ikr_0}}{r_0} \right) \right\} dS\]

In this expression, \(r_0\) is the separation of the element of integration on the surface from the origin of the image plane, while \(r\) is the separation of the point of observa-
Figure 6. Synthesis of essentially spherical wave surface

If it is assumed that the dimensions of the surface are small compared to the separation of the surface from the image plane, and likewise that the separation of the point of observation from the origin of the image plane is small, then the cosines can be replaced by -1 and 1 respectively. The variations in $r$ and $r_0$ over the surface will be small compared to $r$ and $r_0$ measured to the intersection of the surface and the optic axis, but because of the large value of $k$, the exponential is a rapidly varying function. Therefore, while it is improper to neglect variations in the $r$'s in the exponential, the product $rr_0$ appearing in the denominator may be removed from under the integral as being essentially a constant. Finally, the terms
l/r and l/r₀ are neglected with respect to the corresponding terms ik because of the large values of k, r, and r₀. With those assumptions, the amplitude function becomes:

\[
\psi = \frac{-ik}{2\pi} \frac{1}{r \cdot r₀} \int e^{i(kN+\alpha)} dS
\]

The separation of the surface from the origin of the observation plane, r₀, has been determined previously as a function of θ and φ. If the observation point is specified in terms of its polar coordinates ρ and γ in the observation plane, the separation of this point and a point (R, θ, φ) on the wave surface with very large R is

\[
r = R + \rho \sin \Theta \cos(\phi - \gamma)
\]

Substitution of the values of r and r₀ leads to:

\[
\psi = \text{const.} \int \int e^{i\left( \frac{k}{\rho^2} \rho \cos \gamma + \frac{\rho}{\rho^2} \sin \gamma \right)} e^{i \rho \sin \Theta \cos(\phi - \gamma)} \sin \Theta \cos d\rho d\gamma
\]

The first step in the evaluation of this integral is the expansion of the first exponential. This leads to the sum of integrals:

\[
\psi = \text{const.} \int \int \left[ \frac{(k \sin \Theta)^{2n}}{n!} \left( \frac{\rho}{\rho^2} \cos \gamma + \frac{\rho}{\rho^2} \sin \gamma \right) \right] e^{i \rho \sin \Theta \cos(\phi - \gamma)} \sin \Theta \cos d\rho d\gamma
\]

To proceed further, use is made of equations concerning the Bessel functions:

\[
J_n(\alpha) = \frac{1}{2\pi} \int e^{i \alpha \cos \phi} d\phi;
\]

\[
J_n(\alpha) = \frac{1}{2\pi} \int e^{i \alpha \cos \phi} d\phi
\]

It is convenient to use the ordinary diffraction integral

\[
F = \int \int e^{i \rho \sin \Theta \cos(\phi - \gamma) - \alpha \rho \cos \gamma} \sin \Theta \cos \theta d\rho d\gamma = \text{const.} \frac{J_n(\alpha \rho \cos \gamma)}{\alpha \rho}
\]
If now \( \cos(\psi - \tau) \) is rewritten as \( \cos \phi \cos \tau - \sin \phi \sin \tau \), and the definitions of the rectangular coordinates in terms of the polar coordinates in the observation plane are used, \( x_1 = \phi \cos \tau, \ y_1 = \phi \sin \tau \), \( F \) assumes the form:

\[
F = \int e^{i (k a \sin \phi (x_1 \cos \phi + y_1 \sin \phi))} \ \sin \phi \cos \phi \ d\phi
\]  

(2.210)

This last expression may be differentiated partially with respect to either \( x_1 \) or \( y_1 \) with the result:

\[
\frac{\partial^{n+m} F}{\partial x_1^n \partial y_1^m} = \int e^{i (k a \sin \phi)^{n+m} (\cos \phi)^n (\sin \phi)^m} e^{i k a \phi \cos \phi} \ \sin \phi \cos \phi \ d\phi
\]  

(2.211)

This is, however, the form of the integrals in the sum of integrals to be evaluated. By substitution of these results, the amplitude function is given as a sum of partial derivatives of a known function. Thus:

\[
\psi = \text{const} \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{i}{k} \right)^n \frac{\partial^n}{\partial x_1^n} \left[ \frac{D_1}{2} \frac{\partial^2}{\partial x_1^2} + \frac{D_2}{2} \frac{\partial^2}{\partial y_1^2} \right] \left( \frac{a \mu k}{\sqrt{x_1^2 + y_1^2}} \right) \left( \frac{a \mu k}{\sqrt{x_1^2 + y_1^2}} \right)
\]  

(2.212)

For the particular case \( D_1 = -D_2 \), this last form can be further simplified by making use of differential relations between Bessel Functions of successive orders, expansion, and recombination of terms. These manipulations are carried out in Appendix B. The final result for the case \( D_1 = -D_2 = 2D \) gives \( \psi \) as a sum of known functions:

\[
\psi = \text{const} \sum_{n=0}^{\infty} \sum_{d=0}^{\infty} \frac{1}{(n+d)!} \left( \frac{1}{n!} \right) \sum_{k=0}^{\infty} \frac{(\gamma)}{(n-d)!(n-2d)!} \sum_{s=0}^{\infty} \frac{(s)}{(n-4-s)!} \frac{J_{n+2-s} \left( \frac{a \mu k}{\sqrt{x_1^2 + y_1^2}} \right)}{s + \kappa + 1}
\]  

(2.213)

where \( s = \kappa \rho \).
This result then indicates that the form of the image formed by an astigmatic lens is determined by the product \( Dk \alpha^2 = C \) and that for a given value of \( C \), the separation of the focal lines is directly proportional to the wavelength of the illumination, and inversely proportional to the square of the aperture. Since \( \sin \alpha \approx \alpha \), \( 2C \) is equal to the range of phase angles (in radians) encountered in the course of the integration. Further, it can be shown that with a given lens and small apertures, the value of \( C \) is essentially independent of the location of the object and corresponding image planes as long as the separation of the object from the object side focal point is large compared to the difference in the astigmatic focal lengths.

Consider a thin astigmatic lens with focal lengths \( f_1 \) and \( f_2 \) with an object at a distance \( p \) to the left. Then the focal lines lie at \( q_1 = \frac{p}{f_1} \) and \( q_2 = \frac{p}{f_2} \). If the radius of the limiting diaphragm is \( h \), then \( \alpha = \frac{2h}{q_1 q_2} \) and

\[
C = \frac{k_1^2 \left[ q_1 q_2 \left[ p \left( f_1 + f_2 \right) - 2f_1 f_2 \right]^2 \right]}{f_1 f_2 \left( p \left( f_1 + f_2 \right) - 2f_1 f_2 \right)^2} \left\{ 1 - \frac{(f_2 - f_1)^2 \left[ p \left( f_1 + f_2 \right) - 2f_1 f_2 \right]}{f_1 f_2 \left( p \left( f_1 + f_2 \right) - 2f_1 f_2 \right)} \right\}
\]

From this last expression, it can be seen that if both focal lines lie in real image space, the value of \( C \) decreases with decreasing values of \( p \) (higher magnification). If the object is more than ten times the difference in focal lengths away from the shorter focal point, the variation in \( C \) from the value obtained neglecting the second term is \( \frac{1}{2} \leq \frac{C}{C} \leq 3.7 \) with \( |f_2 - f_1| \approx f_1. \)
In the electron microscope, \( f_2 - f_1 \leq 10 \) microns, 
\( f = 0.35 \) cm so \( C \) is constant to within 3\% for magnifications up to 120.

3. Numerical Results of the Theory

The numerical evaluation of the integral was carried out for \( C = 2 \) and \( C = 3 \). The results of these calculations are displayed in the following tables and graphs. Figures 7 and 8 show the relative intensity as a function of separation from the origin for various angles \( 2\theta \). It can be seen that there is no longer a sharply defined ring of zero intensity followed by a secondary maximum as is the case for a perfect lens. Instead, in the directions corresponding to those of the focal lines, a plateau is formed, while at an angle of \( 45^\circ \) to these directions the intensity falls off more rapidly. This resultant star shaped pattern is shown more clearly in figures 9 and 10 which give the equi-intensity contours for the cases \( C = 2 \) and \( C = 3 \). In the case of \( C = 3 \), the point at which the intensity is half of the central maximum value is at nearly twice the radial distance from the pattern center in the directions parallel to the focal lines as in the planes at \( 45^\circ \) to them.

Higher values of \( C \) were not considered because of the slowness of convergence of the series.
### Table I

**Relative Intensity as a Function of \( S \) and \( 2\tau \)**

**For \( C = 2 \)**

<table>
<thead>
<tr>
<th>( 2\tau )</th>
<th>( 0^\circ )</th>
<th>( 15^\circ )</th>
<th>( 30^\circ )</th>
<th>( 45^\circ )</th>
<th>( 60^\circ )</th>
<th>( 75^\circ )</th>
<th>( 90^\circ )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0</strong></td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>.820</td>
<td>.820</td>
<td>.819</td>
<td>.817</td>
<td>.817</td>
<td>.816</td>
<td>.816</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>.483</td>
<td>.480</td>
<td>.471</td>
<td>.458</td>
<td>.445</td>
<td>.435</td>
<td>.432</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>.284</td>
<td>.274</td>
<td>.246</td>
<td>.210</td>
<td>.174</td>
<td>.119</td>
<td>.139</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>.217</td>
<td>.203</td>
<td>.166</td>
<td>.065</td>
<td>.031</td>
<td>.010</td>
<td>.024</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>.141</td>
<td>.131</td>
<td>.102</td>
<td>.065</td>
<td>.031</td>
<td>.010</td>
<td>.024</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>.069</td>
<td>.061</td>
<td>.043</td>
<td>.023</td>
<td>.009</td>
<td>.003</td>
<td>.002</td>
</tr>
<tr>
<td><strong>7</strong></td>
<td>.032</td>
<td>.025</td>
<td>.013</td>
<td>.003</td>
<td>.000</td>
<td>.003</td>
<td>.006</td>
</tr>
<tr>
<td><strong>8</strong></td>
<td>.030</td>
<td>.025</td>
<td>.014</td>
<td>.005</td>
<td>.003</td>
<td>.004</td>
<td>.005</td>
</tr>
</tbody>
</table>

**For \( C = 3 \)**

<table>
<thead>
<tr>
<th>( 2\tau )</th>
<th>( 0^\circ )</th>
<th>( 15^\circ )</th>
<th>( 30^\circ )</th>
<th>( 45^\circ )</th>
<th>( 60^\circ )</th>
<th>( 75^\circ )</th>
<th>( 90^\circ )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0</strong></td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>.899</td>
<td>.897</td>
<td>.895</td>
<td>.893</td>
<td>.890</td>
<td>.890</td>
<td>.888</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>.742</td>
<td>.735</td>
<td>.712</td>
<td>.684</td>
<td>.656</td>
<td>.637</td>
<td>.630</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>.680</td>
<td>.660</td>
<td>.602</td>
<td>.527</td>
<td>.452</td>
<td>.398</td>
<td>.379</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>.609</td>
<td>.579</td>
<td>.501</td>
<td>.398</td>
<td>.305</td>
<td>.241</td>
<td>.221</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>.398</td>
<td>.368</td>
<td>.295</td>
<td>.230</td>
<td>.170</td>
<td>.142</td>
<td>.136</td>
</tr>
<tr>
<td><strong>6</strong></td>
<td>.224</td>
<td>.196</td>
<td>.138</td>
<td>.077</td>
<td>.060</td>
<td>.075</td>
<td>.081</td>
</tr>
<tr>
<td><strong>7</strong></td>
<td>.144</td>
<td>.118</td>
<td>.060</td>
<td>.017</td>
<td>.013</td>
<td>.034</td>
<td>.047</td>
</tr>
<tr>
<td><strong>8</strong></td>
<td>.105</td>
<td>.079</td>
<td>.034</td>
<td>.006</td>
<td>.002</td>
<td>.011</td>
<td>.015</td>
</tr>
<tr>
<td><strong>9</strong></td>
<td>.112</td>
<td>.080</td>
<td>.024</td>
<td>.000</td>
<td>.002</td>
<td>.002</td>
<td>.000</td>
</tr>
</tbody>
</table>

In an effort to estimate the resolving power of an astigmatic lens, the intensity pattern for two self-luminous points separated by the distance corresponding to \( s = \frac{1}{4} \) was computed for two cases: first the points lie along a focal line, and second, the points lie on a line at \( 45^\circ \) to a focal line. The resulting equi-intensity contours
Figure 7. Relative intensity as a function of separation from the origin for $C = 2$.  

Graph 1: $C = 2 F_2 - F_1 = 0.32$
Figure 8. Relative intensity as a function of separation from the origin for C 3.
Figure 9.
Equi-intensity contours for 
$C = 2$. Dotted circle represents the 
geometric optic circle of confusion. ($s = 3$)

Figure 10.
Equi-intensity contours for 
$C = 3$. Dotted circle represents the 
geometric optic circle of confusion. ($s = 3$)
for these cases with $C = 2$ are shown in figures 5 and 6.

From these contours it can be seen that in both cases there is a relative intensity minimum of about 80% between the two intensity peaks. This corresponds well to the relative minimum of 75% that is obtained with a perfect lens. On this basis, an astigmatic lens may be considered to be sufficiently corrected when its $C$ value has been reduced to 2 or less. In the case of a typical electron microscope objective with an aperture of 0.005 radians and \( k = 1.17 \times 10^{10} \) (corresponding to an accelerating potential of about 50,000 volts) the permissible variation in focal lengths is .265 microns. This is more than twice the allowable limit considering the Rayleigh quarter wavelength criteria. (This

Figure 11. Equi-intensity contours for two self luminous points separated by \( s = \frac{1}{4} \) (6.8\% for typical microscope lens) with the points lying along a focal line.)
Figure 12. Equi-intensity contours for two self luminous points separated by \( s = 4 \) (6.8\( \mu \text{m} \) for typical microscope lens) with the points lying on a line at 45° to a focal line.

Very broad criteria states that an optical instrument will not fall seriously short of the performance possible with an absolutely perfect system if the difference between the longest and shortest optical paths leading to a selected focus do not exceed one quarter of a wavelength. The corresponding geometric circle of confusion is 13.6\( \mu \text{m} \) in diameter with \( \sigma f = .265 \) microns, \( \alpha = 5\times10^{-3} \).

Thus it can be seen that the wave optical treatment leads to less stringent requirements for correction of a lens than either the Rayleigh criteria or a geometric optic treatment.
4. **Light Optic Experimental Verification**

In order to obtain an experimental verification and extension of the calculations, we consider the like problem in light optics. The advantage of this digression is the ease with which a suitable point source of illumination can be formed. Because of the much longer wavelength of light as compared to the effective wavelength of the electron beam, much larger variations in the path lengths are needed to provide the same total phase variation, or value of \( C \). This means that either larger apertures or larger separations of the focal lines are required in the optical analogue than exist in the electron microscope.

In setting up the optical counterpart, it was necessary first to determine to what extent other aberrations were present with the lenses used. Chromatic aberration was eliminated by using a monochromatic light source. Similarly all of the off axis aberrations may be neglected with proper alignment, since the calculations whose validity we wish to demonstrate are made for an axial object and image. The only remaining aberration in addition to the desired axial astigmatism, namely spherical aberration, is shown to be negligible in Appendix C.

The test bench was fitted with a sodium vapor lamp fitted with a metal piece which together with the lamp's own outer shield allowed no light to be emitted except through a fine pin hole. The test lenses, a four diopter
spherical lens in contact with various cylindrical lenses, are those from an optometrist's test set. The lenses were mounted in a special holder which also had provision for holding various apertures. The apertures were made by drilling holes with numbered twist drills in 0.020" tin plated steel. A second holder was used to hold the correcting cylindrical lens for the demonstration of the effectiveness of correction. Finally, one of two special mounts were used for the observation and recording of the resultant image. The first provided for the mounting of a cartridge of 35mm film on the front surface, and had a mounting tube for a 10x microscope eyepiece on the rear. The second supported a complete 35mm camera. An auxiliary plate provided for the proper mounting of the eyepiece to determine the focus with the camera back removed.

The patterns obtained for $C = 1, 2, 3$ are shown in figure 13. As can be seen, these patterns correspond with the star shaped patterns predicted theoretically. Figure 14 shows the patterns that are obtained for larger asymmetries. If referred to the electron microscope objective, the patterns represent differences in focal lengths of $0.18\mu m$ to $1.84\mu m$ in steps of $0.18\mu m$. In terms of $C$, the variation is from 1.35 to 13.5 in steps of 1.35. As can be seen, these larger values of $C$ lead to patterns having the form of a distorted checkerboard with the number of squares on the board increasing with increasing $C$. This
Figure 13. Diffraction patterns obtained for slightly astigmatic lenses. From left to right the $C$ values are 1.2, and 3. The corresponding differences in focal lengths for a typical microscope objective are 0.137, 0.274, and 0.411 microns.

latter form of pattern exists for very large values of $C$ as can be seen in figure 15 where $C = 29$ in the left hand pattern. These higher values of $C$ were not considered theoretically because of the slowness of convergence of the series and the difficulties of having the numerical answer very much less than the early terms of the series, that is a small difference between large numbers.

Figure 14. Diffraction patterns obtained with a range of values of astigmatism $C = 1.35$ to $C = 13.5$ in steps of 1.35
Figure 15. The effect of correction of an asymmetric lens with a cylindrical lens. In the pattern on the left $C = 29$.

Finally figure 15 shows the effect of correcting a very bad lens, $C = 29$, by the introduction of a cylindrical correcting lens. In the right hand pattern, the ellipticity is the result of the effectively elliptical aperture as viewed from the image plane through the correcting lens.
III PROPOSED ELECTRIC CYLINDRICAL LENS FOR CORRECTION OF OBJECTIVE LENS ASYMMETRY

1. Magnetic Correction of Objective Lens Asymmetry

The elimination of the pole piece asymmetry by the introduction of an additional magnetic field has been considered in some detail by Hillier. Since the results of this work have been made available commercially, and the microscope that has been used in the present work makes use of this type of correction, its principle and operation are reviewed.

It is assumed that the asymmetric lens can be represented by the superposition of a cylindrical magnetic lens on an accurately axially symmetric lens. The correction therefore consists of adding a second cylindrical lens of proper strength at right angles to the existing cylindrical component. This is accomplished by means of eight soft iron screws inserted in the brass spacer separating the pole pieces.

Compensation of the lens consists of the adjustment of various pairs of screws until no discernable asymmetry remains in the Fresnel fringes formed about the image of a hole in a thin film. (See figure 21.) Each of the screw adjustments require that the microscope column be opened and the lens removed, and hence the process is rather lengthy. Fortunately, the resulting correction is rather stable, and after adjustment, the lens will remain corrected.
for possibly several weeks of operation by a single operator. Of course any asymmetries introduced by the specimen remain uncorrected.

Because of the length of time which is required for correction when this method is used, it seemed desirable to have a rapid auxiliary method, preferably one that could be controlled without the necessity of opening the column. Since it has been shown in the preceding sections that the magnetic objective still leads to line foci, this correction need not be in the form of another magnetic lens. Further, the location of the correcting lens is more or less arbitrary.

From the standpoint of mechanical convenience, this lens is most easily placed in the region between the objective and projector lens which is provided with a demountable bellows arrangement. Since at this location there is mechanical freedom of rotation, the lens may be a simple cylinder lens whose orientation with respect to the cylindrical component of the asymmetric objective lens can be determined mechanically.

2. Paraxial Electron Optics of the Correcting System

The auxiliary correction of the astigmatism of the microscope objective is accomplished by the use of an electrostatic cylindrical lens. This lens is composed of four electrodes with the outer two at the potential of the column and the correcting voltage applied to the center electrodes.
The electrodes are in the form of plane parallel pieces of brass with narrow slits in them. Such a configuration has essentially no effect on the imaging in the plane parallel to the slits, and a converging effect determined by the applied potential in the plane perpendicular to the slits. The axial potential of this configuration of electrodes was determined experimentally in an electrolytic plotting tank. As can be seen from figure 16, the axial potential can be quite closely approximated by a trapezoid.

For the two dimensional electrostatic field,

\[ F(x, y) = \sqrt{\varepsilon \Phi} = \sqrt{\varepsilon \Phi}(1 + \frac{1}{2}(x'^2 + y'^2) - \frac{\Phi'}{\Phi} x') \]

and the Euler-Lagrange equations become:

\[
\left(3.201\right) \frac{d}{ds} \left(\sqrt{\varepsilon \Phi} y'\right) = 0; \quad \frac{d}{ds} \left(\sqrt{\varepsilon \Phi} x'\right) + \frac{\Phi''}{2\varepsilon \Phi} x = 0 \quad \text{on} \quad x'' + \frac{\Phi'}{\varepsilon \Phi} x' + \frac{\Phi''}{2\varepsilon \Phi} x = 0
\]

The integration of these equations is carried out in a number of successive steps. In the case of the trace in the YZ plane, the integration is carried out as the sum of integrals corresponding to the three regions of constant field. In the case of the XZ trace, the change in slope must also be determined at the four points of discontinuity. In the latter case the integration takes the form of a step-by-step process.

The equation in x in the case of field free space reduces to:

\[
\left(3.202\right) x'' = 0 \quad \text{hence} \quad x' = x' \quad \text{and} \quad x = x + x'(z - z_o)
\]
Figure 16. Axial potential of the four slit lens. Only half of the symmetric plot is shown. The dotted lines represent the approximate field.

The equation for $x$ in a constant field reduces to

$$x'' = -\frac{(\phi - \phi_0)x'}{2L\phi_0 + 2(\phi_0 - \phi_1)(x - x_0)}$$

where $\phi_0$ is the potential at the beginning of the region, and $\phi_1$ is the potential at the end. The region is of width $L$. Making use of the linear form of the potential function:
Finally there is to be considered the change of slope at a break point. At such a point, the potential and its first derivative are finite, while the second derivative becomes infinite. If the integration is carried out over a small increment about the break point, and this increment is allowed to approach zero, the contribution of the second term may be neglected with respect to the last. Further, in the latter, x and φ may be considered as a constant. Therefore, at a breakpoint:

\[ x_0' = x_0 + \frac{2x_0' \sqrt{\phi_0}}{\sqrt{\phi_0} + \sqrt{\phi_0}} \]

By using these equations in turn, the trajectory may be traced. Assuming that \( x' = 0 \) to the left of the lens, and that \( x = 1 \) in this region, the values of \( x \) and \( x' \) at the breakpoints \( z_a, z_b, z_c, \) and \( z_d \) are summarized in Table II. In this table, use is made of the abbreviation \( \alpha = \sqrt{1 + \frac{1}{\phi}} \). \("
The focal length of the lens is given by the reciprocal of $x_d$. The location of the image side principal plane from $z_d$ is given by $\frac{1-x_d}{x_d}$.

If it is now assumed that $\frac{V}{D}$ is small compared to one, the quantity $\alpha$ may be expanded to give:

\[
(3.206) \quad x_d = 1 - \left(\frac{V}{D}\right)^2 \left(1 + \frac{1}{\frac{E}{D}}\right) + \left(\frac{V}{D}\right)^2 \left(1 + \frac{3}{\frac{E}{D}}\right) - \left(\frac{V}{D}\right)^4 \left(\frac{15}{E^2} + \frac{3D}{E^2}\right) + \cdots
\]

\[
x_d' = -\frac{1}{d} \left(\frac{V}{D}\right)^2 \left[1 + \frac{V}{4D} - \left(\frac{1}{4} - \frac{D}{E^2}\right)(\frac{V}{D})^2 + \cdots\right]
\]

From this, it can be seen that to the lowest order the focal length is:

\[
(3.207) \quad f = d \left(\frac{\phi}{V}\right)^2
\]

and the principal plane is located at $D_1 = \frac{V}{\phi} \left(\frac{5d + 2D}{4}\right)$ to the right of the center of the lens. Since the lens is
symmetrical, the separation of the principal planes is 
\[ \frac{V}{\phi} (\frac{D}{2} + D) \]

Integration of the paraxial equation for the trace in the YZ plane is \[ y'(\phi) = y_0'(\phi) \quad \text{or by substituting the proper potential function for each of the three regions:} \]

\[ (3.208) \quad y_d - y_a = y_0' \sqrt{\phi} \left( \int \frac{dz}{\sqrt{\phi + z^2}} + \int \frac{d\phi}{\sqrt{\phi + \phi^2}} + \int \frac{d\phi}{\sqrt{\phi + \phi^2 - \frac{1}{2} (s - D^2)}} \right) \]

and carrying out the integrations

\[ y_d - y_a = y_0' \left\{ \frac{D}{1 + \sqrt{\phi}} + \frac{\phi_d (\sqrt{1 + \frac{1}{\phi}} - 1)}{V} \right\} = y_0' (D + 2d) - \frac{V}{\phi} y_0' \left( \frac{D}{2} + d \right) + \ldots \]

This corresponds to a shift of the image by an amount

\[ \Delta^2 = \frac{(y_d - y_a) - y_0'(D + 2d)}{y_0'} = \frac{V}{\phi} (\frac{D + d}{2}) \]

Both the shift of the y focal line and the separation of the principal planes are small compared to the other dimensions in image space. Even with \( (V/\phi) = 10^{-1} \), the values are less than 1 mm, or referred to object space, the shift is less than .1 micron. Therefore, in estimating the effect of the correcting lens, these quantities will be neglected, and the lens considered to be thin.

3. **Geometric Optics of the Correcting System**

Under these conditions, the total objective system may be considered to consist of an astigmatic spherical lens with focal lengths \( f_1 \) and \( f \) in the \((y,z)\) and \((x,z)\) planes respectively, and a cylindrical lens with focal
Figure 17. Geometry and notations used in developing the geometric optics of the correcting system.

length $F$ in the $(x,z)$ plane and infinite focal length in the $(y,z)$ plane. (See figure 17.)

If use is made of the simple lens formula $\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$ where $f$ is the focal length of the lens, $p$ is the location of the object to the left of the lens, and $q$ is the location of the image to the right of the lens, the image in the YZ plane is formed at:

$$q = \frac{pf}{p-f}$$

(3.301)

In the XZ plane, the image formed by the objective lens alone lies at:

$$q_o = \frac{pf}{p-f} = \frac{p(f+s)}{p-f-s} \text{ where } f_o = f+s$$

(3.302)
If the correcting lens lies at a distance $S$ from the objective, the object for this lens may be considered to lie at $P_2 = S - q_1$. The final image formed by the system in the $XZ$ plane therefore lies at a distance

$$q_2 = \frac{F}{p_2 - F} = \frac{S_p - (s+p)(f+s)}{(s-F)p - (s+p-F)(f+s)} F$$

to the right of the correcting lens. In order to remove the primary astigmatism, it is thus necessary that the strength of the cylindrical lens be such that $q = q_2 + S$.

Substituting the values of $q$ and $q_2$ determined above, the required focal length is found to be given by:

$$F = \frac{[s(p-f) - pf][p - s(p-f)]}{p^2(f-ft)} = \frac{[p - s(p-f)]^2 + (s+p)[p - s(p-f)]}{p^2}$$

If the magnification in the $YZ$ plane is introduced ($M = q/p$; $p = \frac{W}{M}$) and $S$ is expressed as $v/f$, then $F$ assumes the form:

$$F = \left[\frac{M+1-v}{M+1}\right]^2 \frac{1}{f} + \frac{[M+1-v][M+1+M]}{(M+1)^2} f$$

For the values $M = 80$, $f = 35$ cm, $v = 30$, $\delta = 1$ micron, the error in neglecting the second term is less than $1\frac{1}{2}$%. With this simplification, the required ratio of lens voltage to accelerating potential for correction of a difference $\delta$ between focal lengths becomes: (See (3.206))

$$\frac{V}{\phi} = \left[\frac{M+1}{M+1-\nu}\right]^2 \frac{\delta}{f} \sqrt{\delta} = 0.015 \sqrt{\delta} \quad \text{in microns}$$
4. Physical Construction of the Correcting Lens

The physical construction of the correcting lens is illustrated schematically in figure 18 and in section to scale in figure 19. The lens consists of 4 electrodes, the outer two being at the potential of the microscope column and the center two being at a potential determined by the setting of the potentiometer across the power supply. All parts of the assembly are of brass except for the spacers separating the central electrodes from the outer electrodes, the insulating ring which separates the
central electrodes from the lens carriage, and the two bushings through which electrical contact is obtained between the center electrodes and the potentiometer. These parts, except for the inner bushing which is a glass seal, are of polystyrene. The electrodes and their associated spacers are mounted rigidly in the lens carriage which is permitted no motion except one perpendicular to the slits (for centering) with respect to the mounting shell. The mounting shell fits over the projection of the objective pole piece coil assembly of the microscope and may be rotated about it. Thus a rotation of the mounting ring results in a rotation of the slits.

The power supply (figure 20) is a simple full wave voltage doubler type and is capable of delivering approximately 3000 volts to the voltage divider with less than .1 volt ripple.
Figure 19. Sectional drawing of the correcting electrostatic lens. Twice actual size.
Figure 20. High voltage power supply. Capacities in microfarads, unmarked resistors 820 k.
5. **Experimental Results**

The electron microscope that was used for these experiments was calibrated as to variation of focal length with setting of the fine focus potentiometer. This calibration was made by measuring the thickness of a specimen screen, and then applying a collodion film to both sides of the screen. Since these collodion films inherently have a great many small holes in them, these holes were used as objects. The lens current was then varied and the difference in the potentiometer setting to obtain a sharp focus of a hole in the upper film to that required for focus of a hole in the lower film was noted. Since the separation of the films had been determined, this provides a direct calibration of the change in focal length with potentiometer setting. The result of this calibration is that approximately 2 divisions change in the setting of the potentiometer corresponds to a change in focal length of 1 micron. The accuracy of this calibration is reduced because of the very thin specimen screens necessary in order to avoid having to reset one of the coarser controls. The screens used were approximately 5 microns thick. The direct measurement was made using a light microscope.

A partially corrected objective lens was then used to form an image of a hole in a single colodion film, and the difference in settings of the fine focus control for sharp focus in the two perpendicular directions was noted.
A large potential was then applied to the center electrodes of the correcting lens, and the shift in the Fresnel fringes was noted. The lens was then rotated until the fringes occurred in the opposite sense to those of the uncorrected lens. That is, if the uncorrected lens showed fringes on the film side of the hole along a horizontal edge with the vertical edge in focus, the lens was rotated to cause the fringes to be formed along the vertical edge. The potential applied to the lens is then reduced until the fringes become symmetrical, and the whole hole is focused for the same lens current. Figure 21 shows the development of the asymmetric fringe system as the asymmetry of the lens is increased.

Figure 21. Fringe patterns observed in micrographs made with lenses having various amounts of astigmatism.
A. Corrected lens with hole in focus; B. 1.2µ; C. 2.6µ; D. 5.2µ; E. Corrected lens 1.5 out of focus. Magnification in all pictures is 100,000.
Since the rotation of the correcting lens and the changing of the applied potential can be accomplished without opening the microscope column, this correction proceeds very rapidly. The limit on the totality of correction is determined by the visibility of the fringe system. With the microscope operating properly, it is possible to make corrections to less than one division of the focus control, or to less than .5 microns.

As a check on the numerical calculations concerning the correcting lens, observations were made using a magnetically compensated lens. Various potentials were applied to the electrostatic lens, and the resulting amounts of astigmatism measured in terms of the difference in the settings of the fine focus control for focus in the two directions. The results of these measurements are shown in figure 22. The solid curve gives the theoretical astigmatism as a function of the applied voltage. The dashed curve gives the theoretical form if the constant is increased by 33%. It can be seen that the experimental points fit the dashed curve very well. An error of this magnitude could easily have been introduced by incorrect calibration. Further, the magnification has been estimated as 80 in the theoretical considerations on the basis of dimensions of the microscope. An equally acceptable estimate, $M = 70$, results in a curve that differs from the experimental points by only a few percent. For these
Figure 22. Astigmatism introduced by the electrostatic cylindrical lens as a function of voltage.

reasons it can be seen that the effect is of the proper magnitude.

6. Conclusions

This work has resulted in two important contributions, the first to the field of optics as a whole, and the second to the field of electron microscopy. Up to this time, the diffraction patterns which are formed in the imaging of a point source by an astigmatic lens had not been completed for the case of circular apertures except for extremely small amounts of astigmatism. While this has been of little but academic interest in the field of light optics where
lenses may be ground with great precision, it is of very great importance in the field of electron optics where asymmetry is invariably present. While estimates had been made as to the effect of this defect in terms of geometric optics, there does not seem to be any published data in which such a basic principle as the Rayleigh Quarter wave limit had been applied. As is usually the case, the Rayleigh limit is more severe than necessary, and it has been shown in the above discussion that reasonable correction leads to a resolving power that corresponds to that of a perfect lens.

The numerical theory has been verified and extended by resorting to the light optical analogy. The characteristic four pointed star shaped patterns that had been computed were indeed found to exist for small amounts of aberration. As the aberration was increased, the form of the pattern changed very noticeably, and then settled to a system of discreet lines which form a distorted checkerboard type of pattern. While it is not apparent in the prints and only barely so on the original film, the sides of the checkerboards are extended by very low intensity tails. As the asymmetry is increased, the lines become more numerous. Of course the overall size of the pattern increases with the central portion corresponding roughly to the geometric circle of confusion.

The second contribution is the electrostatic system
of correction. It was shown theoretically that if the finite length of the magnetic lens field is taken into account, there exist two focal lines, each perpendicular to the optic axis, and perpendicular to each other. This then indicates that correction can be accomplished by an electrostatic lens of proper strength and orientation.

Such a lens was constructed and its axial potential determined by use of an electrolytic plotting tank. The resulting potential was found to be of a form that allows relatively rapid computations by use of an approximate method. The focal length of the correcting lens having been determined in terms of the ratio of the applied potential to the accelerating potential of the electron beam, attention was then paid to the geometric optics of the whole lens system, the objective and the correcting lens. Computations were carried out including the geometry of the lens position within the microscope and estimating the magnification of the objective in terms of the separation of the objective and projector pole pieces. The resulting data was plotted in terms of microns difference in focal lengths as a function of applied voltage.

The operation of the correcting lens was then checked experimentally. An initially astigmatic objective was corrected by the method discussed above: introducing extreme astigmatism to determine the relative orientations of the existing astigmatism and that of the electrostatic
lens, rotation of the lens, and reduction of the applied potential. In the case of constant use, the microscope operator would be able to omit the first step, since the location of the correcting field with respect to the physical lens is of course fixed, and the rotation of the image formed by the objective is rather rapidly acquired by experience. A check on the numerical calculations was made by determining the amount of astigmatism introduced by the correcting lens using a magnetically corrected objective. The experimental results agree with the theoretical predictions to the magnitude of correction obtained with a given potential applied to the correcting lens.

The relative ease and rapidity of correction using the electrostatic lens as compared to the magnetic process of correction, together with the removal of doubt as to the validity of using such a system, both on a theoretical basis and by demonstration, recommend the inclusion of this apparatus to existing electron microscopes.
Since this work was completed, a paper was presented by John H. Reisner of RCA at the November meetings of the ESMA. From the abstract, it appears that Reisner proposes the electrostatic compensation of the objective by the introduction of an electrode system within the pole piece. In this case, the rotation is accomplished electrically and the rotation can be determined separately from the lens strength. This would certainly seem to be a convenient system to use if the controls are indeed independent. However, it requires modification of the lens assembly itself which is a task usually undertaken with misgivings. Further, in this case the electron motion takes place in combined magnetic and electric fields, the latter of which would seem to present great difficulty in theoretical treatment. It is the opinion of the writer that this development will be of principal value commercially where the whole lens assembly can be sold as a unit.
APPENDIX A

Formal Analogy Between Light and Electron Optics

The Newtonian equations of motion for an electron in electric and magnetic fields can be written as:

\[ (A-1) \quad m\ddot{x} = -e \frac{\partial \Phi}{\partial x} + \frac{\Phi}{c} (\dot{y}B_x - \dot{z}B_y) \]

\[ m\ddot{y} = -e \frac{\partial \Phi}{\partial y} + \frac{\Phi}{c} (\dot{x}B_y - \dot{z}B_z) \]

\[ m\ddot{z} = -e \frac{\partial \Phi}{\partial z} + \frac{\Phi}{c} (\dot{x}B_z - \dot{y}B_x) \]

If these equations are multiplied by \( \dot{x}, \dot{y}, \dot{z} \), respectively and added, the resulting expression can be integrated with respect to time to give the energy integral

\[ (A-2) \quad \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + e \Phi = K \]

The potential is normally measured relative to the source of electrons. Therefore, if it is assumed that the emission energy of the electrons is given by \( K = -e \Phi_e \), this must be the total energy. Thus the kinetic energy of the electron at any point in the field is given by

\[ (A-3) \quad \frac{1}{2} m \dot{v}^2 = -e (\dot{\Phi} + \Phi_e) \equiv -e \Phi \]

If now \( ds \) represents the incremental displacement of an electron whose total energy is \( -e \Phi_e \) along its trajectory, then

\[ (A-4) \quad ds^2 = (1 + x'^2 + y'^2)dz^2 + v^2 dt^2 = \frac{2e \Phi}{m} dt^2 \quad ; \quad x' = \frac{dx}{dz}, \quad y' = \frac{dy}{dz} \]
and therefore

\[(A-5) \quad \frac{d}{dt} \frac{d\mathbf{z}}{dz} = \sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}} \frac{d\mathbf{z}}{dz}\]

This last result may be used in the first two Newtonian equations to give \(x\) and \(y\) in terms of the independent variable \(z\). The resulting equations are:

\[(A-6) \quad m\sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}} \frac{d\left(x'y'\sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}}\right)}{dz} = -e \frac{\partial \phi}{\partial x} + e \sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}} (y'B_x - x'B_y)\]

\[(A-7) \quad m\sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}} \frac{d\left(y'\sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}}\right)}{dz} = -e \frac{\partial \phi}{\partial y} + e \sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}} (x'B_x - x'B_y)\]

At this point, it is convenient to introduce the vector potential \(\mathbf{A}\) which is partially defined by the equation

\(\nabla \times \mathbf{A} = \mathbf{B}\) or

\[(A-7) \quad B_x = \frac{\partial A_y}{\partial y} - \frac{\partial A_y}{\partial z} ; \quad B_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_x}{\partial x} ; \quad B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\]

Substituting these values in the above equations, and adding and subtracting \(x' \frac{\partial A_x}{\partial x}, y' \frac{\partial A_y}{\partial y}\) in the first and second equations respectively, leads to the equations:

\[(A-8) \quad \frac{d}{dz} \left(x'y'\sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}}\right) = -e \sqrt{\frac{m(1+x'y')}{2x}} \frac{\partial \phi}{\partial x} + e \left(x'y' \frac{\partial A_x}{\partial x} + y' \frac{\partial A_y}{\partial y} + \frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x}\right)\]

\[
\quad \frac{d}{dz} \left(y'\sqrt{-2\alpha \cdot \frac{\mathbf{z}}{m(1+x'y')}}\right) = -e \sqrt{\frac{m(1+x'y')}{2y}} \frac{\partial \phi}{\partial y} + e \left(x'y' \frac{\partial A_x}{\partial y} + y' \frac{\partial A_y}{\partial x} + \frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x}\right)\]
These equations are nothing more than the Euler equations corresponding to the variation problem:

\[(A-9) \quad \oint F dx = 0 \quad ; \quad F = \sqrt{2meq} \sqrt{1 + x'^2 + y'^2} - \frac{q}{e} \left( y'A_x + y'A_y + A_z \right)\]

or of the variation problem:

\[(A-10) \quad \oint F' ds = 0 \quad ; \quad F' = \sqrt{2meq} - \frac{q}{e} \vec{A} \cdot \vec{r} \quad ; \quad \vec{r} = \left( \frac{x'}{\sqrt{1 + x'^2 + y'^2}}, \frac{y'}{\sqrt{1 + x'^2 + y'^2}}, 1 \right)\]
APPENDIX B

The Evaluation of the Integral

\[ R = \int \int e^{i k \sin \theta (\cos \phi - \sin \phi)} e^{i k \sin \phi \cos (\phi - \gamma)} \sin \phi \cos \phi d\phi d\theta \]

The first step in the evaluation is the expansion of the first exponential as a series. This reduces the problem to determining the value of the sum of integrals

(B.1) \[ R = \sum \int \int \frac{(i k \sin \theta)^n (\cos \phi - \sin \phi)^n}{n!} e^{i k \sin \phi \cos (\phi - \gamma)} \sin \phi \cos \phi d\phi d\theta \]

To evaluate the individual integrals in the sum, the derivatives of a known integral are considered. The integral

\[ \int e^{i k \sin \phi \cos (\phi - \gamma)} \sin \phi \cos \phi d\phi = 2\pi \frac{\mathcal{J}(\Delta \phi)}{\Delta k} \]

is the well known integral giving the amplitude distribution in a plane, the origin of which is the geometric focus of a converging spherical wave. If \( \cos(\phi - \gamma) \) is expanded in terms of the angles \( \phi \) and \( \gamma \), and \( \eta \) and \( \xi \) defined as below, this integral may be differentiated partially with respect to \( \xi \) and \( \eta \) any number of times with the result:

(B.2) \[ \frac{\partial^n}{\partial \xi^n \partial \eta^m} \int \int e^{i k \sin \phi [\cos \eta - \eta \sin \phi]} \sin \phi \cos \phi d\phi d\theta = \int \int \frac{(i k \sin \phi)^n \cos \xi \sin \eta}{\sin \phi \cos \phi} \sin \phi \cos \phi d\phi d\theta \]

\( \xi = r \cos \theta \quad \eta = r \sin \theta \)
These last integrals are of the same form as those to be evaluated in the sum of integrals on the previous page. By substituting these results, the evaluation of the original integral therefore becomes a problem of evaluating the sum of partial derivatives of a known function:

\[ A = \sum_{n=0}^{\infty} \left( -\frac{e^{i\theta}}{\pi} \right)^n \left( \frac{2^2}{\beta^2 n!} - \frac{2^2}{\eta^2 n!} \right) 2\pi \alpha^2 \frac{J_0(\alpha \kappa \eta)}{\alpha \kappa} \]

In attempting the final evaluation, it is convenient to make use of the substitutions \( x = \alpha \kappa \xi \); \( y = \alpha \kappa \eta \); \( s = \alpha \kappa \kappa = \sqrt{x^2 + y^2} \) hence:

\[ A = 2\pi \alpha^2 \sum \left( -\frac{\left(\alpha \kappa \xi\right)^n}{\pi^n} \left( \frac{2}{\beta^2} - \frac{2}{\eta^2} \right)^n \frac{J_0(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}} \right) \]

First the quantity:

\[ (B-5) \left( \cos \alpha \varphi \frac{J_0(\xi)}{s^n} \right) \text{ or } \left( \frac{2}{\beta^2} - \frac{2}{\eta^2} \right) \cos \alpha \varphi \frac{J_0(\xi)}{s^n} \]

and \( \alpha, \beta, \varphi \) are arbitrary integers, is evaluated.

\[ (B-6) \frac{2}{\beta^2} \left( \cos \alpha \varphi \frac{J_0(\xi)}{s^n} \right) = \frac{2}{\beta^2} \left\{ \left( \alpha \phi \right)^n \frac{J_0(\xi)}{s^n} \right\} \]

\[ = \frac{2}{\beta^2} \left\{ 2 \alpha \left( \alpha \phi \right)^n s^{\alpha - \beta} \frac{J_0(\xi)}{s^n} + \left( \beta - \alpha \right) \left( \alpha \phi \right)^n s^{\alpha - \beta} \frac{J_0(\xi)}{s^n} \right\} \]

The last term arises from the properties of Bessel functions:

\[ (B-7) \frac{d}{dx} \left( \frac{J_0(x)}{x^n} \right) = -\frac{J_{n+1}(x)}{x^n} \text{ hence } \frac{d}{dx} \left( \frac{J_0(x)}{x^n} \right) = -\frac{J_{n+1}(x)}{x^n} \left( \frac{x}{s^n} \right) = -\frac{J_{n+1}(\xi)}{x^n} \]
Now carrying out the second partial differentiation, and collecting in terms of $(\cos 2\tau)^n$

\[(B-8) \frac{d^2}{d\tau^2} \left\{ \cos 2\tau \frac{F(\tau)}{3^\tau} \right\} = (\cos 2\tau)^{n-2} \left\{ 4\alpha(x^2 - y^2) \frac{F(\tau)}{3^\tau} + \right\} + (\cos 2\tau)^{n-1} \left\{ 2\alpha \frac{F(\tau)}{3^\tau} + 4\alpha(x^2 - y^2) \frac{F(\tau)}{3^\tau} - 4\alpha x^2 \frac{F(\tau)}{3^\tau} \right\} + (\cos 2\tau)^n \left\{ (x^2 - y^2) \frac{F(\tau)}{3^\tau} + \right\} \]

If \( \frac{d^2}{d\tau^2} \left\{ \cos 2\tau \frac{F(\tau)}{3^\tau} \right\} \) is carried out, the result is found to be the same as above except that the sign of the second term is negative and \( x \) is replaced by \( y \).

Therefore, combining these results:

\[(B-9) \frac{d^2}{d\tau^2} \left\{ \cos 2\tau \frac{F(\tau)}{3^\tau} \right\} = (\cos 2\tau)^{n-2} \left\{ 4\alpha(x^2 - y^2) \frac{F(\tau)}{3^\tau} + \right\} + (\cos 2\tau)^{n-1} \left\{ 2\alpha \frac{F(\tau)}{3^\tau} + 4\alpha(x^2 - y^2) \frac{F(\tau)}{3^\tau} - 4\alpha x^2 \frac{F(\tau)}{3^\tau} \right\} + (\cos 2\tau)^n \left\{ (x^2 - y^2)(-x^2 - 2\alpha) \frac{F(\tau)}{3^\tau} - 2(x^2 - y^2)(-x^2 - 2\alpha) \frac{F(\tau)}{3^\tau} + (x^2 - y^2) \frac{F(\tau)}{3^\tau} \right\} \]

Making use of the definition \( \cos 2\tau = \frac{x^2 - y^2}{s^2} \)

\[(B-10) \frac{d^2}{d\tau^2} \left\{ \cos 2\tau \frac{F(\tau)}{3^\tau} \right\} = \cos 2\tau \left\{ \frac{F(\tau)}{3^\tau} - 2(x^2 - y^2) \frac{F(\tau)}{3^\tau} + (x^2 - y^2)(-x^2 - 2\alpha) \frac{F(\tau)}{3^\tau} + (x^2 - y^2) \frac{F(\tau)}{3^\tau} \right\} \]

Having evaluated \( D \) in terms of the arbitrary integers \( \alpha, \beta, \gamma \), the value of \( \frac{F(\tau)}{3^\tau} \) will be proved by induction.
It is therefore asserted that:

$D^n \frac{J(n)}{x} = (\frac{\delta^x}{2\pi} - \frac{\delta^y}{2\pi})^n \frac{J(n)}{x}$

(B-11)

$$= \sum_{j=0}^{\infty} \frac{(-1)^j 2j^n(n-2j)!}{(n-\delta)(n-2\delta)!} (\cos 2\tau)^{n-2j} \leq \frac{(-1)^j (n-2j)!}{2^j j! (j-\delta)! (n-2j-2\delta)!} \frac{J_{an+1-2j}}{\delta^{j+\delta}}$$

where $\frac{\delta^x}{2\pi}$ if $n$ is even and $\frac{\delta^y}{2\pi}$ if $n$ is odd.

This equation is true for at least one value of $n$: $n = 0$ since then $n = j = k = 0$. There remains only to be shown that:

(B-12)  \( (\frac{\delta^x}{2\pi} - \frac{\delta^y}{2\pi}) D^n \frac{J}{x} = D^{n+1} \frac{J}{x} \)

According to the evaluation of $D \left\{ \cos 2\tau \frac{J}{x} \right\}$ only two terms in the summation over $j$ in $D^n \frac{J}{x}$ can contribute to the term in $(\cos 2\tau)^{n+1-2\delta}$ in $D^{n+1} \frac{J}{x}$. These are the terms for which $j = \delta$ and $j = \delta - 1$. Therefore, the coefficient of $(\cos 2\tau)^{n+1-2\delta}$ in $D(D^n \frac{J}{x})$ is:

(B-13)

$$C_{n+\delta} = \frac{S \delta! \eta^l(n-2\delta)!}{(n-\delta)! (n-2\delta)!} \leq \frac{(-1)^l (n-\delta-k)!}{2^l l! (l-k)! (n-2l-2\delta)!} \left\{ \frac{J_{an+1-k}}{\delta^{l-k+\delta}} + \frac{J_{an+1-k}}{\delta^{l+k}} \right\}$$

Collecting in terms of a single summation:
In this rewritten form, the contribution from the \( j = \delta - 1 \) coefficient is summed over an extra term, \( k = \delta \). However, in the two places where it appears, it is to be noted that the coefficient is zero for this case so the value of \( C_{n+1-2\delta} \) is unaltered.

This last summation may be expanded and recollected in terms of \( \frac{J_{2n+3-\delta-k}}{s^{k+k+1}} \). Doing this:

\[
(B-15) \quad C_{n+1-2\delta} = \left( -1 \right)^{\delta} \frac{2^\delta n!(2n-2\delta)!}{(n-\delta)!(n-2\delta)!} \sum_{k=0}^{\delta} \left( -1 \right)^k \frac{(n-\delta-k)!}{2^k k! (s-k)!(2n-2\delta-2k)!} \frac{J_{2n+3-\delta-k}}{s^{k+k+1}}
\]

\[
x \left\{ 1 + \frac{2(s-k)(2n-2\delta+1)(2n-2\delta-2k+1)}{(n-2\delta+1)(2n-2\delta+2k+1)(n-2\delta+1)} \frac{1}{2n-2\delta-2k+1}
\right. 
\]

\[
+ \left. \frac{4(s-k)(2n-2\delta+1)(2n-2\delta-2k+1)}{(2n-2\delta+2k+1)(2n-2\delta-2k+1)(2n-2\delta+2k+1)} \right\}
\]

With algebraic manipulation, the quantity in the brackets can be shown equal to \( \frac{(n+\delta)(2n-2\delta-1)}{(n-2\delta+1)(2n-2\delta-2k+1)} \) and therefore:
Therefore, we obtain:

(B-17) \[
D(B^\alpha \frac{a}{k}) = \sum_{d=0}^{\left\lfloor \frac{n+1}{2} \right\rfloor} \left( (-1)^d \frac{2^d (2n-2d)!}{(n-d)! (n-2d)!} \right) \frac{k^d}{2^k B^\alpha (n-d)! (n-2d)!} \frac{J_{\alpha (n+1-\frac{d}{2}-k)} (2n+1-k)}{\Gamma (2n+1-2k) k^{n+1}}
\]

Since we have shown the assumed form to be correct for \( n = 0 \), it must hold for all \( n \).

Therefore, the evaluation of the integral \( A \) is reduced to determining the sum of known functions, the Bessel functions.

(B-18) \[
A = \sum_{n=0}^{\infty} \sum_{d=0}^{\left\lfloor \frac{n+1}{2} \right\rfloor} \left( (-1)^d \frac{2^d (2n-2d)!}{(n-d)! (n-2d)!} \right) \frac{k^d}{2^k B^\alpha (n-d)! (n-2d)!} \frac{J_{\alpha (n+1-\frac{d}{2}-k)} (2n+1-k)}{\Gamma (2n+1-2k) k^{n+1}}
\]
APPENDIX C

Consideration of the Effect of Spherical Abberation on the Images Formed in the Light Optical Analogue

Since transverse spherical abberation is a function of the cube of the aperture, we would expect it to have little effect if small apertures are used. Indeed this is the case for the lenses actually used in making the abberation pictures, and it will be shown below that the spherical abberation is entirely negligible.

To determine the effect of the spherical abberation if it alone is present, we again must first determine the wave surface associated with a bundle of rays so affected. Since the coefficient of spherical abberation normally employed in theoretical discussions refers to transverse abberation, we will start using this term, and later transform the results to terms of longitudinal abberation which is the form in which it appears in ray tracing.

Referring to the diagram (figure 14), we have from the geometry:
\[ s-x=\rho \cos \Theta \; ; \; y=\rho \sin \Theta \]
\[ s-S=L \; ; \; s/S=r/Cr^3 \]

and solving the last two equations for \( s \) and \( S \) we have
\[ S=sCr^2 \; ; \; s=L/(1+Cr^2) \], and thus
\[ \tan \Theta = r(1+Cr^2)/L \; ; \; \sin \Theta = \frac{r(1+Cr^2)}{\sqrt{r^2(1+Cr^2)^2 + L^2}} \]
\[ \cos \Theta = \frac{L}{\sqrt{r^2(1+Cr^2)^2 + L^2}} \]
Figure 23. Geometry and notation used in developing the aspherical wave surface

If the first equation is solved for $x$, the value of $s$ in terms of $L$ and $r$ substituted, and finally, differentiated with respect to $r$:

\[(C-1) \quad \frac{dx}{dr} = \frac{ds}{dr} - \frac{dp}{dh} \cos \theta + p \frac{dp}{dh} \sin \theta\]

Differentiating the equation for $y$ with respect to $r$:

\[(C-2) \quad \frac{dy}{dr} = \frac{dp}{dh} \sin \theta + p \frac{dp}{dh} \cos \theta\]

By definition, the derivative of $y$ with respect to $x$ gives the value of the tangent of the angle formed by the intersection of the tangent to the wave surface and the optic axis. Therefore:

\[
\frac{dy}{dx} = \frac{dy}{dr} / \frac{dx}{dr}, \text{ and solving for } \frac{dy}{dr}
\]
(C-3) \( \frac{dy}{dn} = \frac{dx}{dn} \cot \theta \)

If equations (C-1) and (C-2) are substituted in (C-3)

\[
\frac{ds}{dr} \cot \theta - \frac{dc}{dr} \cos^2 \theta + \frac{dp}{dr} \cos \theta + \frac{dp}{dr} \sin \theta + \frac{dp}{dr} \cos \theta \cdot
\]

If this last equation is slightly rearranged, there results the differential equation for the wave surface:

(C-4) \( \frac{d\phi}{dn} = \frac{ds}{dn} \cos \theta \)

If the value of \( s \) and \( \cos \theta \) as determined from the geometry, are substituted in the last equation:

(C-5) \( \frac{d\phi}{dr} = \frac{-2CLr}{[1 + Cn^2]^2 [1 + \frac{n^2(1+Cn^2)}{L^2}]^3} \)

or expanding by the binomial theorem:

(C-6) \( \frac{d\phi}{dr} = \frac{-2CLr}{(1+Cn^2)^3} \left[ 1 - \frac{n^2(1+Cn^2)}{2L^2} + O_6 \right] \)

where \( O_6 \) represents terms of the sixth and higher orders in \( r \).

Integrating this equation with respect to \( r \), and making use of the condition that with \( r = 0, \phi = L \):

(C-7) \( \phi = \frac{L}{1+Cn^2} + \frac{Cn^4}{4L} + O_6 \)

This last result gives the geometrical length of the path from the particular wave surface chosen to the intersection of the ray with the optic axis. To obtain the optical path length, it is necessary to multiply this distance by the index of refraction. For the purpose at
hand however, it is more convenient to determine the phase distribution in the plane perpendicular to the optic axis which intersects it at the same point as the chosen wave surface. The additional distance traversed by the ray is given by:

\[(C-8) \quad (n^2 + s^2)^{\frac{1}{2}} \cdot \rho, \quad \text{or} \quad \frac{D}{\sin \theta} - \rho, \quad \text{and by} \]

expansion:

\[(C-9) \quad s = -\frac{n^2 (1 + CR^2)}{2L} + \frac{n^4}{8L^3} + O_6 \]

The minus sign indicates the lag in phase at the surface.

To determine the total difference in path lengths, we assume the plane to act as a secondary source of radiation, with the phase at various points given by the last equation. We now consider a point lying between the points of intersection of the marginal and paraxial rays as determined geometrically. The paraxial rays of course intersect at a distance \(L\) from the plane, and the marginal ray intersects at a distance \(L - S_m\), where \(S_m = \frac{CR^2L}{1+CR^2}\), from the plane. In the last equation, \(R\) represents the separation of the marginal ray from the optic axis, and \(S_m\) is determined from the original geometric equations. Further, we express \(r = nR\), and \(D = mS_m\).

With this notation, the difference in path length of a ray leaving the plane at a separation \(r\) from the axis from the path length along the axis to a point \(L - D\) from
the plane is:

\[(C-10) \quad \sqrt{(nR)^2 + (L-D)^2} - (L-D)\]

If the value for \(D\) is substituted in this equation, and the root expanded, we find the total path difference, which is the algebraic sum of the two path differences, as:

\[(C-11) \quad dL = -\frac{n^2R^2}{2L} \left(1 + \frac{Cn^2R^2}{2} \right) + \frac{n^4R^4}{8L^3} + \frac{n^2R^2}{2L} \frac{mcR^2}{1 + cR^2} \left[\frac{1}{1 - \frac{mcR^2}{1 + cR^2}}\right] - \frac{n^4R^4}{8L^3} \left[\frac{1}{1 - \frac{mcR^2}{1 + cR^2}}\right]^2\]

If \(S_m\) is solved for \(CR^2\), the latter is found to have the value \(S_m/(L-S_m)\) which is a small quantity that may be neglected compared to one. Making use of this, the denominators in the last two terms may be again expanded giving:

\[(C-12) \quad dL = -\frac{n^2R^2}{2L} \left(1 + \frac{Cn^2R^2}{2} \right) + \frac{n^4R^4}{8L^3} + \frac{n^2R^2}{2L} \left(i + mcR^2\right) \frac{n^4R^4}{L^3} \left(i - \cdots\right)\]

By rearranging the above

\[(C-13) \quad dL = \frac{R^2C}{4L} n^2(2m-n^2) = \frac{S_m}{4L^2} n^2(2m-n^2)\]

The first value gives the difference in path lengths with respect to the path length along the axis in terms of the coefficient of transverse spherical aberration \(C\), and the second in terms of the longitudinal spherical aberration \(S_m\).

In both cases, the values depend on the parameters \(n\) and \(m\). By setting the differential of \(n^2(2m-n^2)\) with respect to \(n\) equal to zero, a relative maximum difference is found for \(n^2 = m\). This is a maximum with respect to zero however, and it is to be noted that for values of \(m\) less than \(\frac{3}{2}\) the range
of path length differences must be computed from \( m^2 - 2m + 1 \),
that is the difference between the path length at the
relative maximum and that for \( n=1 \). For \( m \) greater than or
equal to the extreme value in the negative direction is
zero, hence the range is given by \( m^2 \). From this last
result, it is obvious that the optimum \( m = \frac{1}{2} \). The equation
to be considered for \( m \leq \frac{1}{2} \) decreases monotonically \( 0 \leq m \leq 1 \).
Hence the optimum value of \( m \) must be \( \frac{1}{2} \). With this optimum
value:

\[
(C-14) \quad dL = \frac{S_m R^2}{16 L^2}
\]

This is to be compared to \( \frac{Dm^4}{L} \) for astigmatism. In the
analogue, \( S_m \) is of the order of \( 0.005 \) cm while \( D \) is of the
order of \( 5 \) cm. This result then indicates that with reason­
ably small apertures, rather large numerical values of
longitudinal spherical aberration may be tolerated with the
actual path differences still not exceeding the Rayleigh
quarter wavelength limit.

It is to be noted that although \( (C-14) \) was derived
with a tacit assumption of spherical symmetry, the form
would be identical for cylindrical symmetry with respect
to a line focus instead of a point focus. However, in the
latter case, we are restricted to considering one plane
at a time.

The rays were traced for the lenses used, and the
notation and results appear on the following page.
Figure 24. Geometry and notation used in ray tracing

Assuming the light to travel from left to right,

$L$ separation of intersection of initial ray, produced if necessary, with the optic axis, from the pole of the lens surface. Positive to right.

$L'$ separation of intersection of refracted ray, produced if necessary, with the optic axis from the pole of the lens surface. Positive to right.

$\theta$ angle formed by intersection of initial ray and optic axis, measured counterclockwise from ray to optic axis.

$\theta'$ as above but referring to refracted ray.

$I$ angle formed by ray and outward drawn normal, measured counterclockwise from ray to normal.

$I'$ as above but referring to refracted ray.

$R$ radius of curvature of surface, positive if center of curvature is to right of surface.

$N$ index of refraction in space of unretracted ray.

$N'$ index of refraction in space of refracted ray.

With this notation, the ray tracing is based on the following formulae:
\[
\begin{align*}
\sin I &= \sin \frac{U(L-R)}{R} \quad \text{law of sines} \\
\sin I' &= \frac{N'}{N} \sin I \quad \text{law of refraction} \\
U' &= U - I - I' \quad \text{ext angle of triangle equals sum of remote interior angles set up for } U, I, Q, U', I', Q \quad \text{and } Q \text{ eliminated} \\
L' - R &= R \frac{\sin I'}{\sin U'} \quad \text{law of sines}
\end{align*}
\]

From these basic formulae, we can obtain for paraxial rays

\[
\begin{align*}
U' &= U \\
L(N' - N) &= \frac{NR}{P} \\
L' &= \frac{N'LR}{(N' - N)} \frac{L - NR}{P}
\end{align*}
\]

In the case of a plane surface \( U = I; U' = I' \); and

\[
L' = L \frac{\tan U}{\tan U'}
\]

In addition to the ray tracing formulae, the following trigonometric expansions are used:

\[
\begin{align*}
\sin x &= x - \frac{x^3}{6} + \frac{x^5}{120} \\
\sin^{-1} x &= x + \frac{x^3}{6} + \frac{3x^5}{140} \\
\tan x &= x + \frac{x^3}{6}
\end{align*}
\]
Numerical Values Obtained in Tracing Various Rays through a Combination of a Bi-convex Spherical Lens and a Plano-convex Cylindrical Lens in Axial Contact.

Index of refraction of glass in both lenses: \( N = 1.53 \)
Object point located 50 centimeters to left of first lens surface; angle formed by marginal ray and axis is \(-0.004\) radians.

At the first surface:

\[
\begin{array}{l}
L = 50 \text{ cm}, \quad U = -0.004 \text{ radians} \\
\sin U = 0.003 \quad 989 \quad 333 \\
\sin I = 0.011 \quad 514 \quad 722 \quad 266 \\
\sin I' = 0.007 \quad 514 \quad 722 \quad 266 \\
I = 0.011 \quad 514 \quad 722 \quad 266 \\
I' = 0.007 \quad 514 \quad 722 \quad 266 \\
U' = 0.000 \quad 000 \quad 174 \quad 320 \\
L' - R = 114.731.222 \quad 24 \\
L' = 114.757.722 \quad 24 \text{ cm} \\
\end{array}
\]

Thickness of lens at center 0.21 cm

At second surface:

\[
\begin{array}{l}
L = 114.757.512 \quad 24 \\
U = 0.000 \quad 000 \quad 174 \quad 320 \\
\sin U = 0.000 \quad 000 \quad 174 \quad 320 \\
\sin I = -0.007 \quad 514 \quad 722 \quad 266 \\
\sin I' = -0.011 \quad 514 \quad 722 \quad 266 \\
I = -0.007 \quad 514 \quad 722 \quad 266 \\
I' = -0.011 \quad 514 \quad 722 \quad 266 \\
U' = 0.004 \quad 002 \quad 189 \quad 351 \\
\sin U' = 0.004 \quad 002 \quad 189 \quad 351 \\
L' - R = 76.493 \quad 331 \\
L' = 49.993 \quad 331 \quad 5 \\
\end{array}
\]

Separation of lenses on axis: zero

At third surface (a plane):

\[
\begin{array}{l}
L = 49.993 \quad 331 \quad 5 \\
U = I = 0.004 \quad 002 \quad 189 \quad 351 \\
\sin I = 0.004 \quad 002 \quad 189 \quad 351 \\
\sin I' = 0.002 \quad 615 \quad 803 \quad 050 \\
I' = U' = 0.002 \quad 615 \quad 816 \quad 033 \\
\tan U = 0.004 \quad 002 \quad 210 \quad 719 \\
\tan U' = 0.002 \quad 615 \quad 841 \quad 999 \\
L' = 76.489 \quad 270 \quad 86 \\
\end{array}
\]
Thickness of lens at center .17 cm, at margin
.162 452 937 662 cm.

Using values of tan U and tan U' from the previous page,
we find in the plane of zero curvature:
at margin:

\[
\begin{align*}
L & = 76.326 \text{ 817 92} \\
L' & = 49.887 \text{ 152 37 measured from surface} \\
L' & = 49.879 \text{ 605 31 measured from pole}
\end{align*}
\]

In plane of zero curvature through optic axis:

\[
\begin{align*}
L & = 76.319 \text{ 270 86} \\
L' & = 49.882 \text{ 219 61}
\end{align*}
\]

Faraxial ray in plane of zero curvature:

\[
\begin{align*}
L_p & = 76.33 \text{ exactly} \\
L_p' & = 49.888 \text{ 868 688}
\end{align*}
\]

In plane of curvature \( R = 26.6 \)

\[
\begin{align*}
L & = 76.319 \text{ 270 86} \\
U & = .002 \text{ 615 816 033} \\
\sin U & = .002 \text{ 615 803 050} \\
\sin I & = -.010 \text{ 149 243 865} \\
\sin I' & = -.015 \text{ 528 343 106} \\
I & = -.010 \text{ 149 418 113} \\
I' & = -.015 \text{ 528 967 231} \\
U' & = .007 \text{ 995 364 151} \\
\sin U' & = .007 \text{ 995 278 996} \\
L' - R & = 51.463 \text{ 009 11} \\
L' & = 24.988 \text{ 009 11} \\
L_p & = 24.972 \text{ 163 57}
\end{align*}
\]

We have as the final result of the ray tracing:

\[
dL < 10^{-8} = \frac{1}{5280} \text{ wave lengths or less than } 1^\circ \text{ difference in phase at near focal line.}
\]

\[
dL < 2 \times 10^{-8} = \frac{1}{2290} \text{ wave lengths or less than } 1^\circ \text{ difference in phase at near focal line.}
\]
REFERENCES


8. Born, Max, Optik, Berlin: Springer (1933)


I, Robert Arnold Watkins, was born in Boston, Massachusetts, August 3, 1926. I received my secondary school education in the public schools of the town of Needham, Massachusetts. My undergraduate training was started at Brown University. After completing three semesters, I enlisted in the Naval Reserve. The first fourteen months of my service were spent in the Radio Technician Training program and the last nine months were spent at The Ohio State University as a member of the V-12 program. After my discharge, I returned to Brown University from which I received the degree Bachelor of Science Magna Cum Laude in 1947. From The Ohio State University, I received the degree Master of Science in 1948. The following year, I was employed by The West Virginia University as an Instructor of Physics. From September of 1949 until March of 1952, I was in residence at The Ohio State University while completing some of the requirements for the degree Doctor of Philosophy. During this period, I was employed first as a Graduate Assistant, and later as a Technical Assistant responsible for repair and upkeep of the electron microscopes. Since March of 1952, I have been employed in the color television group at the Zenith Radio Corporation of Chicago, Illinois.