UNIVERSITY OF LONDON

Imperial College of Science and Technology,
Physics Department,
Applied Optics Section.

NEW TECHNIQUES FOR THE EVALUATION AND DESIGN OF OPTICAL SYSTEMS

by

William Bailey King, A.R.C.S., D.I.C., M.Sc.

Thesis submitted for the Ph.D Degree
of the University of London
1966

ABSTRACT

This thesis describes new techniques for the evaluation and design of high quality optical systems, based on the application of canonical pupil variables, and employing the Strehl ratio as the criterion of image quality. Maréchal's treatment of tolerance theory shows that, in designing high quality systems, one should aim at minimising the mean squared deviation, E, of the wave surface from the spherical form, and this approach has been adopted in the present work. As part of the investigation, a numerical integration procedure has been used to obtain the Strehl ratio, and the accuracy and range of usefulness of the Maréchal's criterion have been studied.

A brief account is given of traditional design methods and automatic design techniques. The advantages of using modern techniques of image-assessment as the criterion of image quality is stressed.

Using the canonical pupil variables, the exit-pupil co-ordinates of any ray are equal to the entrance pupil co-ordinates, at least in a nominally corrected system. For this reason, the wave-aberration of an optical system may be expressed as a function of the entrance pupil co-ordinates, and used in this form directly in the diffraction integrals.

The exploration and specification of the pupil shape in terms of these co-ordinates are also described, it being shown that the vignetted pupil shape may be well approximated by

an ellipse, which is then 'scaled' to correspond to a circular aperture. The wave-aberrations may then be expressed as a polynomial in the corresponding equivalent circular pupil co-ordinates, and pattern of rays traced can then remain the same for different image fields and for different systems of the same class.

It is shown that the variance E (for both the axial and extra-axial images) may be calculated very simply by means of a set of universal coefficients P(i,j), Q(i,j) and R(i,j), once the wave-aberrations of selected rays are known. The values of these coefficients are uniquely determined by the form of polynomial assumed for the wave-aberration, and by the pattern of rays traced. Tables of P(i,j), Q(i,j) and R(i,j) have been calculated for the different cases that arise in practice.

A useful criterion of image quality for the whole field is the sum of suitably weighted values of E for a typical set of image points. Optimisation of a design then requires this sum to be minimised with respect to the different design parameters. For this purpose the parameter derivatives of the wave-aberrations are needed. A method of computing these derivatives is described, and it is also explained how they may be used to give an optimisation procedure having the weighted sum of the variances of the wave-aberration for different field angles as the criterion of image quality.

The computing procedures arising from these studies have been programmed for use with digital computers. Flow diagrams and numerical results are presented. The programmes described in this thesis include : (i) ray-tracing, (ii) pupil exploration, (iii) numerical evaluation of the Strehl ratio, (iv) evaluation of the coefficients P(i,j), Q(i,j), R(i,j), (v) calculation of the parameter derivatives of the aberrations of a system.

The optimisation procedure has been considered only to the point of establishing its feasibility. Also, of course, it is desirable to include the properties of the detector in the assessment of image quality. It was not contemplated, however, to include this aspect of the problem in the present studies. A detailed discussion of the photographic spread, the photographic noise and their effects on the image quality may be found in the book "Fourier Methods in Optical Image Evaluation" by E.H.Linfoot (Focal Press, 1964).

The candidate's own contribution to this thesis includes the material of Chapter 3, the pupil exploration technique of Chapter 4, the whole of Chapter 5, the computing procedure and programming of Chapters 6 & 7, and contributions in discussion to the material of Chapter 8. Chapter 1 is essentially an introduction, and Chapter 2 is a summary of a recent publication of the candidate's supervisor, the reason for the inclusion of which is given in the text.

CONTENTS

CHAPTER 1.		Page
(I-i)	Introduction.	1
(I-ii)	Wave-optics and optical design.	3
(I-iii)	Image criteria for lens design.	4
(I-iv)	Lens design by computer.	12
(I-v)	A general linearisation method.	14
	REFERENCES.	19
CHAPTER 2.		
(II-i)	Canonical pupil co-ordinates in geometrical and diffraction image theory.	21
(II-ii)	The applications of canonical co-ordinates in lens design.	27
	REFERENCES.	32
CHAPTER 3.	The calculation of the Strehl Intensity Ratio, I, by means of a numerical integration method.	
	Introduction.	33
(III-i)	The use of equivalent circular pupil in diffraction theory.	34
(III-ii)	The numerical evaluation of the Strehl Intensity Ratio I.	35
(III-iii)	The flow diagram for evaluating I.	43
(III-iv)	Checking of the numerical integration results.	45

CHAPTER 3 -contd. Page.

(III-v) The comparison of the Strehl Intensity Ratios and the validity of Maréchal's approximate formula. 53

(III-vi) The special property of Strehl Intensity Curves. 71

REFERENCES. 80

CHAPTER 4 The use of canonical co-ordinates in aberration calculations, ray-tracing formulae and pupil exploration.

(IV-i) Object and image space variables. 81

(IV-ii) Finite ray trace. 100

(IV-iii) The pupil exploration. 101

REFERENCES. 129

CHAPTER 5

(V-i) Specification of pupil co-ordinates in ray trace. 130

(V-ii) Wave-aberration polynomials. 135

(V-iii) Evaluation of the variance E of the wave-aberration. 140

(V-iv) The computation of the coefficients P(i,j), Q(i,j) and R(i,j). 145

CHAPTER 6 Wave-aberration derivatives.

(VI-i) The change in aberration with respect to the original reference sphere. 175

(VI-ii) Focal shift derivatives. 182

(VI-iii) Parameter derivatives for curvature change. 184

CHAPTER 6 -contd. Page.

(VI-iv) Distortion and variation in gaussian image height. 189

REFERENCES. 190

CHAPTER 7 The calculation of the wave-aberration and the aberration derivatives.

General introduction. 191

(VII-i) Input stage 193

(VII-ii) Gaussian optics, distortion and primary aberrations. 200

(VII-iii) Finite ray trace. 210

(VII-iv) Pupil ray subroutine. 238

(VII-v) T and $\bar{T}$-rays subroutine. 245

(VII-vi) Axial ray subroutine. 254

(VII-vii) Extra-axial ray subroutine. 262

(VII-viii) Output and discussion. 276

REFERENCES. 277

CHAPTER 8 The evaluation of the merit function, its parameter derivatives and the optimisation procedure.

(VIII-i) Introduction. 279

(VIII-ii) Definition of the merit function Ψ. 284

(VIII-iii) Parameter derivatives of Ψ. 289

(VIII-iv) Optimisation procedure. 299

Discussion. 302

Page.

Special supplement 303

General Summary 304

ACKNOWLEDGEMENTS. 307

Chapter One

(1-i) Introduction.

In most optical systems an object lying in some surface (usually a plane) is required to be imaged on some other surface (also generally a plane), with some prescribed magnification, and the image is also usually required to be geometrically similar to the object.

In designing optical systems our goal is, in each case, to produce a perfect system. When the rays from any given object point are homocentric in the image space and if these rays also intersect at the required position in the image plane, then according to geometrical optics an image identical with the object is obtained. Such an optical imaging system may be referred to as a perfect system in the geometrical optics sense only, since we are ignoring the effects on the image due to diffraction. By a perfect system, we mean one in which diffraction is the only limitation. It is well known that this goal cannot be achieved, but it gives direction to our endeavours.

The design of a lens generally follows a pattern[1]. When confronted with a set of specifications, the designer first chooses a particular configuration of an optical system as a starting point. His first calculations make use of thin lens formulae[2], the usefulness of which depends on the fact that the power of a thin lens can be kept constant with respect to bending.

At this stage, the primary aberration theory[1],[3] proves to be useful for it leads to certain general conclusions of the type of lenses that can give corrected systems. Thus for lenses of zero axial thickness the two primary chromatic aberrations and the Petzval sum are invariant as the lens shape is altered, keeping its power constant. The other four primary aberration terms relating to spherical aberration, coma, astigmatism and distortion depend on the lens shape. It is necessary to calculate and to reduce the total primary aberrations of the system. A system of thin lenses having first order correction may therefore be designed by choosing a set of lens powers and separations having the required total power and satisfying the three shape invariant conditions, the remaining four primary aberrations being corrected by appropriate bending. The system is then modified by introducing thicknesses, passing from a thin lens system to a thick lens of equivalent property[6], keeping the focal length constant and total primary aberrations small. Rays are then traced through the system and the total aberrations then found are interpreted as the sum of the computed primary, plus higher order, aberration terms[4],[5]. In many cases, the designer can make use of a set of design data taken from a suitable design in the literature, particularly in the patent literature, where the success of an earlier designer gives some hope of further success, and this is used as a starting point[8]. The total aberrations of the system are reduced by making suitable changes in the

design parameters of the system, thereby introducing into a first-order corrected system small amounts of primary aberrations to balance higher order ones, and rays are again traced. Successive attempts to improve the system continue until the original specifications are satisfied.

In traditional optical design methods, the results obtained from ray tracing are usually expressed in graphical form. The most common perhaps is to calculate and plot the following: transverse spherical aberration, coma, distortion, and the sagittal and tangential field curvatures[9]. Chromatic aberration is most easily determined by Conrady's method where the path difference between two wave fronts corresponding to two different colours is calculated. Other methods for representing the aberrations graphically, e.g. spot diagrams, will be considered later.

It must be emphasized that apart from the preliminary thin lens work, the traditional method of optical design is essentially a process of trial and error in which variations of the system are made largely on the basis of intuition and previous experience of the designer.

(1-ii) Wave-optics and optical design.

Until recently lens designers have tended to assess the image quality indirectly in terms of the aberrations, rather than directly in terms of the influence of the aberrations on the image itself. As

the aberrations become small, there is a divergence between the image actually produced by an optical system and that predicted by ray optics. Thus the image criterion based on the consideration of the size of the ray intersection pattern on the image plane bears little relation to the actual image in a well-corrected system because of diffraction effects at the pupil. For proper evaluation of optical systems, wave-optics is widely agreed to be needed - more necessary the better the correction. Thus in more modern design methods, applicable to the determination of the properties of a lens system when the physical properties of light are considered, one calculates the shape of the emerging wavefront, or rather the difference between the emerging wavefront and a reference sphere centred on the Gaussian image point. One also needs to consider, at the design stage, what ratio of a given primary to a given higher order aberration will lead to a "best form of correction", and what magnitude is permissible for the residual aberrations after a best balance has been achieved. In what follows a brief account will be given of three definitions of image-quality, and of the means of calculating the image-quality from the computed aberrations of the system.

(1-iii) Image Criteria for Lens Design.

(A) Image-assessment based on images of point sources.

The diffraction image of a point source formed by a system with aberration usually has a complex structure whose amplitude can be expressed by the formula

$$F(u,v) = \iint_{\text{Pupil}} e^{i(kW(x,y) + ux + vy)}\, dx\, dy \qquad (1\text{-}1)$$

where $k = \frac{2\pi}{\lambda}$, $W(x,y)$ is the wave front aberration at the point (x,y) of the pupil, and (u,v) are suitably chosen rectangular coordinates in the image plane.

When the aberrations are small, the image criterion based on the so-called Strehl intensity ratio (Definitionshelligkeit)[10] gives one a single number which expresses the effects of aberration, and has a significance in relation to the images of extended objects. The Strehl ratio I is defined as the ratio of the image intensity at the diffraction focus, with the aberration present, to that without aberration. It is given by the formula

$$I = \left| \frac{\iint_{\text{Pupil}} e^{ikW(x,y)}\, dxdy}{\iint_{\text{Pupil}} dxdy} \right|^2$$

$$= \left| \iint_{\text{Pupil}} e^{ikW(x,y)}\, dA \right|^2 \qquad (1 - 2)$$

where $dA = dxdy / \iint_{\text{Pupil}} dxdy$ is the element of area expressed as a fraction of the total area of the pupil. I is always less than 1.0 if $W(x,y) \neq o$, and the Strehl ratio measures the loss of intensity at the centre of the image of a point source, due to the presence of aberration. Maréchal first showed that aberration toler-

NOTE

✱ In chapter 3, the Strehl ratio is evaluated for a number of cases of spherical aberration using numerical integration. These results, compared with those given by Maréchal's approximate formula, indicate to what extent this latter is useful when I is less than 0.8 . The question of the best form of correction of images with 2-term spherical aberration has been studied by a number of authors. In particular, Linfoot (Optica Acta, vol. 10, p.77-104, 1963) has published results similar to those given in the present work, but for larger values of the secondary spherical aberration coefficient.

ances[11] can easily be calculated if the abberration W(x,y) is small. Briefly this consists of using

$$\overline{W}(x,y) = \iint_{\text{Pupil}} W(x,y)\, dA$$

to denote the mean value of W (x,y), and rewriting (1-2) as

$$I = \left| e^{ik\overline{W}(x,y)} \right|^2 \cdot \left| \iint_{\text{Pupil}} e^{ik\left[W(x,y) - \overline{W}(x,y) \right]} dA \right|^2 \qquad (1-3)$$

If $[W(x,y) - \overline{W}(x,y)]$ is sufficiently small, then

$$e^{ik\left[W(x,y) - \overline{W}(x,y)\right]} \approx 1 + ik\left[W(x,y) - \overline{W}(x,y)\right] - \frac{k^2}{2}\left[W(x,y) - \overline{W}(x,y)\right]^2$$

substituting this in (1-3) leads to the condition

$$I \geq \left| 1 - \frac{2\pi^2}{\lambda^2} E \right|^2 \qquad (1-4)$$

where $$E = \iint_A \left[W(x,y)\right]^2 dA - \left[\iint_A W(x,y)\, dA \right]^2 \qquad (1-5)$$

with A denoting the region of pupil. E is the variance of W(x,y), or the mean squared deviation of the wave surface from a spherical form. For the equality sign in (1-4) to be valid, the aberration must be small enough to make I ≥ 0.80. * [SEE NOTE, L.H.S.] The combination of aberration coefficients which minimises E, will give a maximum value for I. Since E is a quadratic combination of these coefficients, a simple method is available for obtaining the best balance of primary and higher order terms (see Chapter 3). Moreover, provided the aberration is cal-

culated in terms of optical path lengths along rays, the value of E is easily found (see Chapter 5). In designing highly corrected systems such as microscope and telescope O.G.'s, one can use I as a criterion of image quality. The procedure involved is essentially to minimize the sum of the weighted E values, calculated for a set of image points over the field to be covered, with the aim that $I > 0.8$ in the finished design (see Chapter 8).

(B) Image-assessment based on images of extended objects

When the aberrations of an optical system are larger than those for which $I \geqslant 0.8$, the Strehl ratio is insufficient to describe the complex effects of aberration on the diffraction image of a point source. Thus, when one is dealing with most photographic objectives or systems that are not so highly corrected as, for example, microscope objectives a different image criterion, based on the properties of the Optical Contrast Transfer Function, is called for.

Frieser[12] has shown that for a one-dimensional object whose distribution of intensity is expressed by

$$B(u) = \alpha + \beta \cos \frac{2\pi}{\lambda} s\, u \qquad (1\text{-}6)$$

the image has always an intensity distribution given by

$$B'(u') = \alpha + T(s) \cdot \beta \cdot \cos \left(\frac{2\pi}{\lambda} s u' + \Theta(s) \right), \qquad (1\text{-}7)$$

allowing for the fact that the ratio of the photometric units used in the image and object respectively are so chosen that the intensity in the image of a uniform object has the same numerical value as the intensity in the object itself. T(s) is the [redacted] ratio of [redacted]

the contrast in the image to that in the object, and $\Theta(s)$ is a phase-shift corresponding to a lateral displacement of the image relative to the ideal image. The Optical Transfer Function (O.T.F.) of the lens system is then the complex number defined[13] by

$$D(s) = T(s)\, e^{i\Theta(s)} \qquad (1\text{-}8)$$

given as a function of reduced spatial frequency, s.

Unlike the Strehl ratio, the O.T.F., or D(s), depends on the reduced spatial frequency, s, and azimuth, ϕ, of the sinusoidal grating being considered. Moreover, it requires the condition of isoplanatism to be satisfied, i.e. the aberration of the system must remain constant over the small region of the image plane within which the image of a point source has a significant intensity. The usefulness of the O.T.F. lies in the fact that it completely describes the image forming properties of the system once the response D(s) is known for the neighbourhoods of a typical set of points in the image plane, and for different orientations and reduced spatial frequencies of sinusoidal gratings considered as test objects. This unique property of D(s) is due to the fact that, to obtain the image of any object, one merely needs to analyse the intensity distribution in the object into its Fourier spectrum, multiplying each component by the response for that frequency, and then synthesize the image by means of the sum of these modulated components. Duffieux[14] (1946) first showed that for self-luminous objects, the O.T.F. is simply the Fourier transform of the distribution of intensity in the image of a narrow line source.

We now consider a criterion of image-quality based on the O.T.F. . A tolerance condition, parallel to that of the Strehl ratio, has been developed by Hopkins[15](1957). This makes use of the formula[13] for the O.T.F. in monochromatic light

$$\left.\begin{aligned} D(s) &= \frac{1}{A}\iint_S e^{iksW(x,y;s)}\,dx\,dy \\ W(x,y;s) &= \frac{1}{s}\left[W(x+\tfrac{s}{2},\,y) - W(x-\tfrac{s}{2},\,y)\right] \end{aligned}\right\} \quad (1-9)$$

where s, the reduced spatial frequency, is defined by

$$s = \frac{\lambda}{n.\sin\alpha}\cdot N = \frac{\lambda'}{n'.\sin\alpha'}\cdot N'$$

where α, α' are the angular radii of the circle $x^2+y^2 = 1$ of the entrance and exit pupils respectively, and subtended at the object and image planes respectively; n, n' are refractive indices; and N, N' denote the number of cycles per unit length as measured in the object and image spaces respectively. A is the area of the pupil, $k = \frac{2\pi}{\lambda}$, and S denotes the region common to two "pupil areas" displaced to the points $(\pm\frac{s}{2}, 0)$ respectively. The value of D(s) for an aberration free system is given by

$$D_o(s) = \frac{1}{A}\iint_S dx\,dy \quad . \qquad (1-10)$$

The influence of aberration on the O.T.F. may be described by the relative function

$$M(s)\,e^{i\theta(s)} = \frac{D(s)}{D_o(s)} \qquad (1-11)$$

If $W_o(s)$ is the mean value defined by

$$W_o(s) = \iint_S W(x,y;s)\,dS \qquad (1-12)$$

where $dS = \frac{dx\,dy}{\iint_S dx\,dy}$ is a reduced element of area, it may be shown that, for sufficiently small values of the argument ksW(x,y;s) in (1-9),

$$\theta(s) = \frac{2\pi s}{\lambda}\cdot W_o(s) \qquad (1-13)$$

gives the spatial phase shift, and for these small values of ksW(x,y;s), also

$$M(s) \geqslant \left[1 - \frac{2\pi^2}{\lambda^2} \cdot s^2 \cdot K(s)\right] \qquad (1\text{-}14)$$

where

$$K(s) = \iint_S \left[W(x,y;s)\right]^2 dS - \left[\iint_S W(x,y;s)\, dS\right]^2 \qquad (1\text{-}15)$$

is the variance of the difference function W(x,y;s) defined in (1-9).

The condition for the equality sign to hold in (1-14) with sufficient accuracy is that $M(s) \geqslant 0.80$. W(x,y;s) can be calculated from the wave-front aberration W(x,y) of the optical system by writing

$$sW(x,y;s) = W\left(x + \frac{s}{2}, y\right) - W\left(x - \frac{s}{2}, y\right) \qquad (1\text{-}16)$$

so that , using the Taylor series,

$$W(x,y;s) = W_x'(x,y) + \frac{1}{3!}\left(\frac{s}{2}\right)^2 W_x'''(x,y) + \frac{1}{5!}\left(\frac{s}{2}\right)^4 W_x^{V}(x,y) + \cdots \qquad (1\text{-}17)$$

From this, M(s) can be calculated. For the condition $M(s) \geqslant 0.80$ to hold, the variance K(s) has to satisfy the inequality

$$K(s) \leqslant \frac{0.10\,\lambda^2}{\pi^2 s^2} \qquad (1\text{-}18)$$

and this is the more easily satisfied the smaller the value of s.

It is important to point out that the argument ksW(x,y;s) may be made small even when the aberrations, and consequently W(x,y;s), are large: one merely needs to select a sufficiently small value for s (see Hopkins, 1957, Proc. Phys. Soc. B70, 449). Consequently, Hopkins' criterion may be applied to a much wider range of aberration magnitudes than that is possible with Maréchal's.

In Chapter 3 a brief account will be given of how a double summation of K(s) over the field and azimuths can effectively serve as an index of image quality, parallel to $\sum$ E' s mentioned earlier. One further advantage of using $\sum_{\tau}\sum_{\phi} K(s_{\phi})_{\tau}$ is that it may be extended to cover different s values, suitably weighted, to give a single overall index of image quality.

(c) Ray optical image-assessment.

A qualitative method of image assessment is the use of Spot Diagrams (HAWKINS & LINFOOT, 1946)*. Here each ray intersection in the image plane is plotted as a point, on a suitably enlarged scale, so that the geometric ray image of a point object is reconstructed. A suitably selected number of rays may be traced and from these one can compute the coefficients in the transverse ray aberration polynomials; then by using the latter any number of points in the Spot Diagram can quickly be obtained. This method was used by Herzberger (1947), Stavroudis and Feder[17] (1954), and some remarkable examples of spot diagrams were given by Linfoot[18] (1955) [SEE ALSO FOOT NOTE]

When the aberrations are large qualitative agreement is obtained with the experimentally found diffraction images. Spot Diagrams, however, give only a visual appreciation of the image and are not readily adapted to the formulation of an index of quality.

* D.G. HAWKINS & E.H. LINFOOT, M.N.R.A.S. 105 (1946) 334-344.

(1-iv) Lens Design by Computer.

Since the advent of high-speed computers, workers in the field of optical design have been intrigued by the idea of automating the lens design process[30]. The tracing of rays through an optical system, which is a fundamental part of any optical designing and is a relatively simple iterative process which has to be repeated a great number of times, is obviously suitable for programming on an electronic digital computer and it is to this that these machines were first applied. The ideas on differential correction techniques, first developed by McAuley and Cruikshank[19](1945), could readily be incorporated into a design programme. The automatic fine correction of optical systems is based on the approximate expression

$$W(p_1+\delta p_1,\ p_2+\delta p_2,\ \ldots\ldots p_n+\delta p_n)=W(p_1,p_2\ldots\ldots p_n) + \sum_{i=1}^{n} \frac{\partial W}{\partial p_i}\ \delta p_i \qquad (1\text{-}19)$$

for the aberration along any ray, expressed as a function of the constructional parameters $p_1, p_2, \ldots\ldots p_n$. The differential correction method mentioned above required the calculation of the derivatives of the classical ray-aberrations by differentiating ray tracing formulae, and then a set of transfer formulae had to be used to obtain the effect of a parameter change on the ray-aberration in the image space. Due to the storage limitation of earlier computers, programmes were developed in which the calculation of the aberration changes of each of the

chosen rays for changes of the various design parameters were obtained by "finite differences" method. This consisted of altering each curvature, separation, or glass, one at a time, by a small amount, altering the next following surface to restore the paraxial ray to its previous condition before the change. In each case the ratio of increments $\delta W/\delta p$ was used as an approximation for the derivative $\partial W/\partial p$. Improved differential correction techniques, based on the use of optical path method, have recently been developed by Hopkins. These new techniques (see Chapter 6) are particularly suitable for automatic lens design on a large storage computer.

The first serious attempt at automatic optical design was made in 1951 by a study group at Harvard University[20]. Another method of automatic optical designing has since been described by Girard[21] (1958) and Wynne[22] (1959). The method used is an iterative one, using a linear approximation to the change of aberrations with design parameters (see equation 1-19). Since, in fact, the dependence of the aberrations of an optical system on the design parameters is highly non-linear if large changes are made, the linear approximation can only be used for steps of the iteration of limited size. Of the numerous optimization methods proposed for the treatment of automatic lens-design problems, the damped least-squares approach first used by Girard (1958) seems to have been the one most frequently applied by optical designers[21-28].

(1-v) A General Linearization Method - statement of the problem.

The formulation of the optical design problem may be summarized as follows:

Let $W_1, W_2 \ldots\ldots W_m$ be a set of m weighted aberrations, which are functions of n design parameters $p_1, p_2, \ldots\ldots p_n$. By expanding into a Taylor series and retaining the zeroth- and first-order terms only, we obtain the approximations

$$Wi = \sum_{j=1}^{n} A_{ij} \cdot p_j + W_{io} \qquad i=1,2\ldots\ldots m \qquad (1\text{-}20)$$

where $$A_{ij} = \frac{\partial Wi}{\partial p_j} \qquad (1\text{-}21)$$

and W_{io} are the values of W_i for the initial point (p_{jo}), which is considered to be the origin of our coordinate system.

The conditions for the elimination of the aberrations are thus expressed by a set of m linear equations of n unknowns:

$$\sum_{j=1}^{n} A_{ij}\, p_j + W_{io} = o \qquad i=1,2\ldots m \qquad (1\text{-}22)$$

Since $m > n$, we apply the least squares method which solves for those p_j which minimize the merit function

$$\phi = \sum_{i=1}^{m} \left[\sum_{j=1}^{n} A_{ij}\, p_j + W_{io} \right]^2 = \sum_{i=1}^{m} W_i^2 \qquad (1\text{-}23)$$

To minimize (1-23) we set the partial derivatives $\partial\phi/\partial p_j$ equal to zero, obtaining a system of n equations of n unknowns. These normal equations may be written:

$$[Aji]\cdot[Aij]\cdot p_{(j} + [Aji]\cdot W_{(io}=0 \qquad (1\text{-}24)$$

where $[Aji]$ is the transpose of $[Aij]$,

$P_{(j}$ is the column vector $\{p_1 p_2 \cdots p_j \cdots p_n\}$

$W_{(io}$ is the column vector $\{W_{10}\ W_{20} \cdots W_{io} \cdots W_{mo}\}$

Because of the assumption made in writing equations (1-20), the solution of equations (1-24) can easily lead to a system which is worse corrected than the initial one. This is because , for larger parameter changes the aberration differentials are not linear, and can only be reasonably so considered for small values of each p_j, while the normal equations obtained in the process of lens design usually have their determinant very small, resulting in a solution having large values of p_j. The ill-conditioned equations arise from the nearly linear dependence of the aberration equations[22].

The above difficulties can be overcome by applying the concept of least squares to the aberration functions W_i and the independent variables p_j simultaneously, thus keeping the absolute value of both the residuals and the increments small. This approach, known as the damped least-squares method, was suggested by Levenberg[29] (1944) and applied to lens design by Girard[21] and Wynne[22].

The Method of Damped Least-Squares.

This consists of adding the sum

$$S^2 \sum_{j=1}^{n} C_j^2 \cdot p_j^2 \qquad (1\text{-}25)$$

to the merit function ϕ , where S and C_j are weighting

NOTE

* In this context, of course , 'minimizing' implies only seeking a local minimum in the neighbourhood of the starting point. Ψ is a complicated function of a large number of variables, and is to be expected to have many such local minima.

factors. The effect of these is to afford a control over the influence of different variables on the solution of the normal equations and provide a means of limiting the magnitude of the computed changes in the system parameters. Instead of minimizing the function ϕ at each step, we now minimize a new merit function ψ defined as

$$\psi = \sum_{i=1}^{m} W_i^2 + S^2 \sum_{j=1}^{n} C_j^2 \, p_j^2 \qquad (1\text{-}26)$$

Thus by equating $\partial\psi / \partial p_j$ to zero [SEE NOTE, LH.S.] we obtain the damped normal equations

$$[A_{ji}] \cdot [A_{ij}] \cdot p_{(j} + [A_{ji}] \cdot W_{(io} + S^2 C \, p_{(j} = o \qquad (1\text{-}27)$$

where C is a diagonal matrix, depending on the definition of C_j :

(i) if all the system parameters are weighted equally then

$$C_j = 1 \quad , \; 1 \leqslant j \leqslant n \qquad (1\text{-}28)$$

and C is a unit diagonal matrix.

(ii) if each of the system parameters p_j is weighted differently according to the "sensitivity" of that particular parameter on the aberrations, then one uses the definition suggested by Meiron[23], i.e.

$$C_j = \left[\sum_{i=1}^{m} A_{ij}^2 \right]^{\frac{1}{2}} \quad , \; 1 \leqslant j \leqslant n \qquad (1\text{-}29)$$

and the diagonal elements of C are then C_j^2.

General Requirements of a Design Programme.

The essential elements of a design programme may

be summarized as follows:

(i) Input A. Design specifications: including the focal length, aperture, field coverage, number of elements, minimum centre and edge thicknesses, spectral range etc.

Input B. Designer's choice: including the data from a partially completed design, to be used as the initial starting point, the target values of aberrations; the weighting factors, assigning the relative importance of various image defects, such as distortion, chromatic aberration etc.; boundary conditions restricting the range of design parameters; the ray-tracing pattern at the entrance pupil etc.

(ii) Working stage. Start systematic ray-tracing, calculate the first differential of each weighted aberration W_i with respect to each of the design parameters p_j and store a matrix of differentials $\frac{\partial W}{\partial p_j} i = A_{ij}$

Deduce the largest step length S in parameter space with good linearity, set up a new system corresponding to the p_j values found, calculate its actual ψ, recompute the matrix A_{ij} for the new improved system, and iterate the procedure.

Note that the value of S to be selected must satisfy the basic requirement that the resulting new merit function ψ must always be smaller than the initial ψ_0.

Boundary condition must also be strictly observed and these may involve the removal of appropriate lines from the A_{ij} matrix corresponding to changes that have produced violations, resulting in a reduced matrix.

(iii) Output. A completed design satisfying both the design specifications and the performance criterion - i.e. an optimized design.

The term "optimized" design implies a particular definition of the merit function ψ, which, in turn, implies the choice of rays and aberrations to be calculated, of weights to be attached to them, and of course, the image criterion to be used. Many authors have used the sums of the squares of the transverse ray aberrations for this purpose (as in equation 1-26). In the present work the merit function used depends on the sum of the suitably weighted variances of wave aberrations (as mentioned in 1-iiiA), and the Strehl intensity ratio is employed as the image criterion for high quality systems. By making use of new canonical pupil coordinates it is possible to use the variance of wave aberration effectively as a guide in the process of improving the performance at the design stage. This requires also the derivatives of the wave aberration to be used.

The canonical pupil coordinates mentioned above play a vital role in both our ray-tracing and optimization stages. A detailed account of these variables, and how they form an essential part of this work, will be found in the following chapter.

REFERENCES

(1) A.E. Conrady, Applied Optics and Optical Design, Pt. 1, Dover Pub. Inc. (1957).

(2) C.G. Wynne, Optica Acta, Vol. 8, No. 3, 255 (1961)

(3) Seidel, Astr. Nachr. 43, 289-332 (1856)
E.T. Whittaker, The Theory of Optical Instruments, Cambridge (1907).
H.H. Hopkins, Wave Theory of Aberrations, Oxford (1950)

(4) R. Glazebrook, A Dictionary of Applied Physics, Vol. IV, London (1923).

(5) W.B. King, First Order Aberrations with Particular Reference to a Monocentric Optical System, Thesis, London University (1963).
W.B. King, Optica Acta, Vol. 11, No. 2, 107-11 (1964).

(6) M. Herzberger, J.O.S.A., 34, 114 (1944).

(7) T. Smith, Proc. Phy. Soc., 57, 286-93 (1945).
W. Weinstein, Proc. Phy. Soc., 65B, 731-35 (1952).

(8) A. Cox, A System of Optical Design, Focal Press, London (1964). See the section on Lens Performance.

(9) B.K. Johnson, Optical Design and Lens Computation, London (1948), Chapter 1.

(10) K. Strehl, Zeitshrift fur Instrumentenkunde, 22, 213 (1902).

(11) A. Marechal, Rev. Opt. (Theor. Instrum.), 26, 257 (1947).

(12) H. Frieser, Kinotechnik, 17, 165 (1935).

(13) H.H. Hopkins, Proc. Phys. Soc., 79, 889-919 (1962).

(14) P.M. Duffieux, L'integrale de Fourier et ses applications a l'optique (Besancon: privately printed, 1946).

(15) H.H. Hopkins, Proc. Phys. Soc., B70, 449 (1957).

(16) M. Herzberger, J.O.S.A., 37, 485 (1947).

(17) O.N. Stavroudis and D.P. Feder, J.O.S.A., 44, 163, (1954).

(18) E.H. Linfoot, Recent Advances in Optics, Oxford (1955); D.G. HAWKINS & E.H. LINFOOT, M.N.R.A.S. 105 (1946) 334-344.

(19) A.L. McAuley and F.D. Cruikshank, Proc. Phys. Soc. 57, 302 (1945).

(20) J.G. Baker, Technical Reports 1-12, Perkin-Elmer Corp. (U.S. Air Force Contract, 1951-55).

(21) A. Girard, Rev. Opt., 37, 225, 397, (1958).
(22) C.G. Wynne, Proc. Phys. Soc., 73, 777 (1959).
(23) J. Meiron, J.O.S.A., 49, 293 (1959); 55, 1105 (1965)
(24) D.S. Grey, J.O.S.A., 53, 672 (1963).
(25) L. Fialovsky, Optica Acta, 10, 341 (1963).
(26) G.H. Spencer, App. Optics, 2, 1257 (1963).
(27) B. Brixner, App. Optics, 2, 1281 (1963).
(28) K.B. O'Brien, J.O.S.A. 54, 1252 (1964).
(29) K. Levenberg, Quart. App. Math. 2, 164 (1944).
(30) D.P. Feder, App. Optics, Vol. 2, 12 1209-1226 (1963)

Chapter 2.

(II-1) Canonical Pupil Co-ordinates in Geometrical and Diffraction Image Theory.

The canonical pupil co-ordinates in geometrical and diffraction theory were first introduced by Hopkins in a paper presented at the Tokyo Optical Conference in September 1964. Due to the very limited circulation of the proceedings, it seems appropriate that this chapter should begin with a summary of Hopkins' paper.

In Figure 2-1, O and O' are axial object and image points; Q' is the gaussian image of Q. E and E' are the paraxial entrance and exit pupil points, and EB, E'B' are the respective pupil spheres. The ray OB emerges along B'O". The system is free from coma near the axis if the sine-condition is satisfied, i.e.

$$\frac{\sin\Theta'}{u'} = \frac{\sin\Theta}{u} \qquad (2\text{-}1)$$

where u, u' are paraxial angles. If R=EO, R'=E'O', then the paraxial ray heights at E,E' are h=Ru and h'=R'u'. Also from fig. 2-1, Y'=R' sinΘ', Y=R sinΘ. The relation (2-1) may thus be written as

$$y' = \frac{Y'}{h'} = \frac{Y}{h} = y \qquad (2\text{-}2)$$

and remains valid despite of the presence of pupil aberrations.

When the system is not free from coma, then for the ray OBB'O"

$$W'_{coma} = H\left\{\frac{\sin\Theta'}{u'} - \frac{\sin\Theta}{u}\right\} = H\,(y'-y) \qquad (2\text{-}3)$$

where $H = nu\eta = n'u'\eta'$ is the Lagrange invariant; $\eta = OQ$, $\eta' = O'Q'$ are gaussian object and image heights.

If a small change in image height δH results in a change δW in coma, then from (2-3)

$$y' - y = \frac{\delta W}{\delta H} = \delta y' \qquad (2\text{-}4)$$

gives the "transverse aberration" of the ray at the exit pupil sphere. Corresponding to the above, we have

$$x' = \frac{X'}{h'} \quad , \quad x = \frac{X}{h} \qquad (2\text{-}5)$$

and $$G = nu\xi = -n'u'\xi' \qquad (2\text{-}6)$$

The canonical relations between pupil and image plane variables take the forms

$$\delta x' = \frac{\delta W}{\delta G} \quad , \quad \delta y' = \frac{\delta W}{\delta H} \qquad (2\text{-}7)$$

$$\delta G' = -\frac{\delta W}{\delta x'} \quad , \quad \delta H' = -\frac{\delta W}{\delta y'} \qquad (2\text{-}8)$$

(2-7) coming from (2-4), and (2-8) being the known formulae for $\delta G'$ and $\delta H'$, the transverse ray aberrations in the image plane.

Now consider the pupil co-ordinates for an extra-axial image. In Fig. 2-2 we show (for convenience only) a meridian ray, $\bar{Q}BB'$, and $\bar{Q}, \bar{Q}'$ are the points $(\bar{\xi}, \bar{\eta})$, $(\bar{\xi}', \bar{\eta}')$ respectively. A pupil ray from Q (close to $\bar{Q}$) cuts the image plane at Q'. The difference in optical path length between the ray from Q to B' and the ray $\bar{Q}BB'$ results in a difference in wave front aberration for the two rays given by

$$W_{Q'} - W_{\bar{Q}'} = [D'B'] - [DB]$$

$$= \frac{n'x'\xi'}{\bar{R}'} - \frac{nx\xi}{\bar{R}} + \frac{n'y'\eta'}{\bar{R}'} - \frac{ny\eta}{\bar{R}} \qquad (2\text{-}9)$$

where (ξ, η) and (ξ', η') are the co-ordinates of Q and Q', relative to $\bar{Q}$ and $\bar{Q}'$ as origins; (x,y), (x',y') are rectangular co-ordinates of the points B,B'; and $\bar{R} = \bar{E}\bar{Q}$, $\bar{R}' = E'Q'$. We now define the pupil variables for extra-axial image points by

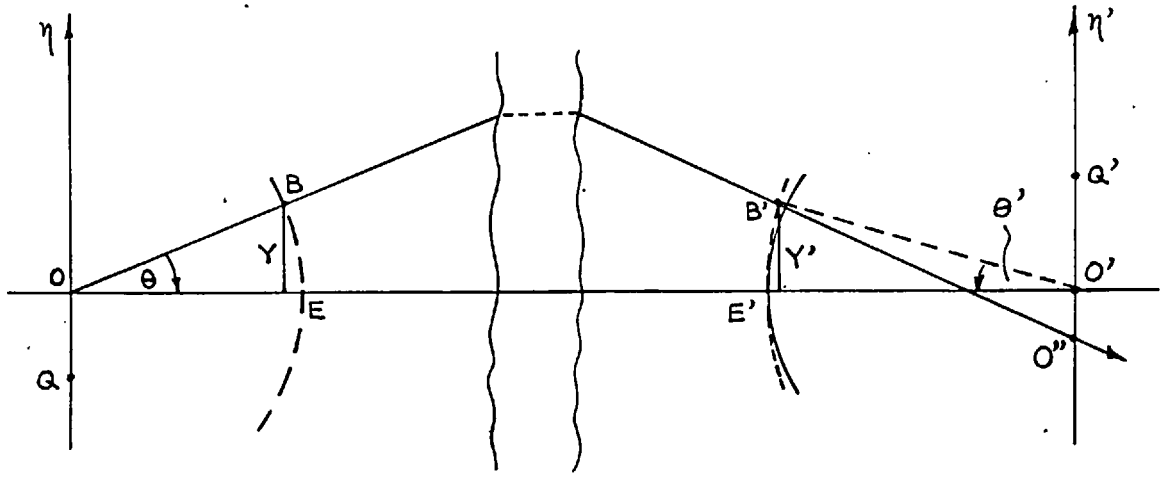


Fig. 2-1 PUPIL CO-ORDINATES FOR THE AXIAL IMAGE

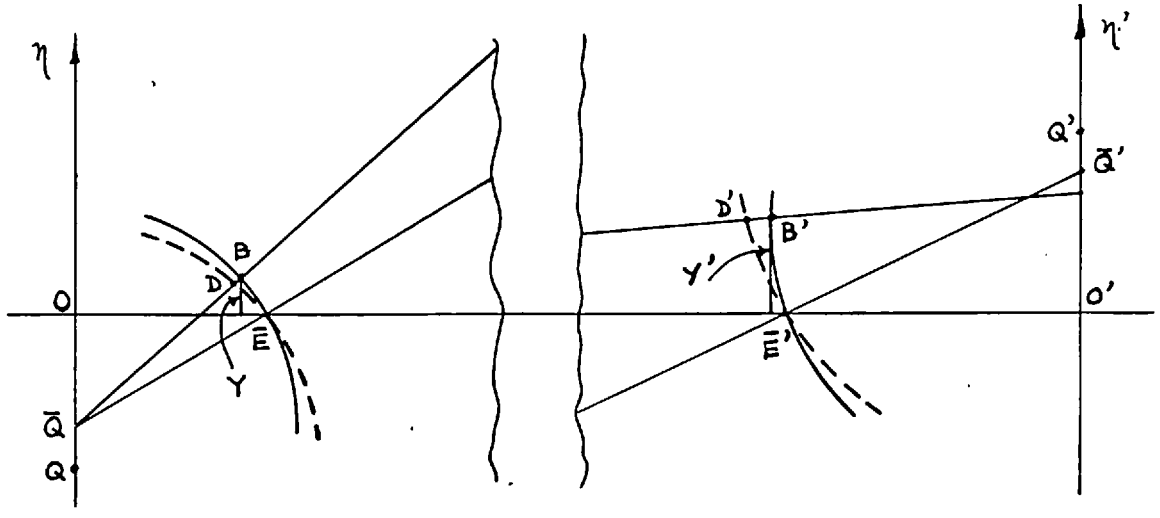


Fig. 2-2 PUPIL CO-ORDINATES FOR AN EXTRA-AXIAL IMAGE

$$x_s = \frac{X}{h_s}\ ,\ x'_s = \frac{X'}{h'_s}\ ;\ X' = h'_s\ x_s + \mathbf{\underline{O}}\,(x_s^2)$$

$$y_T = \frac{Y}{\bar{N}h_T}\ ,\ y'_T = \frac{Y'}{\bar{N}'h'_T}\ ;\ Y' = (\bar{N}'h'_T)\ y_T + \mathbf{\underline{O}}\,(y_T^2) \qquad (2\text{-}10)$$

where h_s, h'_s are the paraxial ray heights at $\bar{E}, \bar{E}'$ of a close sagittal ray; h_T, h'_T are those of a close tangential ray, and $\bar{N}, \bar{N}'$ are direction cosines with respect to the optical axis of $\bar{Q}\bar{E}$ and $\bar{Q}'\bar{E}'$. Substituting the above pupil variables in (2-9), and making use of aperture angles defined by

$$\sin\alpha_s = \frac{h_s}{\bar{R}}\ ;\ \sin\alpha_T = \frac{h_T}{\bar{R}} \qquad (2\text{-}11)$$

we have

$$W_{Q'} - W_{\bar{Q}'} = \left\{(n' \sin\alpha'_s)\,\xi' - (n \sin\alpha_s)\,\xi\right\} x_s + \mathbf{\underline{O}}\,(x_s^2)$$
$$+\left\{(n'\bar{N}' \sin\alpha'_T)\,\eta' - (n\bar{N} \sin\alpha_T)\,\eta\right\} y_T + \mathbf{\underline{O}}\,(y_T^2) \qquad (2\text{-}12)$$

For a limitingly close aperture ray then

$$W_{Q'} - W_{\bar{Q}'} = \mathbf{\underline{O}}\,(x_s^2) + \mathbf{\underline{O}}\,(y_T^2)$$

and (2-12) shows the invariance of G_s and H_T defined by

$$G_s = (n \sin\alpha_s)\,\xi = (n' \sin\alpha'_s)\,\xi' = G'_s$$
$$H_T = (n\bar{N} \sin\alpha_T)\,\eta = (n'\bar{N}' \sin\alpha'_T)\,\eta' = H'_T \qquad (2\text{-}13)$$

If (G_s, H_T) are used now as reduced co-ordinates for Q, Q', equation (2-9) may be written as

$$W_{Q'} - W_{\bar{Q}'} = G_s\,(x'_s - x_s) + H_T\,(y'_T - y_T) \qquad (2\text{-}14)$$

or, for small values of G_s, H_T,

$$x'_s - x_s = \frac{\partial W}{\partial G_s}\ ;\ y'_T - y_T = \frac{\partial W}{\partial H_T} \qquad (2\text{-}15)$$

corresponding exactly to (2-4). The reduced transverse

ray aberration in the image plane can be shown to be given by

$$\delta G'_s = -\frac{\partial W}{\partial x'_s} \quad ; \quad \delta H'_T = -\frac{\partial W}{\partial y'_T} \tag{2-16}$$

If the aberration is stationary with image height (2-15) gives

$$x'_s = x_s \quad ; \quad y'_T = y_T \tag{2-17}$$

The local sagittal and tangential magnifications , M_s and M_T, at the image point $\bar{Q}'$ are defined by equation (2-13) as

$$M_s = \frac{\xi'}{\xi} = \frac{n \sin \alpha_s}{n' \sin \alpha'_s} \quad ; \quad M_T = \frac{\eta'}{\eta} = \frac{n\bar{N} \sin \alpha_T}{n'\bar{N}' \sin \alpha'_T} \tag{2-18}$$

When the derivative of the distortion with respect to image height at $\bar{Q}'$ is not zero, the anamorphotic ratio ρ' defined by

$$\rho' = M_T / M_s \tag{2-19}$$

will not equal to unity. Combining (2-18) and (2-19) and making use of (2-11), one has the important result

(2-20)

$$\rho' = \frac{(h'_s / h_s)}{(\bar{N}' h'_T / \bar{N} h_T)} = \frac{\text{Sagittal Pupil Magnification}}{\text{Tangential Pupil Magnification}}$$

Hopkins (1964) has shown[2] that when $\rho' \neq 1$, the scale of pupil is contracted or dilated in the y' direction in the inverse ratio with which the scale of image is dilated or contracted in the η' direction. However, this scaling difficulty is eliminated by

defining the pupil and image variables in the form given in (2-10) and (2-13), and the results given in (2-17) are valid even in the presence of large amounts of distortion and consequent distortion of the pupil shape.

Defining the image variables with the wavelength as unit of length we have

$$u_S = \frac{G_S}{\lambda} \quad ; \quad v_T = \frac{H_T}{\lambda} \tag{2-21}$$

and the Fraunhofer diffraction integral then has the form

$$F'(u_s', v_T') = \iint_{\text{Pupil}} f(x'_s, y'_T) \exp\left\{ i2\pi (u_s' x_s' + v_T' y_T') \right\} dx_s' \, dy_T' \tag{2-22}$$

For a reasonably corrected system, the results given in (2-17) are valid with good approximation※, and (2-22) may then be expressed in terms of object space variables:

$$F(u_s, v_T) = \iint_{\text{Pupil}} f(x_s, y_T) \exp\left\{ i2\pi (u_s x_s + v_T y_T) \right\} dx_s \, dy_T \tag{2-23}$$

where the pupil function

$$f(x_s, y_T) = \exp.\left\{ i2\pi \; W(x_s, y_T) \right\} \tag{2-24}$$

uses the wavefront aberration as a function of the entrance pupil co-ordinates. When (2-23) is used in the transfer function theory, $u_s' = u_s$ and $v_T' = v_T$ give the image of (u_s, v_T) even though the anamorphotic ratio at $\bar{Q}'$ departs from unity. It may be noted that, according to (2-15), the degree to which $x_s' = x_s$, $y_T' = y_T$ determines the degree to which the isoplanatic conditions hold, and hence the degree to which the application of transfer function theory to image formation is valid.

※ See, for example, pp 303A, 303B.

(II-ii) The Applications of Canonical Co-ordinates in Lens Design.

A detailed account of the ray tracing methods based on the variables introduced in (II-i) will be found in Chapter 4. It is necessary here to introduce new pupil variables which are normalised by $h_s = \bar{N}h_T$. We define:

$$y_s = \frac{Y}{h_s} = \left(\frac{Y}{\bar{N}h_T}\right) = y_T = y_T' = \left(\frac{Y'}{\bar{N}'h_T'}\right) \qquad (2\text{-}25)$$

but in general,

$$y_s' = \frac{Y'}{h_s'} \neq y_T' \qquad (2\text{-}25)A$$

This is because when $h_s = \bar{N}h_T$, the expression for ρ' from (2-20) is

$$\rho' = \left(\frac{h_s'}{\bar{N}'h_T'}\right)$$

so that $h_s' \neq \bar{N}'h_T'$ unless the anamorphotic ratio ρ' is unity.

Having introduced the variable y_s', we may now write an alternative expression for (2-16) as

$$\delta G_s' = -\frac{\partial W}{\partial x_s'} \; ; \; \delta H_s' = -\frac{\partial W}{\partial y_s'} \qquad (2\text{-}26)$$

where H_s' is aperture normalised in the same way as G_s', i.e.

$$H_s' = n' \sin \alpha_s' \eta' \qquad (2\text{-}26)A$$

An introductory account is given below of some useful results arising from the use of the canonical pupil variables x_s, y_s.

As a first step in the lens design process the vignetted pupil shape for each field point is deter-

mined by using the method to be described in Chapter 4. It has been found that the vignetted pupil can be well represented by a best-fitting ellipse

$$\left\{\frac{x_s}{a_s}\right\}^2 + \left\{\frac{y_s}{b_s}\right\}^2 = 1 \qquad (2\text{-}27)$$

where a_s and b_s are the major and minor semi-axes corresponding to the particular field considered. Next we scale this ellipse into an equivalent circle,

$$x^2 + y^2 = 1 \qquad (2\text{-}27)A$$

by defining

$$x = \frac{x_s}{a_s} \; ; \; y = \frac{y_s}{b_s} \qquad (2\text{-}28)$$

The usefulness of this equivalent circular pupil lies in the fact that one may now adopt a standard ray pattern with prescribed values of (x,y) at the entrance pupil. The actual pupil co-ordinates used for tracing a ray are (x_s,y_s), which are simply the values of (x,y) scaled down by a_s and b_s. This facilitates an analytical formulation of the general problem of extra-axial image formation. The aberration is now regarded as a polynomial in (x,y), but only as an intermediate step. The aberration coefficients will be the true coefficients in a corresponding polynomial in the variables (x_s,y_s). Corresponding to (2-23) and (2-24) we now have

$$F(u,v) = a_s b_s \iint_{\substack{\text{circular} \\ \text{pupil}}} \exp\left\{ikW(x,y)+i2\pi\,(ux+vy)\right\} dx\, dy \qquad (2\text{-}29)$$

where x and y are as defined by (2-28), and

$$u = a_s u_s = a_s u_s'$$
$$v = b_s v_T = b_s v_T'$$

and note also that (2-30)

$$y = \frac{y_s}{b_s} = \frac{y_T}{b_s} = \frac{y_T'}{b_s}$$

If the aberration polynomial is taken to be of the form

$$W = W(x,y) = \sum_n \sum_q W_{n,q} (x^2 + y^2)^n y^q \qquad (2\text{-}31)$$

and if there are J aberration coefficients $W_{n,q}$ to be considered, then this number of rays must be traced, giving the same number of values of the aberration W_j, $j=1,2,\ldots,J$, each represented by (2-31). We then have J equations with J unknowns $W_{n,q}$. By inverting the matrix we have for each aberration coefficient

$$W_{n,q} = \sum_{j=1}^{J} K(n,q;j)\, W_j \qquad (2\text{-}32)$$

and, for a given polynomial, the coefficients $K(n,q;j)$ may be calculated once for all (see Chapter 5). It is essential that the ray pattern used for ray tracing, on which $K(n,q;j)$ depends, must be so chosen that the equations from (2-31) are not ill-conditioned.

We consider now the case with the image criterion based on the Strehl ratio. The coefficients $W_{n,q}$ are to be regarded as intermediate auxiliary quantities: by substituting these in (2-31) and then using the resulting $W(x,y)$ to evaluate the variance E (equation 1-5), we have the double sum expression

$$E = \sum_{i=1}^{J} \sum_{j=1}^{J} P(i,j)\, W_i W_j \qquad (2\text{-}33)$$

where the coefficients $P(i,j)$ are constants (see Chap-

ter 5) once the ray pattern and the aberration polynomial have been selected. Thus, when the aberrations W_j from ray tracing are known, E may be at once calculated merely by substitution of their values in the quadratic expression (2-33).

As mentioned in I-iiiA of Chapter 1, one requires to minimise the value of E with respect to design parameters p and this can need a knowledge of $\partial E / \partial p$. Differentiating (2-33) with respect to p we have

$$\frac{\partial E}{\partial p} = \sum_{i=1}^{J} \sum_{j=1}^{J} p(i,j) \left\{ W_i \frac{\partial W_j}{\partial p} + W_j \frac{\partial W_i}{\partial p} \right\} \quad (2\text{-}34)$$

The minimisation procedure using (2-34) in the design stage and the iterative sequence that follows will be described in Chapter 8. When E has been reduced to the desired order of magnitude E_o, the performance of the design may then be checked by calculating its Strehl intensity ratio I from Maréchal's approximate expression (see equation 1-4)

$$I = \left[1 - \frac{2\pi^2}{\lambda^2} E_o \right]^2$$

and this must exceed 0.8 to be acceptable. An accurate value of I may be found by evaluating (1-2), using a numerical integration technique to be described in the following chapter.

The technique described above for finding the parameter derivatives of E may be applied also to the optical transfer function tolerance theory [3]. Here the variance K(s) defined by (1-15) for the difference function W(x,y;s) must be used [4]. The expression for $\frac{\partial K(s)}{\partial p}$ corresponding to (2-34) is

$$\frac{\partial K(s)}{\partial p} = \sum_{i=1}^{I} \sum_{j=1}^{I} P_s(i,j) \left\{ W_i \frac{\partial W_j}{\partial p} + W_j \frac{\partial W_i}{\partial p} \right\} \quad (2\text{-}35)$$

It is thus possible to design a microscope objective with very good axial correction using the Strehl tolerance technique and with the extra-axial image points optimised for a particular spatial frequency s using K(s).In this manner a best overall correction can be acheived.

In addition to the above applications, the canonical pupil co-ordinates are also employed in aberration theory and in differential correction formulae (see Chapter 5). These pupil co-ordinates are essential if transfer functions are to be used as image quality criteria in lens design, because these are the variables in which the Fourier treatment of image formation is valid. Finally, the new co-ordinates system has advantages over the traditional approach in ray-tracing: they always remain finite, and lead to both simple and perfectly general opening and closing formulae (see Chapter 4).

REFERENCES

(1) H.H. Hopkins, Wave Theory of Aberrations, Oxford (1950).
(2) H.H. Hopkins, I.B.M. Research Report RC-1242, (1964).
(3) N.J. Davidson, Thesis, London University, (1966).
(4) H.H. Hopkins, Proc. Phys. Soc., 3, Vol.LXX, P.449 (1957).
H.H. Hopkins, "Canonical Pupil Co-ordinates in Geometrical and Diffraction Image Theory", Tokyo Optical Conference 1964.

CHAPTER 3

THE CALCULATION OF THE STREHL INTENSITY RATIO, I, BY MEANS OF A NUMERICAL INTEGRATION METHOD.

Introduction

The computer programme to be described below calculates the Strehl Intensity Ratio, I, when the aberration coefficients $W_{n,q}$ are known. One obvious application of the programme is that the influence on the intensity ratio I of different aberrations may be systematically investigated (see section III-V). The minimization procedure, to be described in Chapter 8, is concerned with how to minimize the variance E of the aberration of an optical system; it is therefore useful to know, at one particular stage of design, what is the significance of the value of I as calculated from the Maréchal's approximate expression (1-4) when the value of E is not sufficiently small for $I > 0.80$. In particular, when the Strehl intensity is greater than about 0.7, its value increases monotonically with decreasing values of the variance E. If the variance is to be used as the criterion of image quality, therefore, it is desirable to know whether there exist conditions under which a decrease in the value of E leads to a decrease in the Strehl intensity. This could, for example, possibly happen when the Strehl intensity is very small. Such behaviour would not rule out the use of E as a quality criterion for well-corrected systems, but it could reduce the range of states of poorer correction over which E is a useful criterion of image quality. It is with

this question in view that the present detailed study of the Strehl intensity ratio has been undertaken.

(III-i) The use of equivalent circular pupil in diffraction theory.

For the present purpose, the wave-aberration of any image point is represented by a suitably chosen polynomial in the co-ordinates (x,y) of the equivalent circular pupil (see 2-27A). The aberrations W_j are calculated by ray-tracing using the pupil co-ordinates (x_s,y_s). In consequence the aberration coefficients $W_{n,q}$ in the polynomial have no direct significance in analytical aberration theory - except when $x = x_s$, $y = y_s$ corresponding to the axial image case (see Chapter 5). The advantage of the procedure lies in that it reduces all the extra-axial image cases to the problem of a circular aperture; and the use of $W_{n,q}$ facilitates directly the study of the diffraction images of extra-axial point sources.

Referring to the expression of the Strehl Intensity Ratio, I, as given by (1-1), i.e.,

$$I = \frac{\left| \iint_{\text{pupil}} \exp[ikW(x,y)]\, dxdy \right|^2}{\left| \iint_{\text{pupil}} dxdy \right|^2} \qquad (3\text{-}1)$$

and for the present investigation the argument of the exponential function in (3-1) may be replaced by $ikfW(x,y)$; the quantity f being an aberration scaling factor (see later).

Using polar co-ordinates, the expression for I becomes:

$$I = \frac{1}{A^2} \cdot \left| \iint_{\text{pupil}} \exp\left[ikfW(r,\phi)\right] r\,dr\,d\phi \right|^2 \qquad (3.2)$$

where $k = 2\pi$ if the aberrations are measured in the units of wavelength; A being the area of the pupil (see later). The wave-aberration $W(r,\phi)$ may be expressed as a polynomial as follows :

$$\begin{aligned} W(r,\phi) = {} & W_{2,0}r^2 + W_{4,0}r^4 + W_{6,0}r^6 + W_{8,0}r^8 + W_{10,0}r^{10} \\ & + W_{2,2}r^2\cos^2\phi + W_{4,2}r^4\cos^2\phi + W_{4,4}r^4\cos^4\phi \\ & + W_{1,1}r\cos\phi + W_{3,1}r^3\cos\phi + W_{5,1}r^5\cos\phi \\ & + W_{7,1}r^7\cos\phi + W_{9,1}r^9\cos\phi + W_{3,3}r^3\cos^3\phi \end{aligned} \qquad (3.3)$$

with r=1 for the aperture ray. The choice of the aberration coefficients $W_{n,q}$ will be described in detail in Chapter 5.

(III-ii) <u>The numerical evaluation[1] of the Strehl Intensity Ratio I.</u>

Consider the double integrals given by

$$\iint_{\text{pupil}} \exp\left[ikfW(r,\phi)\right] r\,dr\,d\phi \qquad (3.4)$$

and put $t = r^2$, when (3-4) may be written as

$$\tfrac{1}{2} \iint_{\text{pupil}} \exp\left[ikfW(t,\phi)\right] dt\,d\phi \qquad (3.5)$$

When the argument of the exponential function in (3.5) is increasing or decreasing sufficiently rapidly, the precise

form of its variation with t and ϕ is of little consequence. Over a small enough region of the (t,ϕ) plane, therefore, it will be the first two terms in the Taylor expansion of $kfW(t,\phi)$ that determine the contribution of that region to the value of the integral. Let the region be an elementary area defined by the lines

$$t = t_p \pm \alpha_t \quad , \quad \phi = \phi_q \pm \alpha_\phi \tag{3.6}$$

Then, to the order of approximation contemplated, the value of the integrand of (3.5) over the element will be given by

$$\tfrac{1}{2} \exp\left[ikfW(t_p,\phi_q)\right] \int_{t_p-\alpha_t}^{t_p+\alpha_t} \exp\left[ikfW_t(t_p,\phi_q)(t-t_p)\right] dt$$

$$\times \int_{\phi_q-\alpha_\phi}^{\phi_q+\alpha_\phi} \exp\left[ikfW_\phi(t_p,\phi_q)(\phi-\phi_q)\right] d\phi$$

i.e. by

$$\tfrac{1}{2} \exp\left[ikfW(t_p,\phi_q)\right] \quad \frac{2\sin\{\alpha_t kfW_t(t_p,\phi_q)\}}{\{kfW_t(t_p,\phi_q)\}}$$

$$\cdot \frac{2\sin\{\alpha_\phi kfW_\phi(t_p,\phi_q)\}}{\{kfW_\phi(t_p,\phi_q)\}}$$

If the pupil area is considered to be the sum of the

elementary areas bounded by the lines defined in (3.6), the value of the integral (3.5) is given by

$$2.\sum_{p}\sum_{q} \exp\left[ikfW(t_p,\phi_q)\right].\frac{\sin\{\alpha_t kfW_t(t_p,\phi_q)\}}{\{kfW_t(t_p,\phi_q)\}} \cdot\frac{\sin\{\alpha_\phi kfW_\phi(t_p,\phi_q)\}}{\{kfW_\phi(t_p,\phi_q)\}} \qquad (3.7)$$

The value of t in (3.6) has a maximum of unity, corresponding to r=1 in the equivalent circular pupil. Let this unit length be divided into N equal parts so that $p = 1,2,3,\ldots,N$, with p=o at the origin (t=o). The length t_p in (3.6) may then be expressed as :

$$t_p = \frac{p-\frac{1}{2}}{N} = \frac{2p-1}{2N} \qquad (3.8)$$

The quantity α_t defined in (3.6) is equal to half the length of one division; that is :

$$\alpha_t = \tfrac{1}{2}\left(\tfrac{1}{N}\right) \qquad (3.9)$$

For the general case, we regard the aperture as being annular with $t_n = \frac{n}{N}$ as the lower limit of the clear aperture, corresponding to p = n. For most systems, the obstruction factor is zero, corresponding to the case n=o. The summation for p covers the range $(n+1) \leqslant p \leqslant N$. Using this notation, the effective area A of the pupil in (3.2) is given by

$$A = \int_{r=t_n^{\frac{1}{2}}}^{r=1} \int_{\phi=o}^{2\pi} r \, dr d\phi = \pi (1-t_n) \qquad (t_n = \frac{n}{N}) \qquad (3.10)$$

Substituting (3.10) and (3.7) in (3.2) we have:

$$I = \left| \left(\frac{2\alpha_t \alpha_\phi}{\pi(1-t_n)} \right) \sum_p \sum_q \exp\{ikfW(t_p, \phi_q)\} \cdot \right.$$

$$\left. \frac{\sin\{\alpha_t kfW_t(t_p, \phi_q)\}}{\{\alpha_t kfW_t(t_p, \phi_q)\}} \cdot \frac{\sin\{\alpha_\phi kfW_\phi(t_p, \phi_q)\}}{\{\alpha_\phi kfW_\phi(t_p, \phi_q)\}} \right|^2$$

$$= \left| \frac{2\alpha_t \alpha_\phi}{\pi(1-t_n)} \cdot S \right|^2 \qquad (3.11)$$

where S represents the double summation over p and q.

We now consider the summation over q. The scale of division in the azimuth ϕ depends on the type of aberrations to be considered. Referring to Fig. 3.1, the circular aperture is represented by the four quadrants A,B,C and D. The angle ϕ being measured in a clockwise direction with $\phi=0$ along the positive y axis. The values of ϕ_q and α_ϕ will now be described for the following types of aberration.

Type 1 : $W_{even} \neq o$, $W_{odd} \neq o$

When both even and odd parts of the aberration function are present in the aberration polynomial (as given in 3.3), the

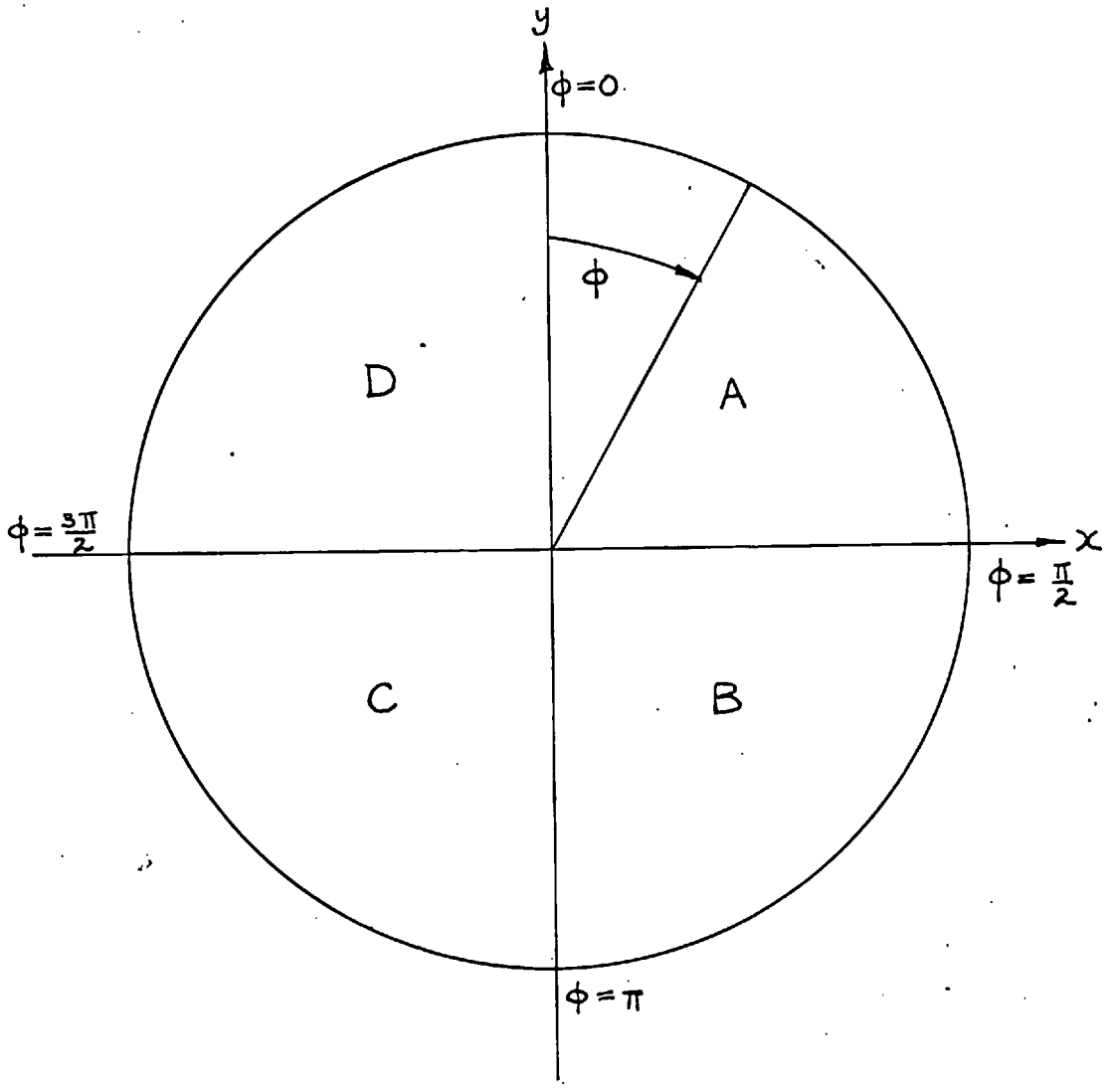


Fig: 3-1

integration with respect to ϕ in (3.5) need only take place in the quadrants A and B, and the result is then doubled to obtain the total contribution - this follows from the symmetry of the aberrations about the y-axis.

In this case, the semi-circular aperture (from $\phi=0$ to $\phi=\pi$) is divided into M equiangular parts, so that $q = 1,2,3,\ldots,M$, with $q = o$ corresponding to $\phi=0$. The value of ϕ_q in (3.6) is thus given by:

$$\phi_q = \left(\frac{q - \frac{1}{2}}{M}\right) \cdot \pi = \left(\frac{2q - 1}{2M}\right) \cdot \pi$$

$$0 < q \leqslant M \qquad (3.12)$$

The quantity α_ϕ defined in (3.6) is equal to half the angular division; that is :

$$\alpha_\phi = \tfrac{1}{2} \left(\frac{\pi}{M}\right) \qquad (3.13)$$

Type 2 : $W_{even} \neq 0$, $W_{odd} = 0$

Here only the even part of the aberration is present in the aberration polynomial, which now has symmetry about the y-axis as well as the x-axis, so that the integration for ϕ in (3.5) need only take place in the quadrant A (see Fig. 3.1). The total contribution from the four quadrants A,B,C and D is four times that from A. The quadrant A with $\phi = 0 \rightarrow \frac{\pi}{2}$ is to be divided into M equal parts, and corresponding to (3.12) and (3.13), we have

$$\phi_q = \left(\frac{2q - 1}{2M}\right) \cdot \frac{\pi}{2} \qquad (3.14)$$

$$0 < q \leqslant M$$

and $$\alpha_\phi = \tfrac{1}{2} \cdot \left(\frac{\pi}{2M}\right) \qquad (3.15)$$

Note for the special case when Type 2 has the even part of the aberration independent of ϕ, i.e. when $W_{2,2}$ and $W_{4,4}$ are all zero, then M=1 may be used.

Type 3: $W_{even} = 0$, $W_{odd} \neq 0$

Here only the odd part of the aberration is present in the aberration polynomial, which now has symmetry about the y-axis and assymmetry about the x-axis. The former implies that as in the Type 1, the integration for ϕ need only take place from $\phi = 0$ to $\phi = \pi$, and the result is to be doubled; the assymmetry about the x-axis implies that the imaginary part of (3.5) has zero resultant contribution to the integration over the range from $\phi = 0$ to $\phi = \pi$ and, over this range of ϕ, the contribution from the real part is equal to twice that obtained by integrating the real part over the range of ϕ from $\phi = 0$ to $\phi = \pi/2$. The resultant effect of these is that for the Type 3 case only the real part of (3.5) need be considered; the integration for ϕ takes place only over the range $\phi = 0$ to $\phi = \pi/2$, and the total contribution for the circular aperture is four times this result. The

values of ϕ_q and α_ϕ are as given in (3.14) and (3.15).

SUMMARY. The values of ϕ_q and α_ϕ defined above for the three types of aberrations require the following definitions for S (see 3.11).

For Type 1 $S = 2 \sum_p \sum_q \ldots$

i.e. S = twice the contribution given by the summations, resulting from the range of ϕ being taken only from $\phi = 0$ to $\phi = \pi$.

For Type 2 $S = 4 \sum_p \sum_q \ldots$

i.e. S = four times the contribution given by the summations, resulting from the range of ϕ being taken only from $\phi = 0$ to $\phi = \pi/2$.

For Type 3 $S = 4.\ \text{Real} \left| \sum_p \sum_q \ldots \right|$

i.e. S = four times the real part of the contribution given by the summations, resulting from the range of ϕ being taken only from $\phi = 0$ to $\phi = \pi/2$.

The range of values for p is $n+1 \leqslant p \leqslant N$ as defined in (3.8) and (3.10). The range of value for q is $0 < q \leqslant M$. The accuracy of the above numerical integration method depends on the values of M and N: the larger these values the higher the accuracy. In general the following values have been found to be suitable:

$$N = 10 \quad \text{(assuming n=o)}$$

$$M = 10 \quad \text{(for Types 2 and 3)}$$

$$= 20 \quad \text{(for Type 1)}$$

(III-iii) The flow diagram for evaluating I.

A computer programme has been written with the following basic applications in mind :

(1) For given values of the aberration coefficients $W_{n,q}$ the programme will calculate the value of I. The method for obtaining the coefficients $W_{n,q}$ in a practical case will be described at the end of Chapter 5 (see also equations 5.17, 5.19 and 5.21).

(2) For given values of $W_{2,0}$, $W_{4,0}$, $W_{6,0}$,....., different values of the scaling factor f are employed; for each f value the corresponding I is calculated. These procedures will be required to be repeated for a range of values of $W_{2,0}$.

(3) Provisions are to be made for similar investigations to be carried out with the odd aberrations, and a range of values of $W_{1,1}$ are then required to be read-in.

NOTE for (1) above, the value of f will be set to unity.

All aberrations are expressed in the units of wavelength, with $k = 2\pi$ (see 3-2). The general expression for $W(t,\phi)$ may be obtained from (3.3) by putting $t = r^2$ as follows :

$$W(t,\phi) = W_{2,0}t + W_{4,0}t^2 + W_{6,0}t^3 + W_{8,0}t^4 + W_{10,0}t^5$$

$$+ (W_{1,1}t^{\frac{1}{2}} + W_{3,1}t^{3/2} + W_{5,1}t^{5/2} + W_{7,1}t^{7/2} + W_{9,1}t^{9/2})\cos\phi$$

$$+ (W_{2,2}t+W_{4,2}t^2)\cos^2\phi+W_{3,3}t^{3/2}\cos^3\phi$$

$$+ W_{4,4}t^2\cos^4\phi \qquad (3.16)$$

Differentiating $W(t,\phi)$ with respect to t, we have

$$W_t(t,\phi)=W_{2,0}+2W_{4,0}t+3W_{6,0}t^2+4W_{8,0}t^3+5W_{10,0}t^4$$

$$+(0.5W_{1,1}t^{-\frac{1}{2}}+1.5W_{3,1}t^{\frac{1}{2}}+2.5W_{5,1}t^{3/2}+3.5W_{7,1}t^{5/2}$$

$$+ 4.5W_{9,1}t^{7/2})\cos\phi$$

$$+(W_{2,2}+2W_{4,2}t)\cos^2\phi+ 1.5\ W_{3,3}t^{\frac{1}{2}}\cos^3\phi$$

$$+ 2W_{4,4}t\cos^4\phi \qquad (3.17)$$

Differentiating $W(t,\phi)$ w.r.t. the azimuth angle ϕ, we have

$$W_\phi(t,\phi)=(W_{1,1}t^{\frac{1}{2}}+W_{3,1}t^{3/2}+W_{5,1}t^{5/2}+W_{7,1}t^{7/2}+W_{9,1}t^{9/2}).(-\sin\phi)$$

$$+(W_{2,2}t+W_{4,2}t^2).(2\cos\phi).(-\sin\phi)$$

$$+(W_{3,3}t^{3/2}).(3\cos^2\phi).(-\sin\phi)$$

$$+(W_{4,4}t^2).(4\cos^3\phi).(-\sin\phi) \qquad (3.18)$$

The above expressions (3.16),(3.17) and (3.18) will be used in the flow diagram. At the input stage, any of the above aberration coefficients may be set to zero when they are not required. The flow diagram also requires the use of the following codes:

(a) NO = 0 if more than one value of $W_{2,0}$ are to be read-in

NO = 1 " " " " " " $W_{1,1}$ " " " " "

(b) K11= the number of $W_{1,1}$'s to be read-in.

(c) K20= " " " $W_{2,0}$'s " " " ".

(d) n, N,M,f, TYPES 1,2 and 3 all have the same significances as defined in Section (III-ii).

(e) Δf = increment of f, so that $f = f+\Delta f$ for the next scaling factor.

(f) the maximum value of f required is $f_{max.}$; $f_{max}=f_o+k(\Delta f)$

where f_o = initial value of f.

k = integer constant.

SEE P.46 FOR FLOW DIAGRAM

(III-iv) Checking of the numerical integration results.

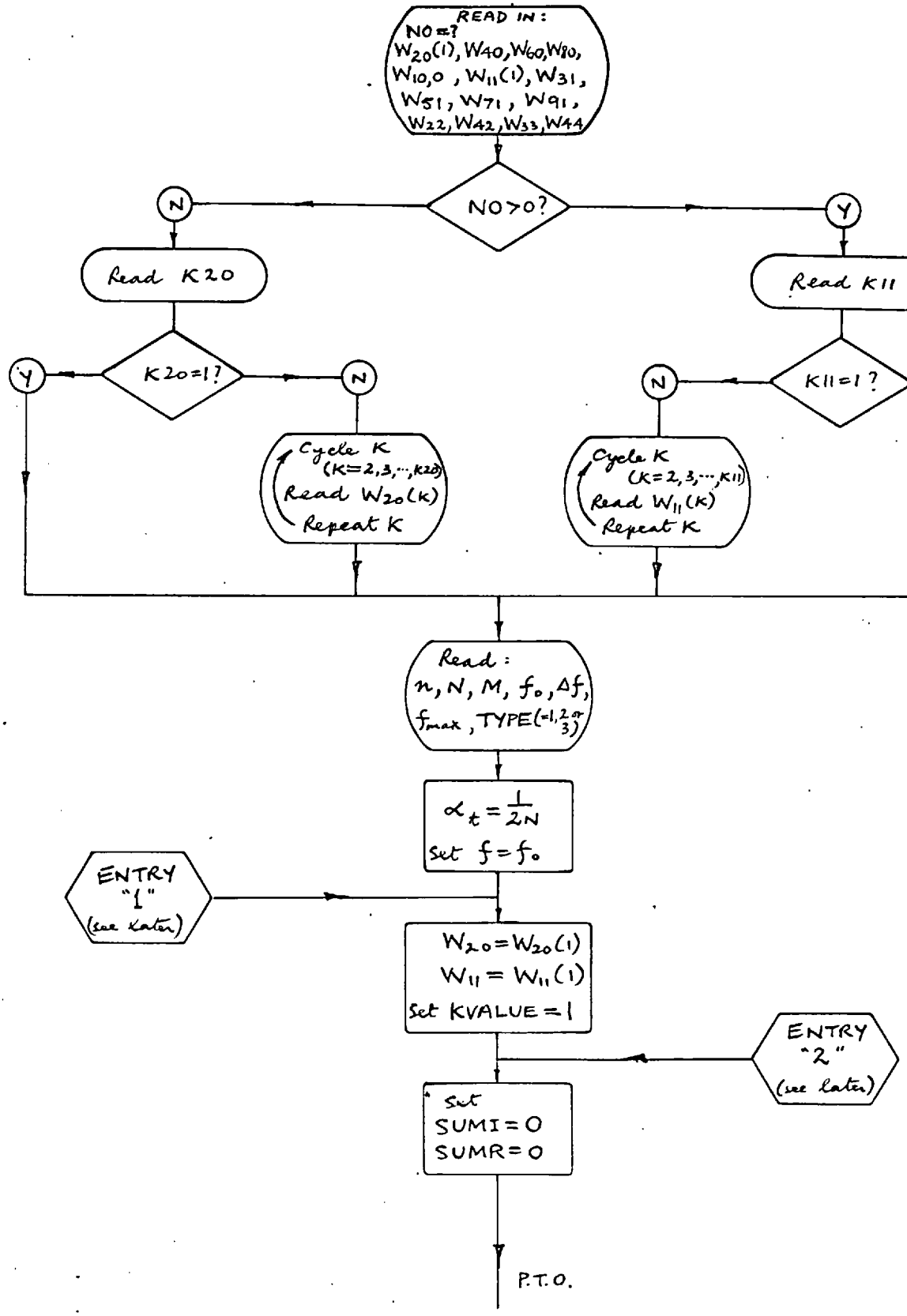
Two independent computer programmes for calculating I have been written for the purpose of cross-checking results. The flow diagram given above represents the version prepared for the IBM 7094 computer, using the Fortian code. The other programme was written for the Elliot 803A computer in Elliott autocode, and was based on a modified flow diagram (e.g. the cycles for p and q described above were intentionally reversed in order). The results from these two independently written programmes agreed completely.

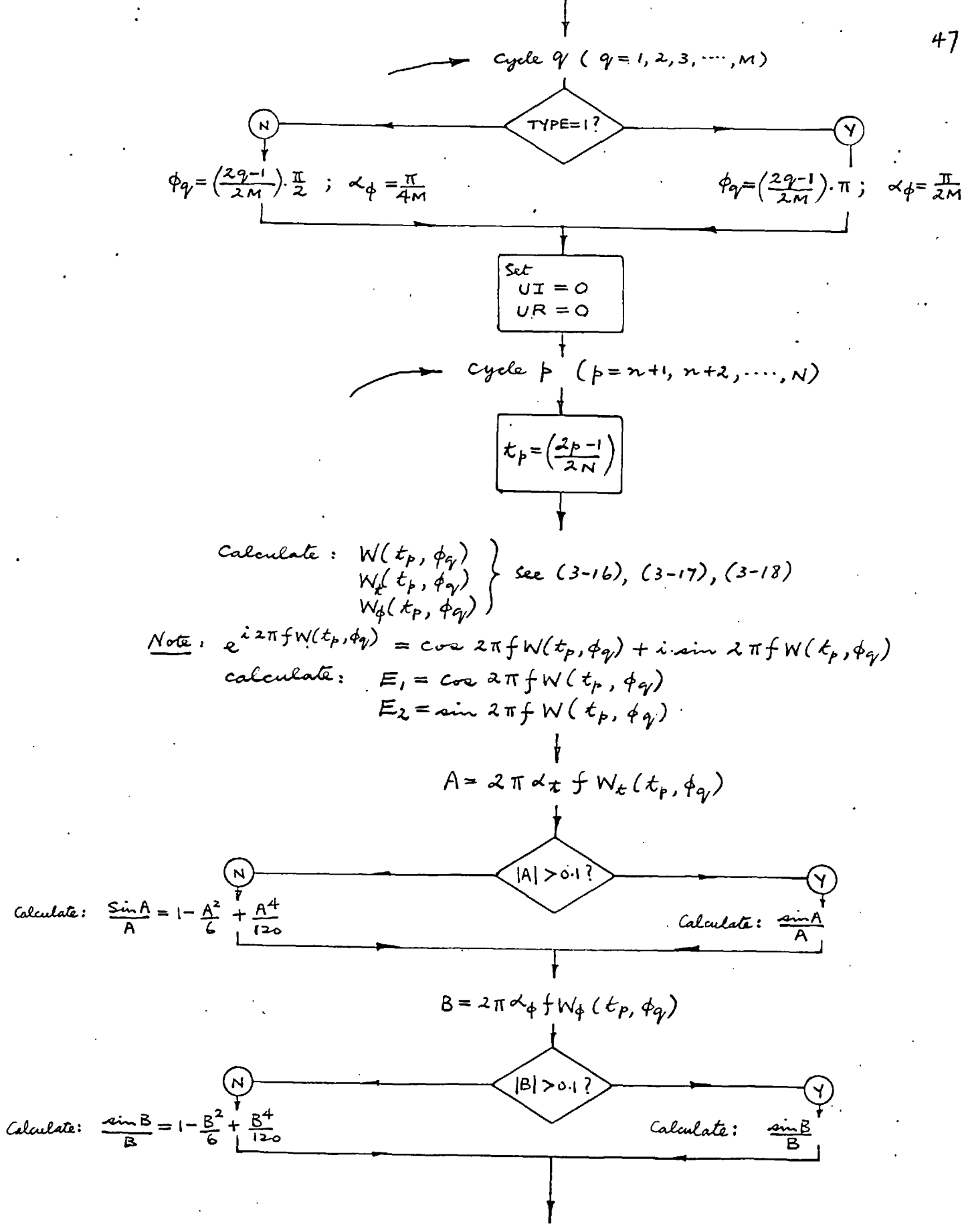
A further check on the validity of the numerical integration technique was made by using the complex Fresnel Integral :

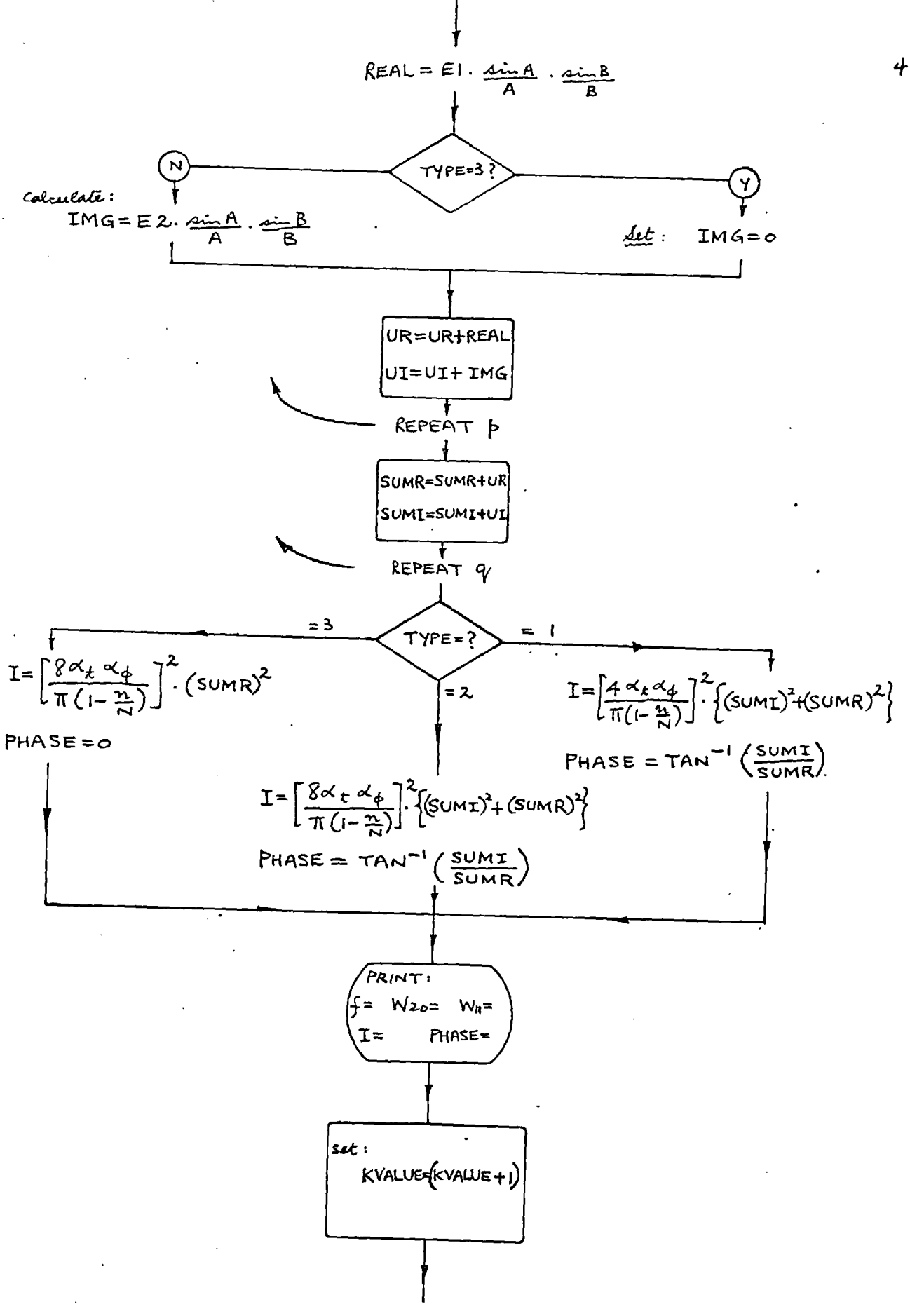
$$\int_o^u \exp\left[i\,\frac{\pi t^2}{2}\right]\,dt \qquad (3.19)$$

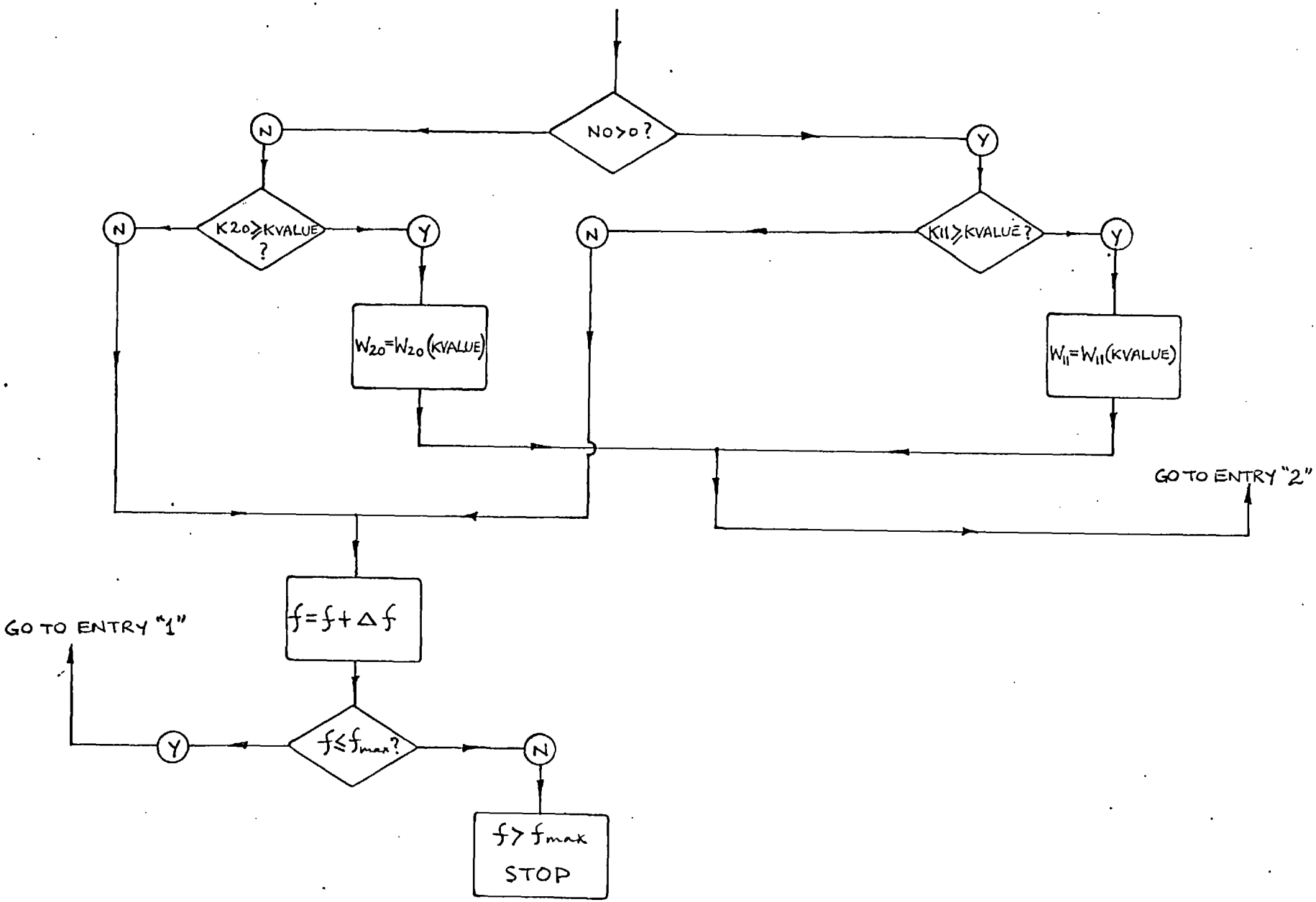
CONTINUED ON P.50

Flow Diagram for evaluating I




Cycle q ($q = 1, 2, 3, \cdots, M$)
TYPE=1?
N
Y
$\phi_q = \left(\frac{2q-1}{2M}\right) \cdot \frac{\pi}{2}$; $\alpha_\phi = \frac{\pi}{4M}$
$\phi_q = \left(\frac{2q-1}{2M}\right) \cdot \pi$; $\alpha_\phi = \frac{\pi}{2M}$
Set
UI = 0
UR = 0
Cycle p ($p = n+1, n+2, \cdots, N$)
$t_p = \left(\frac{2p-1}{2N}\right)$
Calculate: $W(t_p, \phi_q)$
$W_t(t_p, \phi_q)$
$W_\phi(t_p, \phi_q)$
See (3-16), (3-17), (3-18)
Note: $e^{i 2\pi f W(t_p, \phi_q)} = \cos 2\pi f W(t_p, \phi_q) + i \cdot \sin 2\pi f W(t_p, \phi_q)$
calculate: $E_1 = \cos 2\pi f W(t_p, \phi_q)$
$E_2 = \sin 2\pi f W(t_p, \phi_q)$
$A = 2\pi \alpha_t f W_t(t_p, \phi_q)$
|A| > 0.1?
N
Y
Calculate: $\frac{\sin A}{A} = 1 - \frac{A^2}{6} + \frac{A^4}{120}$
Calculate: $\frac{\sin A}{A}$
$B = 2\pi \alpha_\phi f W_\phi(t_p, \phi_q)$
|B| > 0.1?
N
Y
Calculate: $\frac{\sin B}{B} = 1 - \frac{B^2}{6} + \frac{B^4}{120}$
Calculate: $\frac{\sin B}{B}$


REAL = E1 · sin A/A · sin B/B
TYPE=3 ?
N
Y
Calculate:
IMG = E2 · sin A/A · sin B/B
Set: IMG=0
UR=UR+REAL
UI=UI+IMG
REPEAT p
SUMR=SUMR+UR
SUMI=SUMI+UI
REPEAT q
TYPE=?
=3
=2
=1
I = [8 α_t α_φ / π(1 − n/N)]² · (SUMR)²
PHASE=0
I = [8 α_t α_φ / π(1 − n/N)]² · {(SUMI)²+(SUMR)²}
PHASE = TAN⁻¹ (SUMI/SUMR)
I = [4 α_t α_φ / π(1 − n/N)]² · {(SUMI)²+(SUMR)²}
PHASE = TAN⁻¹ (SUMI/SUMR)
PRINT:
f= W20= W11=
I= PHASE=
set:
KVALUE=(KVALUE+1)


NO>0?
N
Y
K20≥KVALUE ?
N
Y
W20=W20(KVALUE)
K11≥KVALUE?
N
Y
W11=W11(KVALUE)
GO TO ENTRY "2"
f=f+Δf
GO TO ENTRY "1"
f≤fmax?
Y
N
f>fmax
STOP

CONTINUED FROM P.45

Referring to equation (3.2), if the aberration scaling factor f is unity, and if the pupil is a circle of unit radius ($A=\pi$), then

$$I = \left| \frac{1}{\pi} \int_{r=o}^{r=1} \int_{\phi=o}^{\phi=2\pi} \exp[i2\pi W(r,\phi)]\, r\, dr\, d\phi \right|^2 \qquad (3.20)$$

If $W(r,\phi)$ is given by

$$W(r,\phi) = W_{2,0}\, r^2 + W_{4,0} r^4 \qquad (3.21)$$

and if U represents the quantities inside the modulus sign, we have

$$U = \frac{1}{\pi} \int_{r=o}^{r=1} \int_{\phi=o}^{\phi=2\pi} \exp[i2\pi(W_{2,0} r^2 + W_{4,0} r^4)]\, r\, dr\, d\phi$$

i.e.

$$U = 2 \int_o^1 \exp\left[i2\pi W_{4,0}\left(r^4 + \frac{W_{20} r^2}{W_{40}}\right)\right] r\, dr$$

$$= 2 \exp\left(-\frac{i\pi W_{2,0}^2}{2W_{4,0}}\right) \int_o^1 \exp[i2\pi W_{4,0}(r^2 + \frac{W_{20}}{2W_{40}})^2]\, r\, dr \qquad (3.22)$$

Let $$\frac{\pi}{2} \cdot t^2 = 2\pi W_{40} \left(r^2 + \frac{W_{20}}{2W_{40}}\right)^2 \qquad (3.23)$$

so that $$\frac{dt}{4\sqrt{W_{40}}} = r\, dr$$

Substituting (3.23) in (3.22), we have

$$U = \frac{\exp\left(\frac{-i\pi W_{20}^2}{2W_{40}}\right)}{2\sqrt{W_{40}}} \int_{\frac{W_{20}}{\sqrt{W_{40}}}}^{(W_{20}+2W_{40})/\sqrt{W_{40}}} \exp\left(\frac{i\pi t^2}{2}\right) dt$$

This may be written as :

$$U = \frac{\exp\left(\frac{-i\pi W_{20}^2}{2W_{40}}\right)}{2\sqrt{W_{40}}} \left\{ \int_0^{(W_{20}+2W_{40})/\sqrt{W_{40}}} \exp\left(\frac{i\pi t^2}{2}\right) dt - \int_0^{W_{20}/\sqrt{W_{40}}} \exp\left(\frac{i\pi t^2}{2}\right) dt \right\}$$

in terms of the Fresnel Integrals. Expressing the results in a standard form[2] we have

$$U = \frac{\exp\left(\frac{-i\pi W_{20}^2}{2W_{40}}\right)}{2\sqrt{W_{40}}} \left\{ C\left[\frac{W_{20}+2W_{40}}{\sqrt{W_{40}}}\right] - C\left[\frac{W_{20}}{\sqrt{W_{40}}}\right] + iS\left[\frac{W_{20}+2W_{40}}{\sqrt{W_{40}}}\right] - iS\left[\frac{W_{20}}{\sqrt{W_{40}}}\right] \right\}$$

Since $I = |U|^2$, we have

$$I = \frac{1}{4W_{40}} \left\{ \left\{ C\left[\frac{W_{20}+2W_{40}}{\sqrt{W_{40}}}\right] - C\left[\frac{W_{20}}{\sqrt{W_{40}}}\right] \right\}^2 + \left\{ S\left[\frac{W_{20}+2W_{40}}{\sqrt{W_{40}}}\right] - S\left[\frac{W_{20}}{\sqrt{W_{40}}}\right] \right\}^2 \right\} \quad (3.24)$$

For given values of W_{20} and W_{40}, C[ ]'s and S[ ]'s in (3.24) may be found from tables of Fresnel Integrals.[2] The values of I so calculated provide the independent means of checking the validity of the numerical integration method used in the computer programmes. For this purpose, several pairs of W_{20}, W_{40} values were used. For the computer in-put, all the aberration coefficients, other than W_{20}, W_{40}, were set to zero. The value of N used was 10, M was set to unity since the aberrations were independent of ϕ (see equation 3.15); in addition, n=o for a clear circular aperture and f=1. The computed values of I were then compared with the values of I calculated from (3.24). Some of the results are shown in Table 3.1: the agreement between the two sets of I values for the corresponding pairs of W_{20}, W_{40} is excellent, thereby proving the validity of the numerical integration method used. Further evidence of this will be described in (III-vi).

W_{40} (λ's)	W_{20}(λ's)	I(computer output)	I(calculated from (3.24))
4	-4	0.08921	0.08919
4	-2	0.17470	0.17469
4	1	0.00860	0.00860

TABLE 3.1

(III-v) The comparison of the Strehl Intensity Ratios and the validity of Maréchal's approximate formula.

Let the aberration polynomial representing the axial image case be

$$W = W_{20}r^2+W_{40}r^4+W_{60}r^6 \qquad (3.25)$$

Substituting this polynomial in the variance formula (1.5) we have

$$E = \frac{1}{\pi}\int_{r=o}^{1}\int_{\phi=o}^{2\pi}(W_{20}r^2+W_{40}r^4+W_{60}r^6)^2\, r\, dr\, d\phi$$

$$-\left[\frac{1}{\pi}\int_{r=o}^{1}\int_{\phi=o}^{2\pi}(W_{20}r^2+W_{40}r^4+W_{60}r^6)r\, drd\phi\right]^2$$

Evaluating this we have the following result for E:

$$E = \frac{W_{20}^2}{12} + \frac{4W_{40}^2}{45} + \frac{9W_{60}^2}{112} + \frac{W_{20}W_{40}}{6} + \frac{W_{40}W_{60}}{6} + \frac{3W_{20}W_{60}}{20} \qquad (3.26)$$

The conditions for E to be a minimum are :

(i) $\frac{\partial E}{\partial W_{20}} = 0$ → best focal plane.

(ii) $\frac{\partial E}{\partial W_{40}} = 0$ → best form of correction.

The values of W_{20} and W_{40} satisfying the above conditions are :

$$W_{40} = -1.5\, W_{60}$$
$$W_{20} = 0.6\, W_{60} \qquad (3.27)$$

Substituting (3.27) in (3.26), and using Maréchal's tolerance

$$(E)_{min} \approx \frac{\lambda^2}{180} \qquad (3.27)A$$

for I = 0.8, we have the following permissible values for the three aberration coefficients:

$$W_{20} = 2.4\lambda$$
$$W_{40} = -6.0\lambda \qquad (3.28)$$
$$W_{60} = 4.0\lambda$$

These results were first given by Maréchal[3] (1948). If we now scale down the aberration so that W_{60} becomes 1.0λ, then from (3.27) we have

$$W_{40} = -1.5 \times 1.0 = -1.5\lambda$$
$$W_{20} = 0.6\lambda$$

Substituting these in (3.25), we have

$$W = 0.6\, r^2 - 1.5\, r^4 + 1.0\, r^6 \qquad (3.29)$$

and this result will be used for later comparison.

The use of the aberration scaling factor f in the Strehl Intensity Ratio investigations.

PART I. The presence of the aberration scaling factor f in (3.2) enables the argument of the exponential function to be varied. This facilitates the study of the influence on the intensity ratio I of aberrations of different magnitude, according to the formula :

$$fW = f(0.6\, r^2 - 1.5r^4 + 1.0r^6)$$

Thus f = 4 gives Maréchal's aberration tolerance values as defined in (3.28).

For the present purpose, let the quantity fW in the Strehl Intensity Ratio integral (3.2) be given by

$$fW = f(W_{20}r^2 + W_{40}r^4 + 1.0r^6) \qquad (3.30)$$

with $W_{60} = 1.0\lambda$.

The Strehl Intensity Ratio computer programme is now used to compute the values of I corresponding to the following cases* [SEE FOOTNOTE]:

Case(A) $f = f_o$, $W_{40} = -1.5\lambda$

$W_{20}=W_{20}(k)$, with k=1, 2,3,...K.

(the K values of W_{20} are to be tried one at a time).

Case (B) Repeat (A) with a range of f values given by:

$f = f_o + p.\Delta f$, $p = 1,2,3,....,P$

(the P values of f are to be tried one at a time).

Case (C) Repeat (A) and (B), but now with $W_{40}= -1.25\lambda$.

The results are to be presented as curves of I against W_{20}. For each W_{40} value, there will be (P+1) of these curves, corresponding to (P+1) values of f. From each curve (each f value), the values of I_{max} and the corresponding $(W_{20})_{Imax}$ may be found. Two curves may then usefully be plotted, namely

(1) $\left[\frac{(W_{20})_{Imax}}{W_{60}}\right]$ versus f

(2) I_{max} versus f

corresponding to the particular W_{40} value used.

* See also the important paper "On the best fine-correction of photographic lenses with 2-term spherical aberration" by E.H.Linfoot (Optica Acta, vol. 10, P77-104, 1963) in which similar results were described for larger values of the secondary spherical aberration coefficient. The present computations extend the part of the investigation which deals with "aerial images" to smaller values of W_{60} (see P.55) .

Discussion of Results for PART 1

(i) $W_{40} = -1.5\lambda$ case

The starting value of $f(=f_o)$ used for the investigation was f = 1.0. Since $W_{60} = 1.0\lambda$, $W_{40} = -1.5\lambda$, direct comparison between expressions (3.29) and (3.30) showed that $W_{20} = 0.6$ would give $I = I_{max}$ for this initial case. The graph I/W_{20} would therefore require a range of W_{20} values with $W_{20} = 0.6$ as roughly the mean. Consequently, the range selected was from -0.2λ to 1.2λ; with $W_{20}(k)$ given by

$$W_{20}(k) = 0.01\,(k-1) - 0.2;\ k = 1,2,3,\ldots,141(=K).$$

The increment Δf was set to 1.0, and the total number of f values investigated were 10 (i.e.P=9). Referring now to the input of the flow diagram for I; all the aberration coefficients, other than W_{20}, W_{40} and W_{60}, were set to zero; the values for n, N and M used were the same as described in (III-iv); f_{max} was 10, and since only even aberrations were considered, the code for TYPE was '2'.

The computer output I for the cases (A) and (B) were plotted. Fig.3.2 shows a total of ten of these curves, corresponding to f = 1,2,3,...,10. For each curve, the regions close to I_{max} were required to be determined accurately: additional spot values of W_{20} were therefore used for this purpose. Referring now to Fig. 3.2:

When f=1, the effective values of fW_{40} and fW_{60}

are only $^1/4$ of that given in (3.28). Consequently the I/W_{20} curve has a broad peak, with I_{max} located at $W_{20}= 0.6$ as expected (see 3.29).

When f = 4, the effective values of fW_{40}, fW_{60} are that given in (3.28). The value of $f.W_{20}$ corresponding to I_{max} is actually found to be $4\times0.605 = 2.42\lambda$, differing slightly from the W_{20} value given in (3.28). This had been anticipated since the relation (3.27)A used for deriving the results in (3.28) is not an exact one. The actual value of I_{max} was found to be 0.80, agreeing with Maréchal's prediction.

When f=10, there are two maxima in the I/W_{20} curve: the broad peak can easily be deduced to be the 'principal maximum', whilst the other maximum corresponds to the secondary. At this stage, due to the large aberrations present, the Strehl ratio is completely insufficient to describe the complex effects of aberration on the diffraction image of a point source (see Chapter 1, section I-ii).

The graph in Fig.3.3 is derived from Fig.3.2, and represents the plot of the W_{20}/W_{60} values corresponding to I_{max} for the ten values of f. The ratio $[W_{20}]_{Imax}/W_{60}$ is clearly shown to vary when the aberration is being scaled-up by f. The reference line $\left(\frac{W_{20}}{W_{60}}\right) = 0.6$ illustrates the case if I_{max} for different f values are associated with the same $\left(\frac{W_{20}}{W_{60}}\right)$ ratio as when f is small ($f \leqslant 1$).

In Fig. 3.4, the values of I_{max} corresponding to f=1,2,3, ...,10 (as obtained from Fig.3.2) are plotted against f. The purpose of this I_{max}/f curve will be described later.

(ii) $W_{40} = -1.25\lambda$case

The computer input data for $W_{40} = -1.25\lambda$ case were the same as that described above for $W_{40} = -1.5\ \lambda$. The I/W_{20} curves for f = 1,2,...,10 are as shown in Fig.3.5. Here the ratio W_{40}/W_{60} is no longer that given in (3.27) for the best form of aberration correction, and as expected, the graphs show that

$$(I_{max})_{W_{40}= -1.25\lambda} < (I_{max})_{W_{40} = -1.5\lambda} \qquad (3.30)A$$

for the corresponding f values, provided that the latter are not too large (i.e. $f \leqslant 7$). From Fig. 3.5, it is seen that when $f > 7$, the inequality sign in (3.30)A is reversed, and the peaks are also more sharply defined as compared with that in Fig. 3.2. It is of interest to note that for the f = 10 curve in Fig. 3.5, the principal maximum, corresponding to $W_{20}= 0.42$, is approximately three times the peak value of the secondary, and about 1.3 times the value of I_{max} of the corresponding curve in Fig.3.2.

Fig. 3.6 represents the plot of the $\left(\frac{W_{20}}{W_{60}}\right)$ values corresponding to I_{max} for the ten values of f as derived from Fig.3.5. Direct comparison between this and Fig. 3.3. shows that as f increases from zero to f=7, the case

Fig. 3-2

$W_{60} = 1.0\ \lambda$

$W_{40} = -1.5\ \lambda$

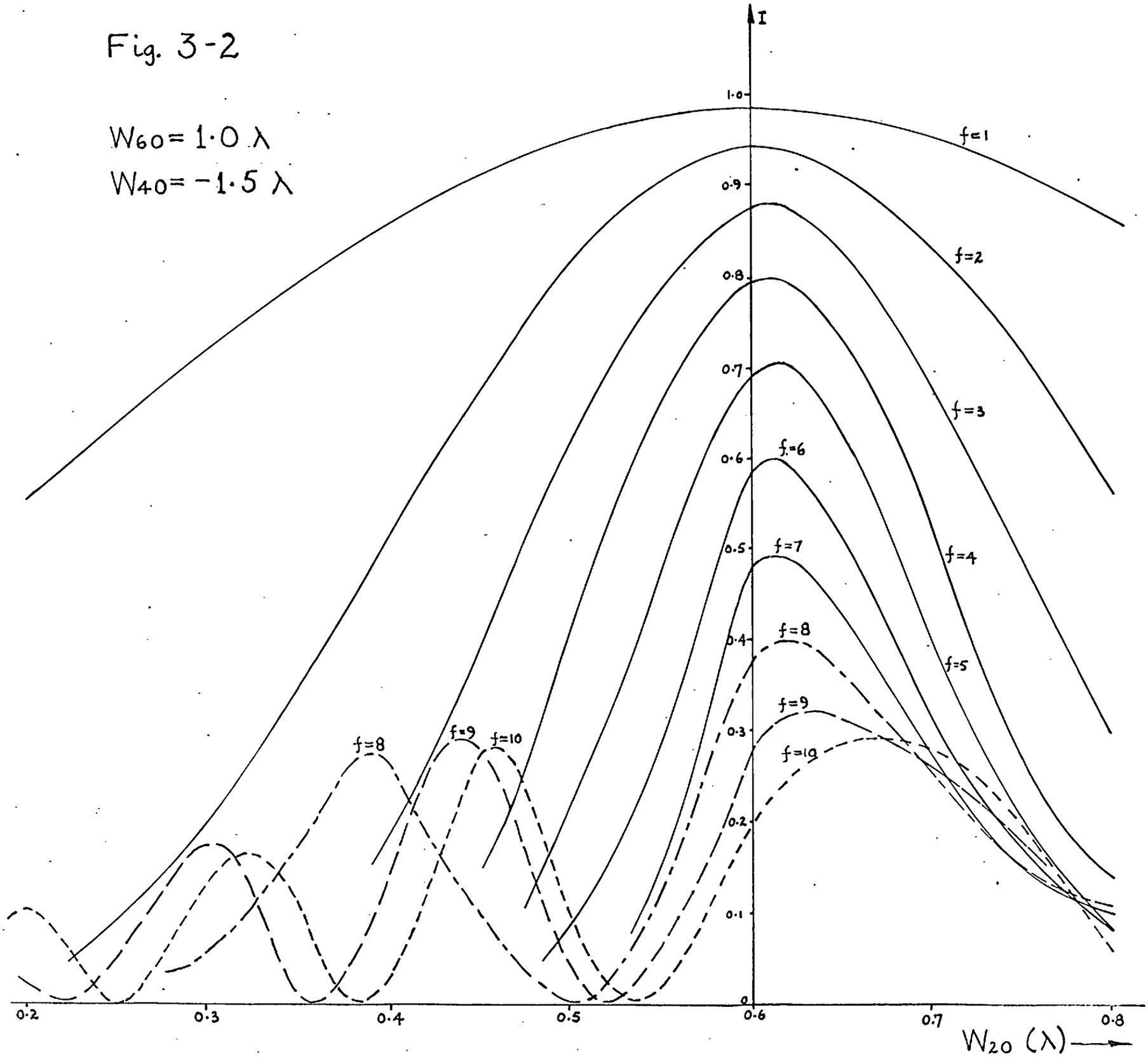
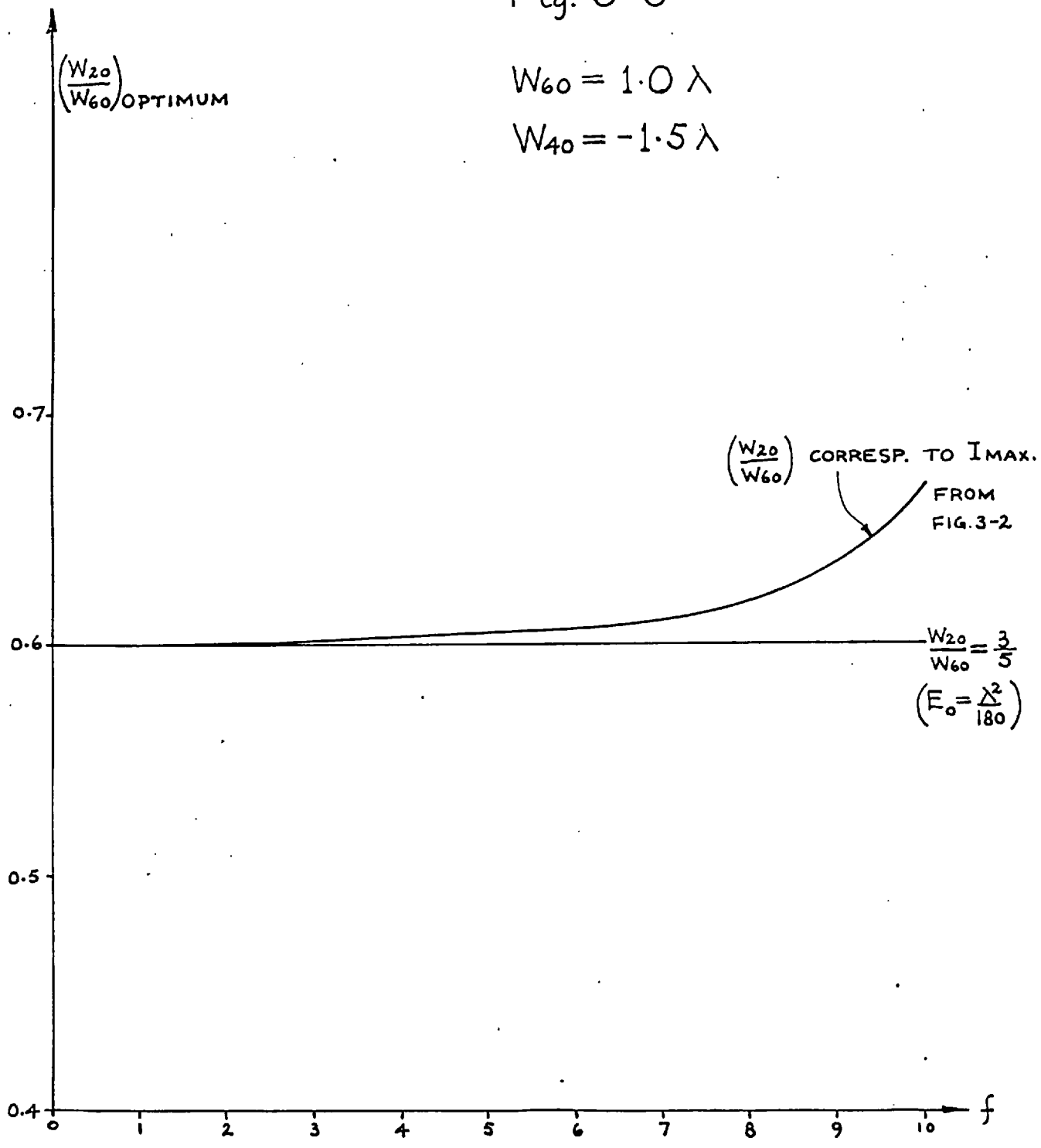


Fig. 3-3

$W_{60} = 1.0\,\lambda$

$W_{40} = -1.5\,\lambda$



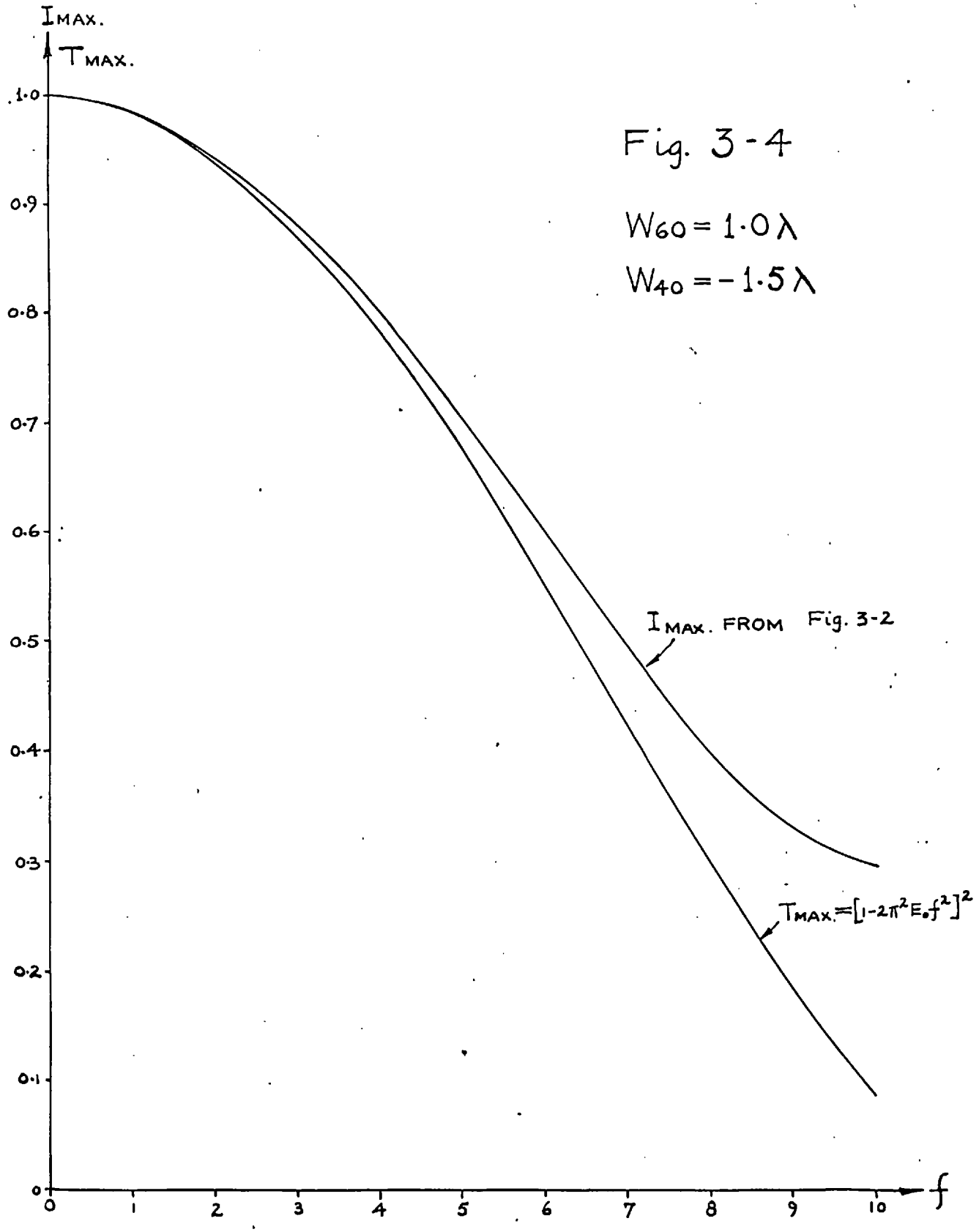


Fig. 3-4

$W_{60} = 1.0\lambda$

$W_{40} = -1.5\lambda$

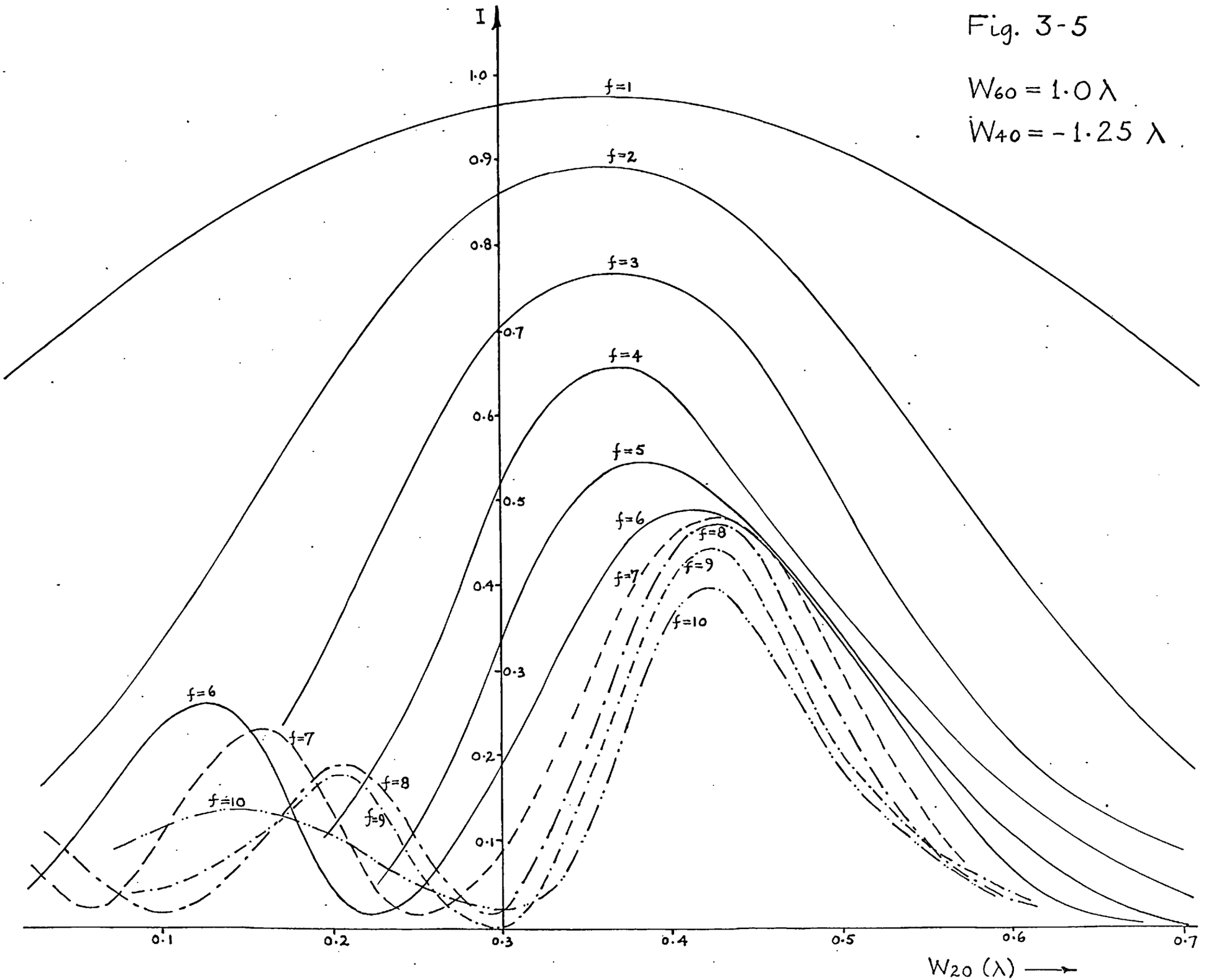
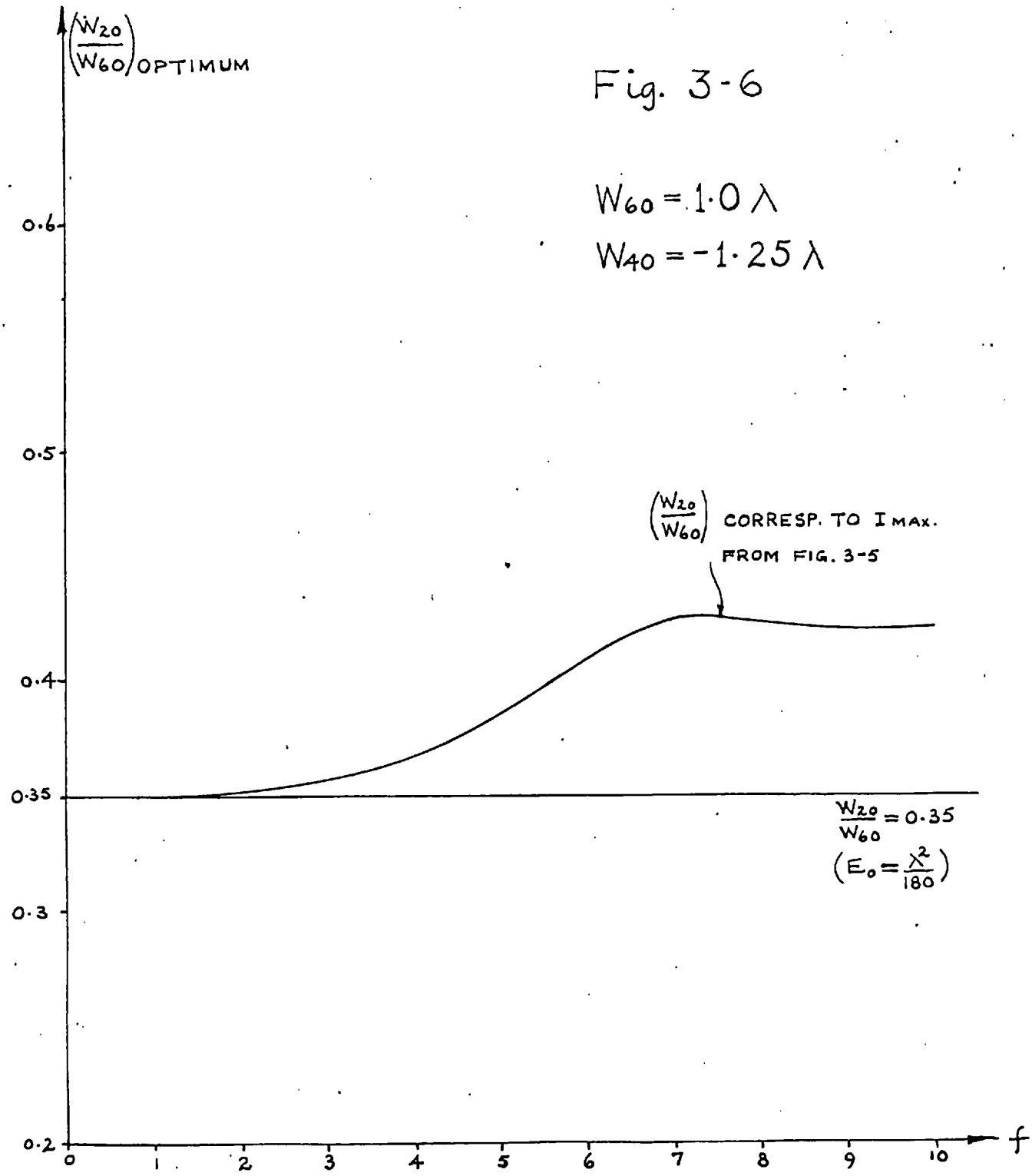


Fig. 3-5

$W_{60} = 1.0\,\lambda$

$W_{40} = -1.25\,\lambda$


$\left(\frac{W_{20}}{W_{60}}\right)_{OPTIMUM}$
Fig. 3-6
$W_{60} = 1.0\,\lambda$
$W_{40} = -1.25\,\lambda$
0.6
0.5
0.4
0.35
0.3
0.2
$\left(\frac{W_{20}}{W_{60}}\right)$ CORRESP. TO I MAX.
FROM FIG. 3-5
$\frac{W_{20}}{W_{60}} = 0.35$
$\left(E_0 = \frac{\lambda^2}{180}\right)$
0
1
2
3
4
5
6
7
8
9
10
f

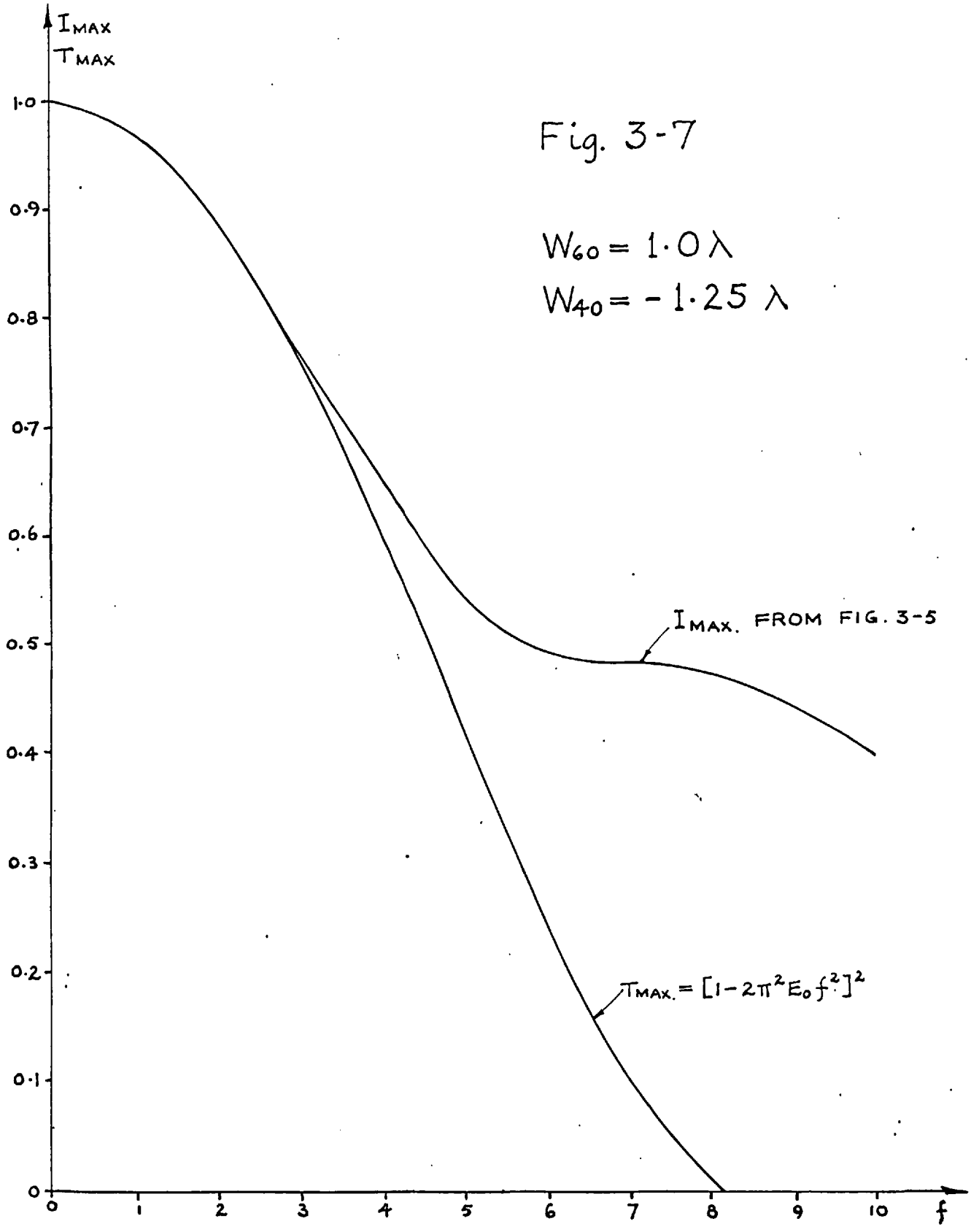


Fig. 3-7

$W_{60} = 1.0\lambda$

$W_{40} = -1.25\lambda$

corresponding to $W_{40} = -1.5W_{60}$ has a smaller variation in the ratio $\left(\frac{W_{20}}{W_{40}}\right)_{Imax}$ compared with that in the case $W_{40} = -1.25W_{60}$.

In Fig. 3.7, the values I_{max} corresponding to f=1,2,...,10 (as obtained from Fig.3.5) are plotted against f. The purpose of this Imax/f curve will be described in PART 2 which follows.

PART 2 It has been described earlier in Chapter 1 that Maréchal's approximate expression (1.4) may be used to calculate I, i.e.,

$$I \approx \left[1 - \frac{2\pi^2}{\lambda^2} E\right]^2 \qquad (3.31)$$

provided that the condition $I \geq 0.8$ (or $E \leq \frac{\lambda^2}{180}$) is satisfied. However, when this condition is violated such that E_o, the smallest value of E, is no longer $\leq \lambda^2/180$, then the quantity given by

$$\left[1 - \frac{2\pi^2}{\lambda^2} E_o\right]^2 \qquad (3.32)$$

will no longer represent the true value of I_{max}, a consequence which follows directly from the approximation made in the derivation of the expression (3.31) - see Chapter 1, (I-ii). Under this circumstance, it is of practical importance to know whether the result given by (3.32) is larger, or smaller, than the true I_{max}. This investigation may be carried out as follows:

As in (3.30), we scale each one of the aberration coefficients by the factor f; the effect of this is that the variance E is scaled by the factor f^2. The expression for f^2E may be obtained from (3.26) as follows :

$$f^2E = f^2\left[\frac{W_{20}^2}{12} + \frac{4W_{40}^2}{45} + \frac{9W_{60}^2}{112} + \frac{W_{20}W_{40}}{6} + \frac{W_{40}W_{60}}{6} + \frac{3W_{20}W_{60}}{20}\right] \qquad (3.33)$$

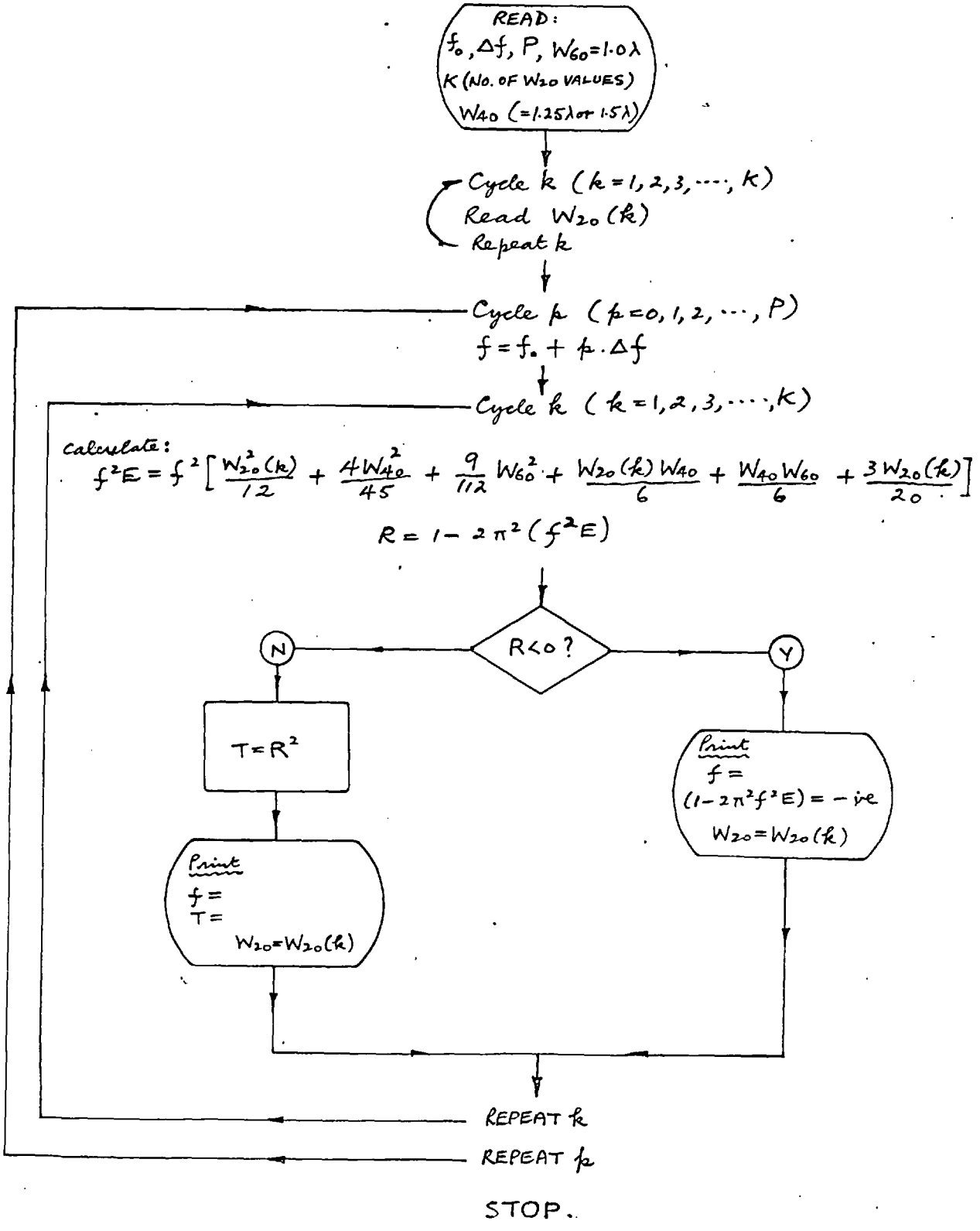
where, for the present purpose, $W_{60} = 1.0\lambda$.

It is now required to compute the values of T given by

$$T = [1 - 2\pi^2(f^2E)]^2 \qquad (3.34)$$

corresponding to the cases (A),(B) and (C) listed for PART 1 above. The factor λ^2 being omitted in (3.34) since the aberrations are measured in units of λ. A short programme has been prepared for this purpose and the flow diagram is as follows :

Flow Diagram for calculating T



The value of T is not calculated when

$$[1-2\pi^2 f^2 E] < 0 .$$

The results from the computer (Elliott 803A) are to be presented as T/W_{20} curves: for each W_{40} value there will be (P+1) of these curves, corresponding to (P+1) values of f. From each curve (each f value), the value of T_{max} may be found. A graph T_{max}/f may then be plotted, corresponding to the particular W_{40} value used.

Discussion of Results for PART 2

The input data used for calculating T were as follows :

For W_{40}= -1.5λ case.

The values of $W_{20}(k)$ used were calculated from

$$W_{20}(k) = 0.5 + (k-1) \times 0.01$$

with = 1,2,3,....,31(=K)

corresponding to the range from 0.5λ to 0.8λ - this was deduced from Fig. 3.2 as the probable range for W_{20} within which the values of T_{max} could be located for the ten values of f.

For W_{40} = - 1.25λ case.

Corresponding to the above case, the values of $W_{20}(k)$ tried were given by

$$W_{20}(k) = 0.25 + (k-1) \times 0.01$$

with k = 1,2,3,....,31 (=K)

and the range of W_{20} values were from 0.25λ to 0.55λ (as deduced from Fig. 3.5).

The other input data were :

$f_0=1$, $\Delta f=1$, $P=9$ corresponding to $f=1,2,\ldots,10$, as described for PART 1.

The computer results were presented as T/W_{20} curves: since, for each W_{40}, there were ten f values and therefore ten T/W_{20} curves. From each one of these curves, T_{max}, the maximum value of T, was determined. A graph of T_{max}/f was then plotted, corresponding to the particular value of W_{40} used. In each case, it was found convenient to present the T_{max}/f curve on the same graph and using the same scales, as that used for I_{max}/f (see Figs. 3.4 and 3.7). The graphical results indicate clearly that the values of T_{max} as calculated from

$$T_{max} = [1 - 2\pi^2 f^2 E_o]$$

are such that

$$T_{max} \leqslant I_{max}$$

for any given values of f. We may thus conclude that provided $I \geqslant 0.8$, the result given by (3.32) will represent the Strehl Intensity Ratio, I, to a good approximation; when $I < 0.8$, the expression (3.32) will generally give a value of I_{max} which is low compared with the value given by accurate numerical integration.

In those cases where a clearly distinguishable principal maximum is present in the curve of Strehl intensity against focal plane, it is found that a decrease in the variance E leads to an increase in the Strehl ratio. On the basis of the present results, therefore, it would seem that for cases where $I \geqslant 0.5$ the variance E of the wave-aberration is useful as an approximate index of the decrease in the Strehl ratio from its theoretical maximum value of unity. In other cases it is possible to find that an increase in the variance E is associated with an increase in the Strehl ratio. For example, fig. 3-4 shows that for f=8 the Strehl ratio has the value 0.395 compared with 0.291 from Maréchal's formula: on the other hand fig.3-7 shows that for f=8 the Strehl ratio is equal to 0.471 while the Maréchal's formula gives 0.011 . Thus an increased variance has led to an increase in the Strehl ratio.

(III.vi) The special property of Strehl Intensity Curves: the conditions for different $\frac{W_{20}}{W_{60}}$ and $\frac{W_{40}}{W_{60}}$ values to give the same I. [SEE FOOT NOTE] (4),(5)

Referring to the Strehl Intensity Ratio expression in (3.20), i.e.

$$I = \left| \frac{1}{\pi} \int_{r=o}^{r=1} \int_{\phi=o}^{\phi=2\pi} \exp\left[i2\pi W \right] \, r dr d\phi \right|^2 \qquad (3.35)$$

Let $W = f(W_{20}r^2 + W_{40}r^4 + 1.0r^6)$ (3.36)

where f is the aberration scaling factor, W_{20} and W_{40} are defined relative to W_{60} as being unity. Substituting (3.36) in (3.35) and integrating w.r.t.ϕ, we have

$$I = \left| 2 \int_{r=o}^{1} \exp\left[i2\pi f(W_{20}r^2 + W_{40}r^4 + r^6) \right] r dr \right|^2$$

Put $t = r^2$

$dt = 2rdr$

then

$$I = \left| \int_{t=o}^{t=1} \exp[i2\pi f(W_{20}t + W_{40}t^2 + t^3)] \, dt \right|^2$$

Put $a = t - \frac{1}{2}$, so that $t = a + \frac{1}{2}$. Then

$$I = \left| \int_{a=-\frac{1}{2}}^{a=\frac{1}{2}} \exp\left\{ i2\pi f \left[W_{20}(a+\tfrac{1}{2}) + W_{40}(a^2 + a + \tfrac{1}{4}) + (a^3 + \tfrac{3}{2}a^2 + \tfrac{3}{4}a + \tfrac{1}{8}) \right] \right\} da \right|^2$$

NOTE. A more elegant proof of this property has been given by R.Richter[4] (1925, Z. InstrumKde, 45, 1) and E.H.Linfoot[5].

Regrouping the terms and simplifying, we have

$$I = \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp\left\{ i2\pi f(W_{40}+\frac{3}{2})a^2 \right\} . \exp\left\{ i2\pi f \left[(W_{20}+W_{40}+\frac{3}{4})a + a^3 \right] \right\} da \right|^2$$

$$= \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\{ \cos 2\pi f(W_{40}+\frac{3}{2})a^2 + i \sin 2\pi f(W_{40}+\frac{3}{2})a^2 \right\} \right.$$

$$\left. \left\{ \cos 2\pi f\left[(W_{20}+W_{40}+\frac{3}{4})a+a^3 \right] + i\sin 2\pi f\left[(W_{20}+W_{40}+\frac{3}{4})a + a^3 \right] \right\} da \right|^2$$

Since the range of integration is between $\pm\frac{1}{2}$, the term $i\sin 2\pi f[(\ )a+a^3]$ will not contribute to the integration, so that

$$I = \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\{ \cos\left[2\pi f(W_{40}+\frac{3}{2})a^2 \right] . \cos 2\pi f\left[(W_{20}+W_{40}+\frac{3}{4})a+a^3 \right] \right.\right.$$

$$\left.\left. + i \sin\left[2\pi f(W_{40}+\frac{3}{2})a^2 \right] . \cos 2\pi f\left[(W_{20}+W_{40}+\frac{3}{4})a+a^3 \right] \right\} da \right|^2$$

Put now $W_{40} = -\frac{3}{2} + \delta W_{40}$, so that $I = I_+$

$$I_+ = \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\{ \cos\left[2\pi f\ \delta W_{40} . a^2 \right] . \cos 2\pi f\left[(W_{20}+\delta W_{40}-\frac{3}{4})a+a^3 \right] \right.\right.$$

$$\left.\left. + i\sin\left[2\pi f \delta W_{40} a^2 \right] . \cos 2\pi f\left[(W_{20}+\delta W_{40}-\frac{3}{4})a+a^3 \right] \right\} da \right|^2 \qquad (3.37)$$

If δW_{40} above is now replaced by $-\delta W_{40}$, then the corresponding I_- will have its modulus the same as that of I_+, provided that the value of ( ) in (3.37) is such that

$$(\ )_+ = (\ )_-$$

i.e. $(W_{20_+} + |\delta W_{40}| - \frac{3}{4}) = (W_{20_-} - |\delta W_{40}| - \frac{3}{4})$

or $(W_{20})_+ = (W_{20})_- - 2|\delta W_{40}|$

This special property applies only when

$$(W_{40})_\pm = -\frac{3}{2} \pm \delta W_{40}$$

The significance of these results may be summarized as follows :

If the aberration function may be represented by

$$W = \text{constant} \times (W_{20}r^2 + W_{40}r^4 + W_{60}r^6)$$

$$= f \cdot \left(\frac{W_{20}}{W_{60}} r^2 + \frac{W_{40}}{W_{60}} r^4 + 1.0r^6 \right) \qquad (3.38)$$

then for a given value of f, we have

Case 1

$$\left(\frac{W_{40}}{W_{60}} \right)_1 = -1.5 - \delta W_{40} \quad \Big\} \text{ corresponding } I = I_1$$

$$\left(\frac{W_{20}}{W_{60}} \right)_1 = \text{a certain value} \qquad (3.39)$$

Case 2

$$\left(\frac{W_{40}}{W_{60}}\right)_2 = -1.5 + \delta W_{40} \quad \Big\} \text{ corresponding } I=I_2$$

$$\left(\frac{W_{20}}{W_{60}}\right)_2 = \left(\frac{W_{20}}{W_{60}}\right)_1 - 2|\delta W_{40}| \qquad (3.40)$$

The above proof shows that

$$|I_1| = |I_2| \text{ for } \underline{\text{all}} \text{ values of f and } \delta W_{40}.$$

This surprising result was discovered in the process of investigating the I/W_{20} curves for different f and W_{40} values. Referring now to the previous section (III-v), PART I, cases (A), (B) and (C); if, in addition to $W_{40} = -1.5\lambda$ and -1.25λ, a new value $W_{40} = -1.75\lambda$ is also to be tried, using the same data for K, P, f_o, Δf etc. as before. Then the I/W_{20} curves corresponding to $W_{40} = -1.75\lambda$ may be plotted for the different f values used. Fig.3.8 shows these I/W_{20} curves for f = 1,2,3,...,10.

At this stage, we have two sets of I/W_{20} curves corresponding to cases (1) and (2) as follows :

Case (1)

$$\left(\frac{W_{40}}{W_{60}}\right)_1 = \left(W_{40}\right)_1 = -1.5 - \delta W_{40} = -1.75$$

$$W_{60} = 1.0\lambda \; ; \delta W_{40} = 0.25$$

I/W_{20} curves are presented in Fig.3.8.

Case(2)

$$\left(\frac{W_{40}}{W_{60}}\right)_2 = (W_{40})_2 = -1.5+\delta W_{40} = -1.25$$

$W_{60} = 1.0\lambda$; $\delta W_{40} = 0.25$ as in case (1) above

I/W_{20} curves are presented in Fig.3.5.

According to (3.39) and (3.40), the condition for $|I_2|=|I_1|$ corresponding to the above cases (1) and (2) is

$$(W_{20})_2 = (W_{20})_1 - 0.5$$

for all f values. This has indeed found to be exactly true (to 7 or 8 places of decimals) for all f values (f=1,2,3,...,10); intercomparisons between Fig.3.5 and Fig.3.8 show that the curves in one graph are the exact replica of the other, with the W_{20} axis shifted by an amount equal to 0.5. Thus, for example, when f=5, we have :

In Fig. 3.8 $(W_{20})_1 = 0.80$; $I = |I_1| = 0.34907649$

In Fig. 3.5 $(W_{20})_2 = 0.30$; $I = |I_2| = 0.34907648$

showing $|I_1|=|I_2|$, $\delta W_{40} = 0.25$

It may be noted that this provides further evidence for the accuracy of the method of numerical integration used for evaluating I.

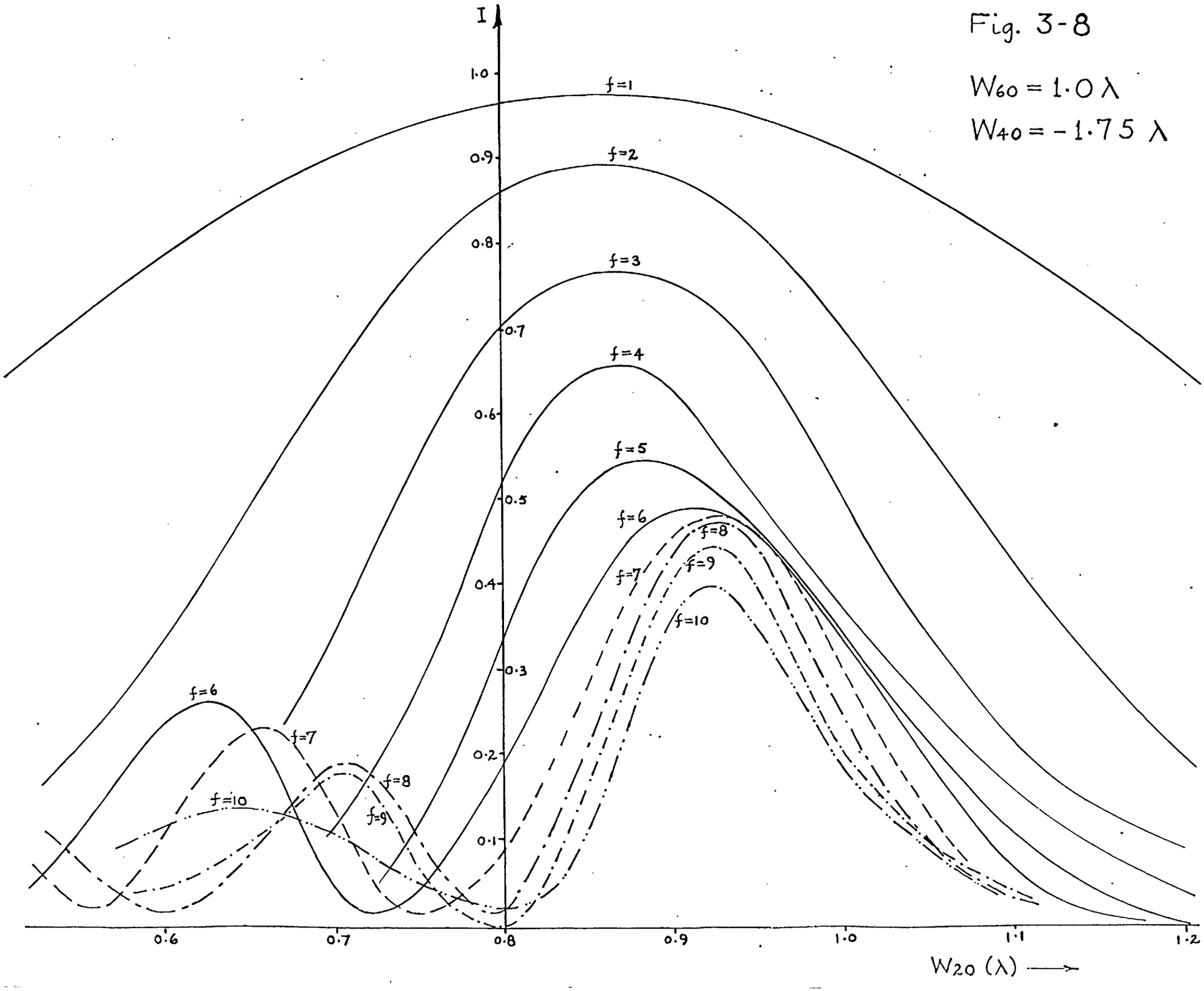
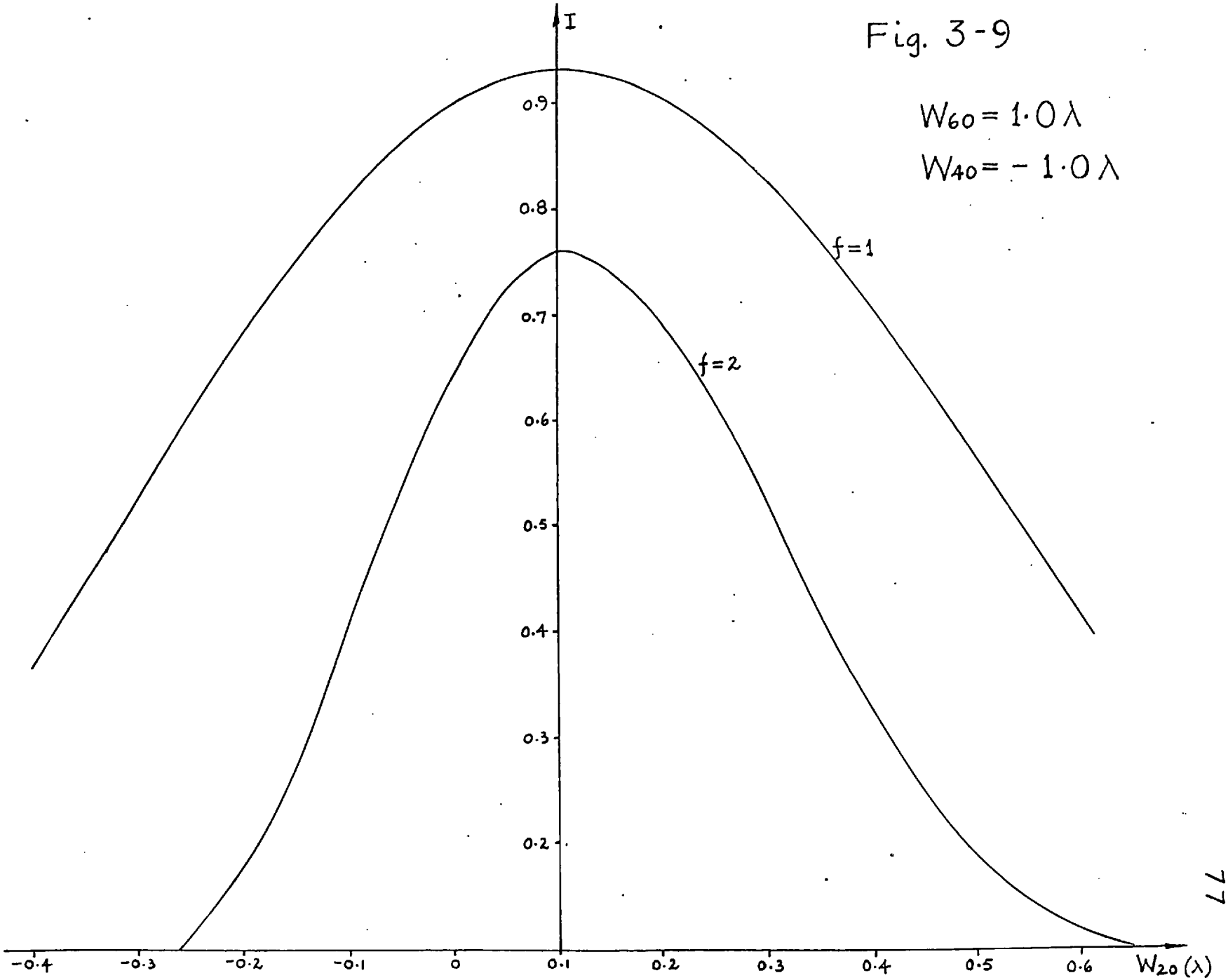
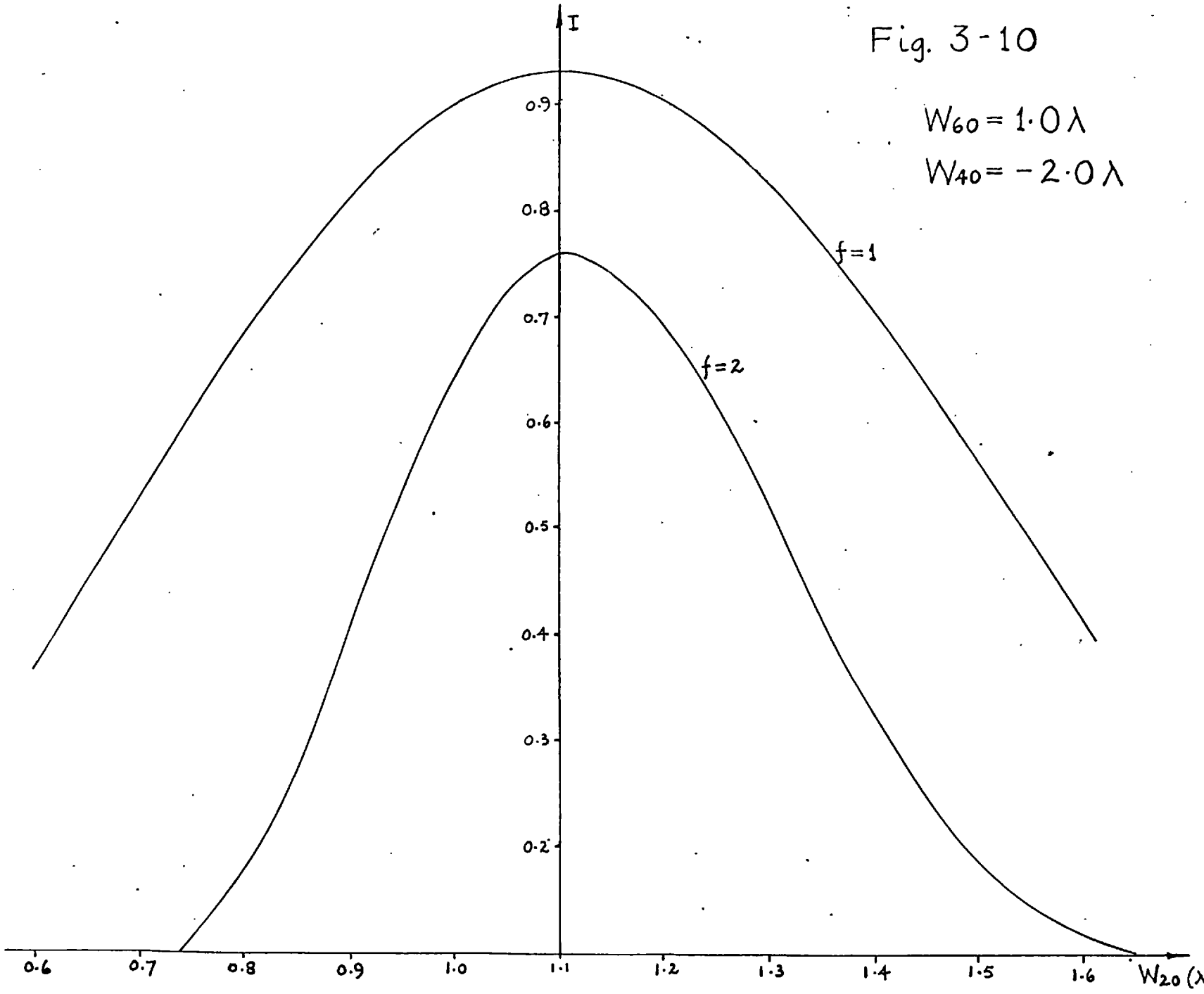


Fig. 3-8

$W_{60} = 1.0\,\lambda$

$W_{40} = -1.75\,\lambda$


Fig. 3-9
$W_{60} = 1.0\lambda$
$W_{40} = -1.0\lambda$
I
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
f=1
f=2
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
$W_{20}(\lambda)$


Fig. 3-10
$W_{60} = 1.0\lambda$
$W_{40} = -2.0\lambda$
I
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
f=1
f=2
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
$W_{20}(\lambda)$

As a further check, δW_{40} was set to 0.5, and two values of f(=1 and 2) were used. For each f value, we have

Case (1)

$(W_{40})_1 = -1.5 - 0.5 = -2.0$

$W_{60} = 1.0\lambda$, and the I/W_{20} curve is as shown in Fig. 3.10.

Case (2)

$(W_{40})_2 = -1.5 + 0.5 = -1.0$

$W_{60} = 1.0\lambda$, and the I/W_{20} curve is as shown in Fig. 3,9 .

Direct comparison between Fig. 3.9 and Fig. 3.10 again confirms the $I_1 = I_2$ property, corresponding to $(W_{20})_2 = (W_{20})_1 - 1.0$. Thus, for example, when f = 2, we have

In Fig. 3.10, $(W_{20})_1 = 1.0$ $\quad I = |I_1| = 0.64697311$

In Fig. 3.9, $(W_{20})_2 = 0.0$ $\quad I = |I_2| = 0.64697310$

showing $|I_1| = |I_2|$, $\delta W_{40} = 0.5$

Conclusion (for section III-vi) [SEE FOOTNOTE ON P. 71]

The special property $|I_1| = |I_2|$ described above is clearly seen to be true for all values of f and δW_{40}, provided that the conditions (3.39) and (3.40) are satisfied. We may thus conclude that, if the axial wave aberrations of two different systems may be represented by the polynomial of the form given in (3.38), and if their aberration coefficients satisfy (3.39) and (3.40) respectively, then the modulus of

the Strehl Intensity Ratio, $|I|$, on the axis will be identical for both systems.

REFERENCES

(1) H.H. Hopkins, The Numerical Evaluation of the Frequency Response of Optical System, Proc.Phy.Soc.B. Vol LXX, P.1002 (1957).

(2) "Table of Fresnel Integral to six dec. places". Compiled by T. Pearcey, Cambridge Press (1956).

(3) M. Maréchal, Thesis, University of Paris (1948).

(4) R.Richter, 1925, Z. InstrumKde, 45, 1.

(5) E.H.Linfoot, Optica Acta, vol.10, P.84, 1963

Chapter 4.

The Use of Canonical Co-ordinates in Aberration Calculations, Ray-tracing Formulae and Pupil Exploration.

(IV-i) Object and Image Space Variables.

In this section, we are not concerned directly with ray-tracing formulae beyond assuming that an algebraic method is employed in which the direction of a ray is specified by means of direction cosines (L,M,N) and intersection points at interfaces are specified by rectangular co-ordinates (X,Y,Z). The formulae considered here relate to specifying rays in the object and image planes and relative co-ordinates for the entrance and exit pupils.

A) Variables and axial object and image points.

The general notations associated with the current surface number "K" may be summarised as follows:

The medium immediately to the left of the K^{th} surface is numbered as the K^{th}, and that to the right is the $(K+1)^{th}$. u and u' refer to the paraxial convergence angles in the K^{th} and $(K+1)^{th}$ media, n and n' the corresponding refractive indices. d' is the separation between the K^{th} and the $(K+1)^{th}$ surfaces, and h, h_{+1} are the corresponding paraxial ray heights at these two surfaces. For a paraxial pupil ray, the notations $(\bar{h},\bar{u})$ are used. C is the surface curvature and, like other quantities, it may be written as C_k to be specific. When a quantity is associated with a general wavelength, a subscript "λ" may be used. The subscript "$\bar{\lambda}$" indicates the mean (or reference) wave-

length.

The surface of the entrance pupil is always identified for any object point O with the reference sphere centred on the object point O and of radius $\vec{EO}$ (Fig.4-1). The exit pupil is identified with the reference sphere, centred at the image point O' for the mean wavelength ($\bar{\lambda}$), and of radius $E\vec{'}O'$. Consider a ray from O intersecting the entrance pupil at the point B (X,Y) in Fig. 4-2. Using the relative pupil co-ordinates introduced in (2-2) and (2-5),we have

$$x = \frac{X}{h} \quad ; \quad y = \frac{Y}{h} \tag{4-1}$$

and similarly

$$x' = \frac{X'}{h'} \quad ; \quad y' = \frac{Y'}{h'} \tag{4-2}$$

defining the relative entrance and exit pupil co-ordinates of the ray. As mentioned in Chapter 2, the values of h and h' are the paraxial incidence heights of the paraxial marginal ray as calculated at the pupil planes E and E'.

At the start of the axial ray trace, one requires in practice only the value of y to be known and x may be set to zero. For completeness, however, formulae will be given for a 3-D case.

In Fig. 4-2, the direction cosines of the ray $OB\underline{P}_1P_1$ are given by

$$\begin{aligned} L_1 &= - u_1 x \\ M_1 &= - u_1 y \\ N_1 &= + \left[1-(L_1^{\,2} + M_1^{\,2}) \right]^{\frac{1}{2}} \end{aligned} \tag{4-3}$$

and this ray intersects the first surface at P_1. $\underline{P}_1$ is the point $(\underline{X}_1, \underline{Y}_1, \underline{Z}_1)$ such that $\underline{P}_1A_1$ is equally inclined to $O_1\underline{P}_1$ and O_1A_1. Noting that the $\triangle$'s OBE and $OB\underline{P}_1A_1$ are

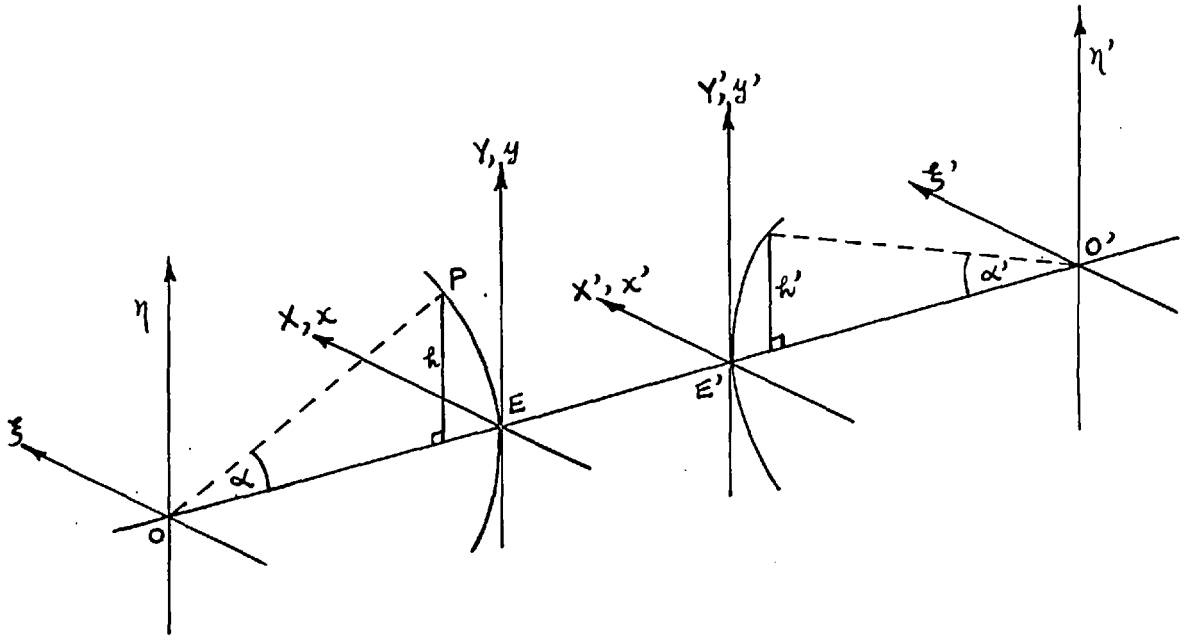


Fig 4-1 CO-ORDINATES FOR AXIAL O & O'

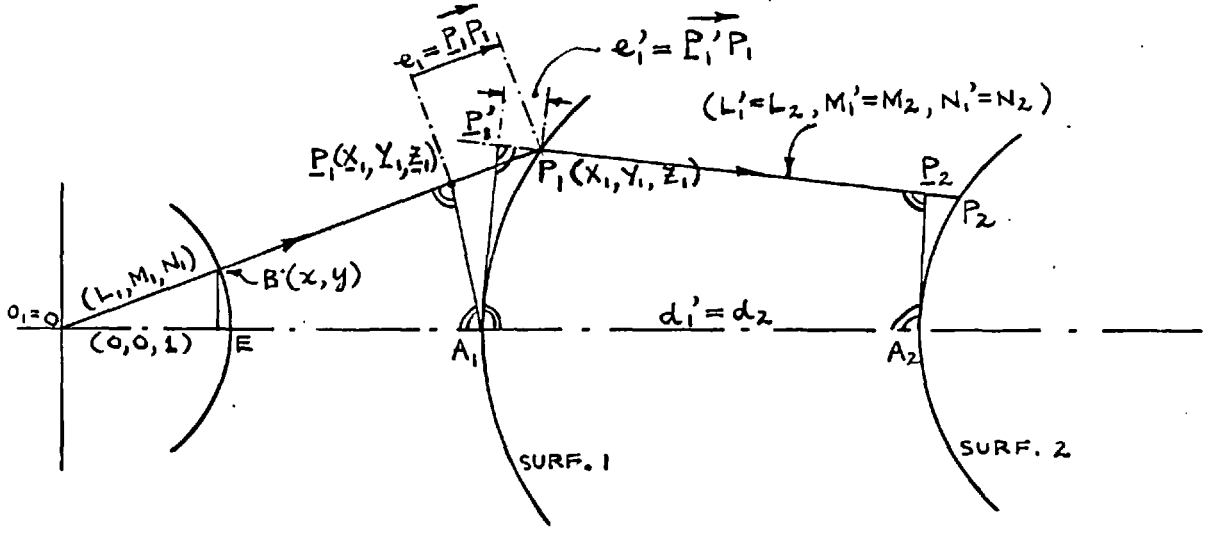


Fig 4-2 OPENING FORMULAE: AXIAL OBJECT POINTS

similar, we have

$$\underline{X}_1 = h_1 x$$
$$\underline{Y}_1 = h_1 y$$
$$\underline{Z}_1 = -\left\{\frac{L_1\underline{X}_1 + M_1\underline{Y}_1}{1 + N_1}\right\} \qquad (4\text{-}4)$$

as the co-ordinates of $\underline{P}_1$ relative to the pole of the first surface A_1, which is now regarded as our new origin of co-ordinates. By using the object space transfer equation to be described later, the length $\vec{\underline{P}_1 P_1} = e_1$ can be calculated and the co-ordinates (X_1, Y_1, Z_1) of P_1 determined. By applying Snell's Law, the direction cosines of the refracted ray $P_1\underline{P}_1'P_2$ may be found. The length $e_1' = \vec{\underline{P}_1' P_1}$ is next to be calculated, where the point $\underline{P}_1'$ is such that $P_1'A_1$ is equally inclined to the refracted ray and the axis. e_1' is given by*

$$e_1' = \frac{L_1' X_1 + M_1' Y_1' + (N_1' + 1)Z_1}{1 + N_1'} \qquad (4\text{-}5)$$

The total monochromatic aberration produced on refraction, referred to the invariant foci,[1] is given by

$$\Omega' - \Omega = n_1' e_1' - n_1 e_1 \quad \text{(at 1st surf.)} \qquad (4\text{-}7)$$
$$= \Delta(ne) \quad \text{(for any surf.)}$$

*NOTE. Equation (4-5) is a special case of the general formula:

$$e = \frac{\sum(L + \bar{L}).(X - \bar{X})}{1 + \sum L\bar{L}} \qquad (4\text{-}6)$$

given by Hopkins (Proc. Phys. Soc., B, Vol. LXV, P934, 1952) where (L, M, N), $(\bar{L}, \bar{M}, \bar{N})$ are the direction cosines of a ray and the pupil ray; (X, Y, Z), $(\bar{X}, \bar{Y}, \bar{Z})$ are the co-ordinates of the points of incidence P and $\bar{P}$ of the two rays (see Fig. 4-2A).

Using Conrady's differential formula, the corresponding chromatic aberration is given by

$$\begin{aligned} \partial \Omega_\lambda' - \partial \Omega_\lambda &= \delta n_1' \cdot e_1' - \delta n_1 e_1 \quad \text{(at 1st surf.)} \\ &= \Delta (\delta n \cdot e) \quad \text{(for any surf.)} \end{aligned} \tag{4-8}$$

The quantity e' is also used in the following transfer formulae (4-9) to locate the co-ordinates of the point $\underline{P}_2$, which is a point corresponding to $\underline{P}_1$ but at the second surface (see Fig. 4-2). The co-ordinates of $\underline{P}_2$ relative to the vertex A_2 as origin are

$$\begin{aligned} \underline{X}_2 &= X_1 + L_2 (d_2 - e_1') \\ \underline{Y}_2 &= Y_1 + M_2 (d_2 - e_1') \\ \underline{Z}_2 &= -\left(\frac{L_2 \underline{X}_2 + M_2 (\underline{Y}_2)}{1 + N_2} \right) \end{aligned} \tag{4-9}$$

since $L_1' = L_2$, $M_1' = M_2$, $N_1' = N_2$. The formula for $\underline{Z}_2$ in (4-9) is a special case of that described in Fig. 4-2A, and always gives an accurate value for $\underline{Z}$ on transfer.

Having determined the point $\underline{P}_2$, one proceeds to calculate e_2 and e_2' in a similar manner to that of the previous surface, and then following this by locating the point $\underline{P}_3$ for the next surface etc. As the ray trace progresses, the aberrations at each surface as given by (4-7) and (4-8) are calculated. The total aberrations along the ray at the end of the system (kth surface) are given by

$$\Omega'_k = \sum_{i=1}^{k} \Delta (ne)_i \tag{4-10}$$

$$\partial W_\lambda = (\partial \Omega_\lambda')_k = \sum_{i=1}^{k} \Delta (\partial n \cdot e)_i \tag{4-11}$$

noting that $\Omega_1 = 0$ for an aberration-free object point. It is desirable to refer the monochromatic aberrations along different rays to a common reference focus (e.g. the gaussian focus of the $\bar{\lambda}$ ray) and for this a focal shift term $\delta\Omega^{(1)}$ has to be added to (4-10). The monochromatic aberration W is thus given by

$$W = \sum_{i=1}^{k} \Delta (ne)_i + \delta\Omega' \qquad (4\text{-}10)A$$

To calculate $\delta\Omega'$, one needs to know the relative co-ordinates $(\underline{x}', \underline{y}', \underline{z}')$ of the point $\underline{B}'$ (Fig. 4-3). $\underline{B}'$ is such that $\underline{B}'\,\underline{E}'$ is equally inclined to the emergent ray $B_k V'$ and the axis $A_k O'$. In the image space, the values of (L_k', M_k', N_k') and the co-ordinates of $\underline{P}_k'$ $(\underline{X}_k', \underline{Y}_k', \underline{Z}_k')$ are known from the ray trace. Since $A_k E' = \underline{P}_k' \underline{B}' = \bar{h}_k / \bar{u}_k'$ (obtained from the paraxial pupil ray), the relative co-ordinates $(\underline{x}', \underline{y}', \underline{z}')$ of $\underline{B}'$ are given by

$$\begin{aligned}
\underline{x}' &= \frac{\underline{X}'}{h'} = \frac{n_k'}{H}\left\{ (-\bar{u}_k')\, \underline{X}_k' - L_k' \bar{h}_k \right\} \\
\underline{y}' &= \frac{\underline{Y}'}{h'} = \frac{n_k'}{H}\left\{ (-\bar{u}_k')\, \underline{Y}_k' - M_k' \bar{h}_k \right\} \qquad (4\text{-}12) \\
\underline{z}' &= \frac{\underline{Z}'}{h'} = -\left\{ \frac{L_k' \underline{x}' + M_k' \underline{y}'}{N_k' + 1} \right\}
\end{aligned}$$

noting that the paraxial marginal ray height h' at E' may be obtained from the Lagrange invariant H given by

$$H = - n' \bar{u}' h' \qquad (4\text{-}13)$$

Referring again to Fig. 4-3, the distance

$$\delta' = \underline{B}'\vec{B}'$$

is measured from $\underline{B}'(\underline{x}', \underline{y}', \underline{z}')$ to the point B' on the

$$e = \frac{\Sigma[(L+\bar{L})\cdot(X-\bar{X})]}{1+\Sigma L\bar{L}}$$

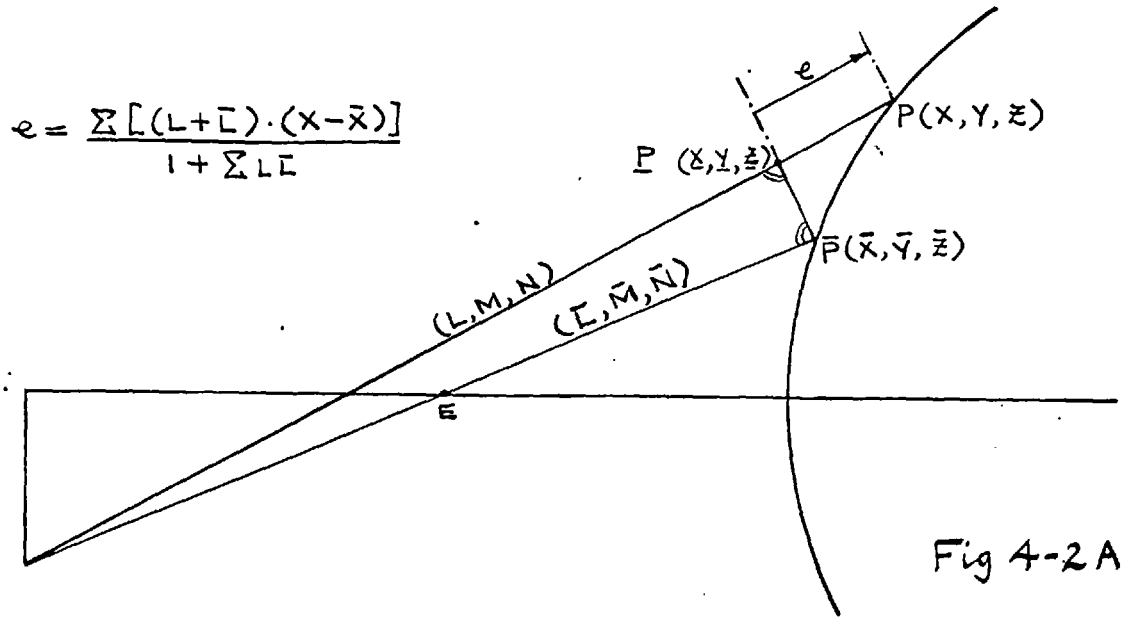


Fig 4-2A

The co-ordinate $\underline{Z}$ of $\underline{P}$ may be calculated by putting $e=0$ (P moves to $\underline{P}$), leading to

$$\underline{Z} = -\frac{(L+\bar{L})(\underline{X}-\bar{X})+(M+\bar{M})(\underline{Y}-\bar{Y})}{(N+\bar{N})} + \bar{Z}$$

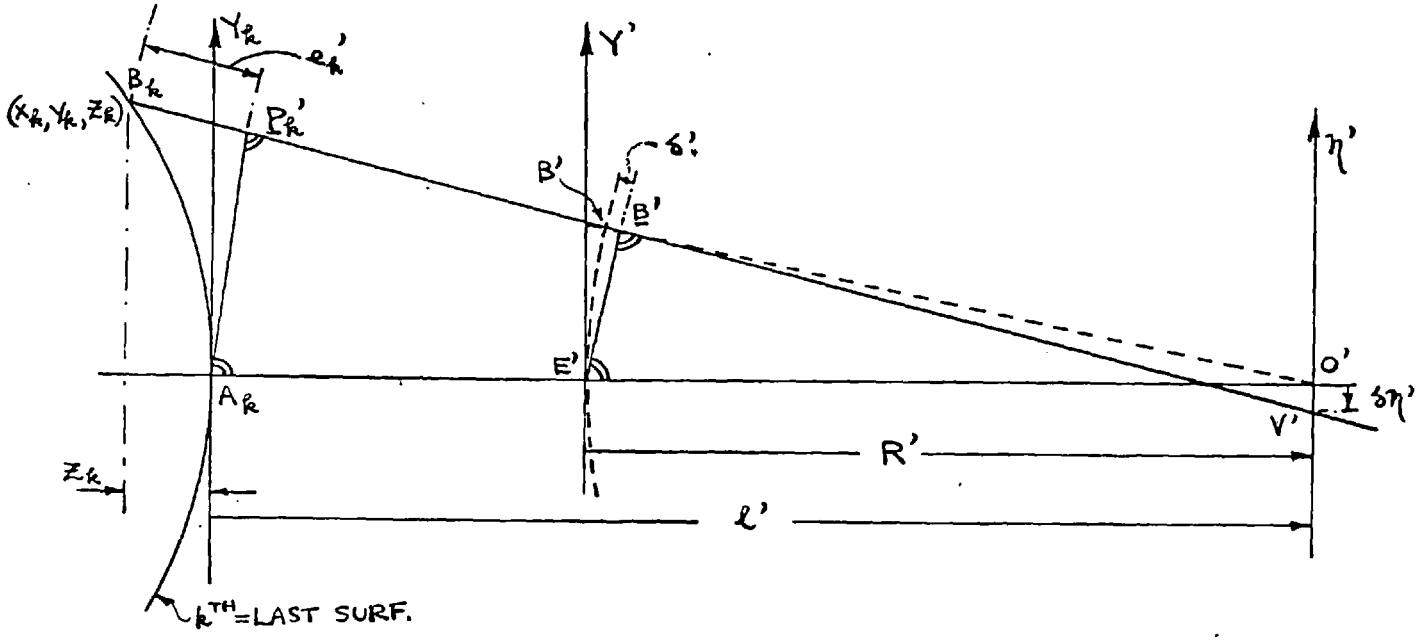


Fig 4-3 CALCULATION OF PUPIL CO-ORDINATES & TRANSVERSE ABERRATION IN AXIAL IMAGE

reference sphere centred on O', which is the gaussian image focus for $\bar{\lambda}$ light. The focal shift term $\delta\Omega'$ in (4-10)A is given by

$$\delta\Omega' = - n'_k \, \delta'$$

Since the point B' is such that E'B' is equally inclined to B'O' and E'O', the distance δ' from the known point $\underline{B}'$ to B' may be found by using a special form of the formula (4-6). It can be shown[2] that

$$\delta\Omega' = - \frac{(\underline{x}' - L'p') \, \delta G'_o + (\underline{y}' - M'p') \, \delta H'_o}{1 + N'} \qquad (4\text{-}14)$$

where $p' = L'\underline{x}' + M' \, \underline{y}' + N'\underline{z}'$

all quantities in (4-14) are referred to the image space, leaving the subscript k to be understood. In (4-14), $\delta G'_o$ and $\delta H'_o$ are the reduced transverse ray aberrations defined by

$$\delta G'_o = n'_\lambda \, u'_\lambda \, \delta\xi'_\lambda \; ; \; \delta H'_o = n'_\lambda \, u'_\lambda \delta\eta'_\lambda \qquad (4\text{-}15)$$

where $(\delta\xi', \delta\eta')$ are the co-ordinates of the λ ray intersection point in the gaussian image plane for $\bar{\lambda}$ (in Fig. 4-3, only $\delta\eta'$ is shown). $\delta G'_o$ and $\delta H'_o$ are calculated from

$$\delta H'_o = \pi' \, H'_o \; ; \; \delta G'_o = \pi' \, G'_o \qquad (4\text{-}15)A$$

with $\pi' = H_\lambda / \tilde{H}'_{o\lambda}$ as the aperture normalizing factor,

$$\tilde{H}'_{o\lambda} = n'_\lambda \, (\bar{h}_\lambda \, u'_\lambda - h_{\bar{\lambda}} \, \bar{u}'_\lambda)_k \qquad (4\text{-}16)$$

$$= n'_{k\bar{\lambda}} \, u'_{k\bar{\lambda}} \, \eta'_\lambda$$

$$H_\lambda = n_\lambda \, u_\lambda \, \eta_\lambda = n'_\lambda \, u'_\lambda \, \eta'_\lambda$$

$$H'_o = n'_{\bar{\lambda}}\, u'_{\bar{\lambda}}\, \delta\eta'_{\lambda}$$

$$= \frac{n'_{\bar{\lambda}}}{N'_k}\left\{ (N'Y - M'Z)\, u'_{\bar{\lambda}} + M'h_{\bar{\lambda}} \right\}_k \qquad (4\text{-}16)A$$

(noting that $\delta\eta' = Y_k + \left\{\frac{\ell' - Z'_k}{N'_k}\right\} M'_k$ in Fig. 4-3)

similarly,

$$G'_o = n'_{\bar{\lambda}}\, u'_{\bar{\lambda}}\, \delta\xi'_{\lambda}$$

$$= \frac{n'_{\bar{\lambda}}}{N'_k}\left\{ (N'X - L'Z)\, u'_{\bar{\lambda}} + L'\, h_{\bar{\lambda}} \right\}_k$$

Finally the relative co-ordinates of B' are given by

$$x' = \underline{x}' + L'\, \frac{\bar{u}'}{H}\, \delta\Omega'$$

$$y' = \underline{y}' + M'\, \frac{\bar{u}'}{H}\, \delta\Omega' \qquad (4\text{-}17)$$

$$z' = \underline{z}' + N'\, \frac{\bar{u}'}{H}\, \delta\Omega'$$

and equation (4-13) again being used for $1/h'$.

It is worth noting that the variables introduced above for axial ray tracing will always remain determinate and meaningful when, for example, either the object, the image or one of the pupils, or both, is at infinity. As mentioned earlier, only meridian rays with $x = x' = o$ are traced in practice for axial object and image points. However the general ray case considered above facilitates comparison with the gen-

eralization to extra-axial points.

B) Variables for extra-axial object and image points.

In the axial ray trace described above, the starting data required are (u_1, h_1) and $(\bar{u}_1, \bar{h}_1)$, and the paraxial marginal and pupil rays are then traced. For an extra-axial object point and its image, a paraxial sagittal ray (traced along the pupil ray) is employed, and the starting data required are $(u_s)_1$, $(h_s)_1$. In practice, this sagittal ray need not be traced explicitly since only the incidence heights h_s, h'_s (in place of h , h') are required here for the purpose of pupil normalisation. A pupil ray, corresponding to the given field, is first traced to provide the necessary data for $(u_s)_1$, $(h_s)_1$. The formulae for the latter are:

$$(u_s)_1 = \frac{\bar{H}_s}{\bar{H}_o} \; u_1 \qquad (4\text{-}18)$$

$$(h_s)_1 = \frac{1}{\bar{N}_1} \cdot \frac{\bar{H}_s}{\bar{H}_o} \; (h_1 - u_1 \bar{Z}_1)$$

where $\bar{N}_1$ is the direction cosine (relative to the axis) of the pupil ray, and $\bar{Z}_1$ is the Z co-ordinate of the intersection point $\bar{P}_1$ of this ray at the first surface (Fig. 4-6). The pupil ray in question has been corrected for assymmetrical vignetting and this will be described in Chapter 7. The quantity $\bar{H}_o$ is the reduced object height normalised by nu , i.e.

$$\begin{aligned} \bar{H}_o = nu\,\bar{\eta} &= n(\bar{h}_1 u_1 - h_1 \bar{u}_1) \\ &= H \\ &= \mathbf{n}u\eta \end{aligned} \qquad (4\text{-}19)$$

since the object height $\bar{\eta} = \eta$ in the object space. In the image space, however, $\bar{\eta}' \neq \eta'$ (i.e. the actual and gaussian image heights are different) due to the presence of distortion, so that

$$\bar{H}'_o = n' u' \bar{\eta}' \neq H \qquad (4\text{-}20)$$

indicating that $\bar{H}_o$ is not an invariant. In fact, for the extra-axial ray case the true invariant of the system is $\bar{H}_s$, and this is defined[3] by

$$\bar{H}_s = nu_s\eta_s = n'u'_s\eta'_s \qquad (4\text{-}21)$$

where η_s, η'_s are the heights of the sagittal foci above the axis in the object and image spaces as shown in Fig. 4-4. It is convenient here to note that alternative expressions for $\bar{H}_s$ are:

(a) $$\bar{H}_s = nh_s \bar{M} = n'h'_s \bar{M}' \qquad (4\text{-}22)$$

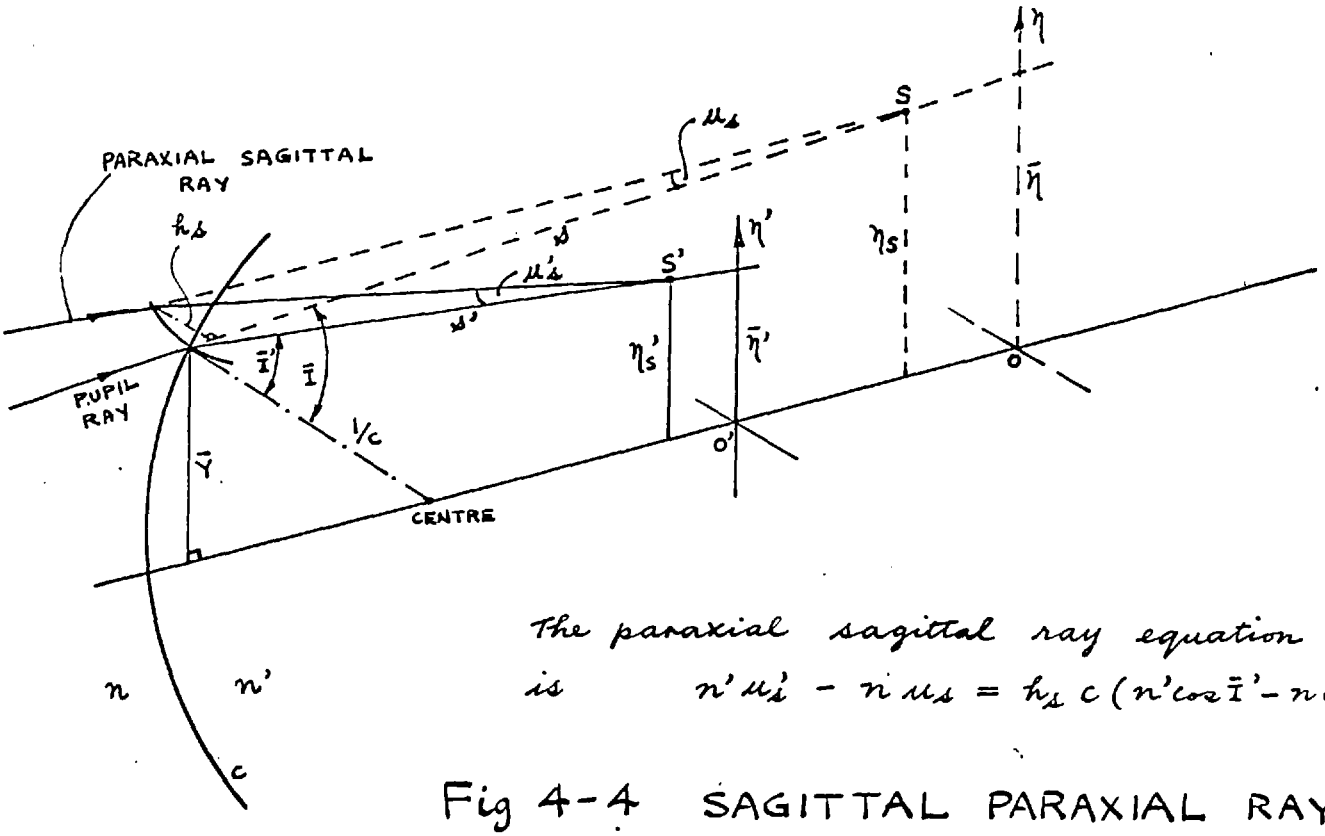
corresponding to (4-13), where $\bar{M}$ is the direction cosine of the pupil ray, and paraxially, $\bar{M} \rightarrow (-\bar{u})$, $\bar{M}' \rightarrow (-\bar{u}')$.

(b) From (4-22) and Fig. 4-5, it is easily seen that

$$\bar{H}_s = n \sin\alpha_s \bar{\eta} = n' \sin\alpha'_s \bar{\eta}' \qquad (4\text{-}23)$$

where the quantities $\sin\alpha_s$ and $\sin\alpha'_s$ are as defined in Chapter 2. Note that (4-23) gives an exact relation between the actual object and image heights. Since $\eta_s = \bar{\eta}$ in the object space, (4-21) and (4-23) requires $(u_s)_1 = u_s = \sin\alpha_s$. It can now be seen that the expression for $(u_s)_1$ in (4-18) follows directly from (4-19) and (4-23). Also from Fig. 4-6, the paraxial sagittal ray height $(h_s)_1$ at surface one is simply given by $(u_s)_1 \cdot \bar{Q}\bar{P}_1$.

The relative pupil co-ordinates (x_s, y_s) to be used in the ray trace starting data, and the calculation of



The paraxial sagittal ray equation is $n'u'_s - nu_s = h_s c(n'\cos\bar{I}' - n\cos\bar{I})$

Fig 4-4 SAGITTAL PARAXIAL RAY

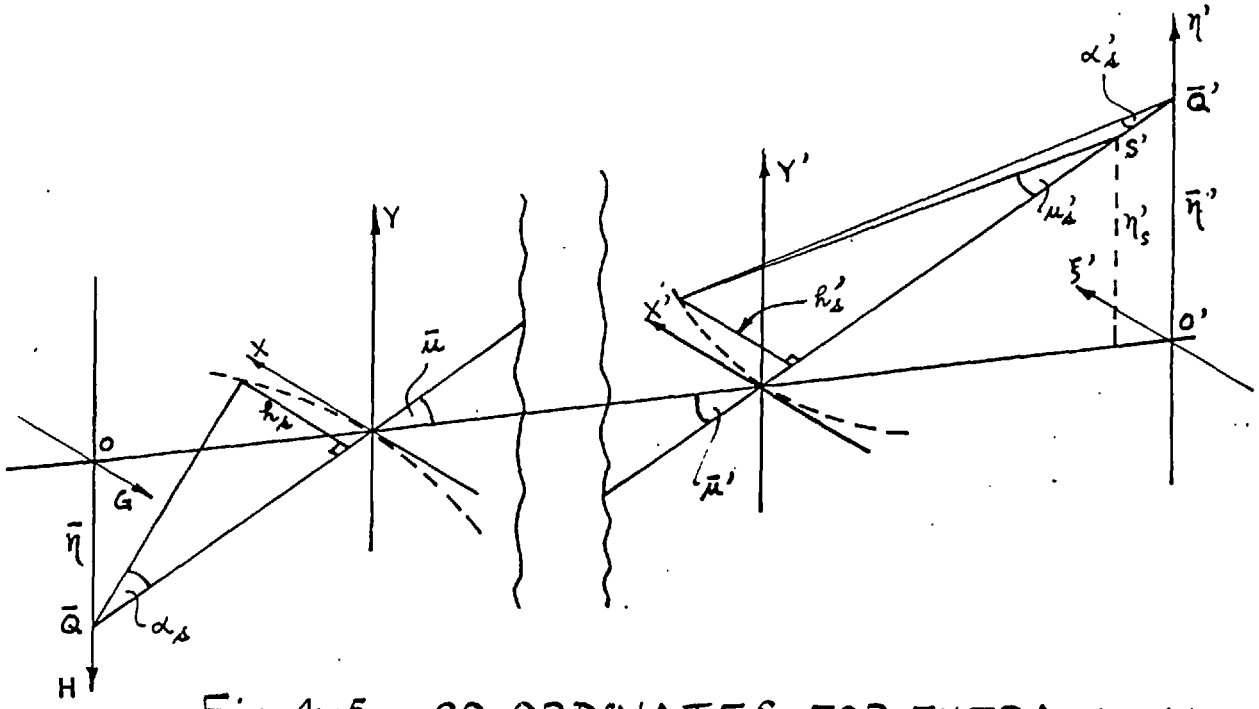


Fig 4-5 CO-ORDINATES FOR EXTRA-AXIAL OBJECT & IMAGE

$\bar{H}_s$ are all required to be modified in the presence of pupil vignetting. This will be discussed in Chapter 7, but at present it may be assumed that (x_s, y_s) and $\bar{H}_s$ have already been so modified.

In Fig. 4-8, the point $\underline{P}\ (\underline{X},\underline{Y},\underline{Z})$ is defined to be such that $\underline{P}\bar{P}_1$ is equally inclined to the ray $\bar{Q}B\underline{P}P_1$ that is desired to trace, and the pupil ray $\bar{Q}\bar{P}_1$. The opening formulae are then:

$$\begin{aligned} L_1 &= -(u_s)_1\ x_s \\ M_1 &= \bar{M}_1 - y_s\ (u_s)_1 \\ N_1 &= +\left[\ 1-(L_1^2+ M_1^2)\ \right]^{\frac{1}{2}} \end{aligned} \qquad (4\text{-}24)$$

$$\begin{aligned} \underline{X} &= (h_s)_1\ x_s \\ \underline{Y} &= \bar{Y}_1 + (h_s)_1\ y_s \\ \underline{Z} &= \bar{Z}_1 - \left[\frac{L_1\underline{X} + (M_1 + \bar{M}_1)(\underline{Y} - \bar{Y}_1)}{N_1 + \bar{N}_1}\right] \end{aligned} \qquad (4\text{-}25)$$

and the distance $\underline{P}P_1 = \tilde{e}_1$ is calculated, and this quantity is defined relative to the pupil ray. It is preferable to trace rays relative to the axis and to have the quantities e_K and e'_K found at each surface to be relative to the axis. In Fig. 4-8, the distance $e_1 = \underline{\vec{P}_1}P_1$ is required; the point $\underline{P}_1$ is such that $\underline{P}_1A_1$ is equally inclined to the ray $\bar{Q}B\underline{P}_1\underline{P}P_1$ and the axis. The co-ordinates $(\underline{X}_1, \underline{Y}_1, \underline{Z}_1)$ of $\underline{P}_1$ are used as the true starting data, and these are

$$\begin{aligned} \underline{X}_1 &= \underline{X} - L_1\ e \\ \underline{Y}_1 &= \underline{Y} - M_1\ e \\ \underline{Z}_1 &= -\left\{\frac{L_1\ \underline{X}_1 + M_1\ \underline{Y}_1}{1 + N_1}\right\} \end{aligned} \qquad (4\text{-}26)$$

where e is the distance $\underline{P}_1P$ given by

$$(u_s)_1 = \left(\frac{\bar{H}_s}{\bar{H}_o}\right) u_1$$

$$(h_s)_1 = (u_s)_1 \left(\frac{\ell_1 - \bar{z}_1}{\bar{N}_1}\right)$$

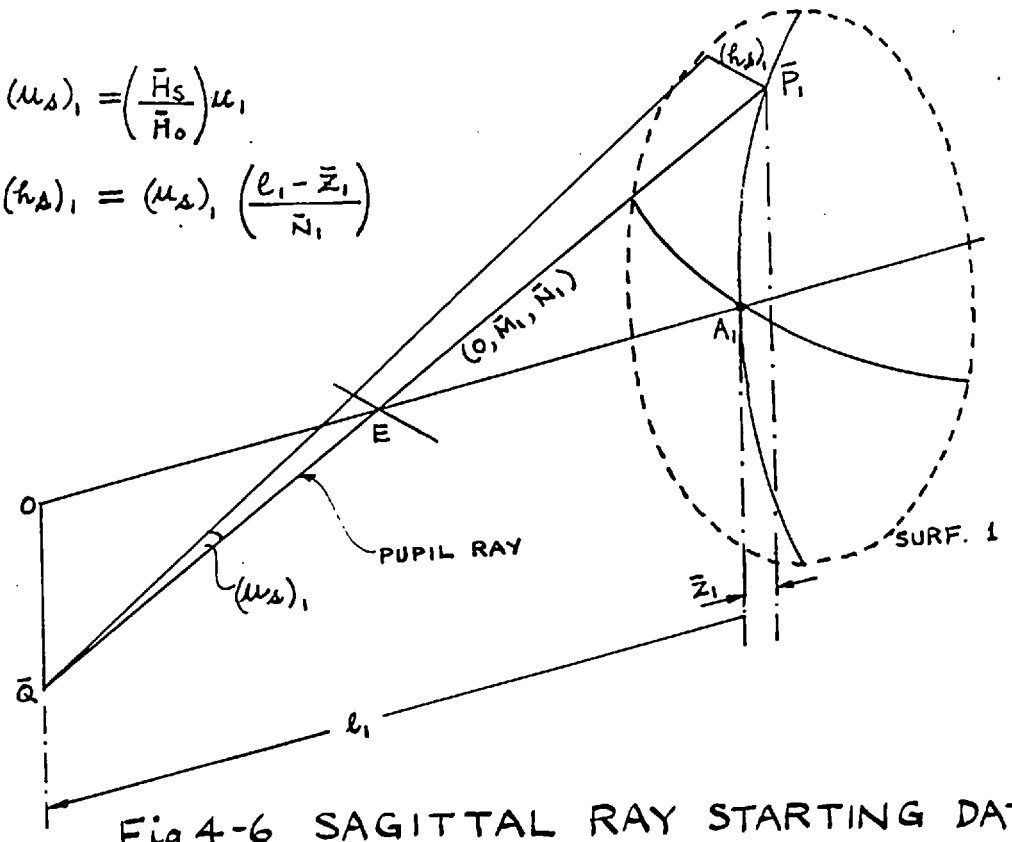


Fig 4-6 SAGITTAL RAY STARTING DATA

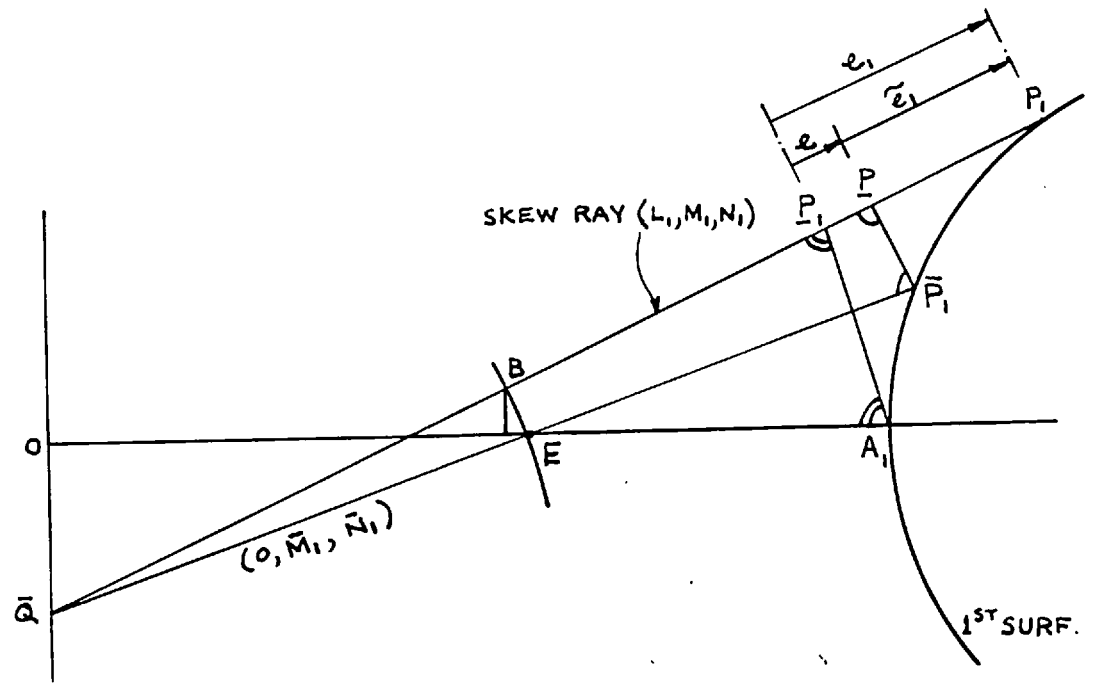


Fig 4-8 OPENING FORMULAE:
EXTRA-AXIAL OBJECT POINTS

$$e = \frac{L_1 \underline{X} + M_1 \underline{Y} + (1 + N_1)\, \underline{Z}}{1 + N_1} \qquad (4\text{-}27)$$

The point $\underline{P}_1$ here is essentially the same as that shown in Fig. 4-2 for the axial ray case, and the ray tracing procedure involved here is also similar to the previous case.

In the image-space (Fig. 4-9), the ray trace will give the values of (L'_k, M'_k, N'_k) and the point P'_k (X'_k, Y'_k, Z'_k) at the last surface. The point $\underline{P}'_k$ is such that $\underline{P}'_k \bar{P}_k$ is equally inclined to the ray $P_k \underline{P}'_k V'$ and the pupil ray $\bar{P}_k \bar{Q}'$, and the distance $\tilde{e}'_k = \vec{P}_k' P_k$ is calculated, relative to the pupil ray. The co-ordinates of $\underline{P}'_k$ are then

$$\begin{aligned} \underline{X}'_k &= X'_k - \tilde{e}'_k L'_k \\ \underline{Y}'_k &= Y'_k - \tilde{e}'_k M'_k \end{aligned} \qquad (4\text{-}28)$$

The point $\underline{B}'$ $(\underline{X}'\ \underline{Y}'\ \underline{Z}')$ shown in Fig. 4-9 is such that $\underline{B}'E'$ is equally inclined to the ray $P_k \underline{B}'V'$ and the pupil ray $\bar{P}_k E' \bar{Q}'$. The relative co-ordinates of $\underline{B}'$ are given by

$$\begin{aligned} \underline{x}'_s &= \frac{\underline{X}'}{h'_s} = \frac{h'_k}{\bar{H}_s} \left\{ \bar{M}'\, \underline{X}' - L'\, \bar{Y} \right\}_k \\ \underline{y}'_s &= \frac{\underline{Y}'}{h'_s} = \frac{n'_k}{\bar{H}_s} \left\{ \bar{M}'\, \underline{Y}' - M'\, \bar{Y} \right\}_k \\ \underline{z}'_s &= \frac{\underline{Z}'}{h'_s} = -\left\{ \frac{L'\, \underline{x}'_s + (M' + \bar{M}')\, \underline{y}'_s}{N' + \bar{N}'} \right\}_k \end{aligned} \qquad (4\text{-}29)$$

noting that $\underline{P}'_k \underline{B}' = \bar{P}_k E' = -\bar{Y}_k / \bar{M}'_k$. The paraxial sagittal ray height h'_s at E' is obtained from equation (4-22).

Corresponding to the axial image case, the distance

$$\delta' = \underline{\vec{B}}'B'$$

is measured from $\underline{B}'$ $(\underline{x}'_s, \underline{y}'_s, \underline{z}'_s)$ to the point B' (x_s, y_s, z_s) on the reference sphere centred on $\bar{Q}'_{\bar{\lambda}}$, which is the point where the $\bar{\lambda}$ pupil ray intersects the $\bar{\lambda}$ gaussian image plane. The value of $\delta\Omega'$ in (4-10)A is given by

$$\begin{aligned}\delta\Omega' &= -\, n'_k\, \delta' \\ &= -\frac{1}{1+\bar{M}\bar{M}+\bar{N}\bar{N}}\Big\{\Big[\underline{x}_s - L\,(L\underline{x}_s + M\underline{y}_s + N\underline{z}_s)\Big]\delta G_s \\ &\quad + \Big[\underline{y}_s - M\,(L\underline{x}_s + M\underline{y}_s + N\underline{z}_s)\Big]\delta H_s \\ &\quad + \Big[\underline{y}_s - \bar{M}\,(\bar{M}\underline{y}_s + \bar{N}\underline{z}_s)\Big]\;\delta\bar{H}_s\Big\} \qquad (4\text{-}30)\end{aligned}$$

and all quantities in (4-30) are referred to the image space, leaving the subscript k and the prime to be understood. $\delta G'_s$ and $\delta H'_s$ are reduced transverse ray aberrations defined as

$$\begin{aligned}\delta G'_s &= n'_{\bar{\lambda}} \sin \tilde{\alpha}'_s \;\delta\xi' \\ \delta H'_s &= n'_{\bar{\lambda}} \sin \tilde{\alpha}'_s \;\delta\eta'\end{aligned} \qquad (4\text{-}31)$$

where $\delta\xi'$, $\delta\eta'$ are the co-ordinates of the intersection point of the ray relative to $\bar{Q}'_{\bar{\lambda}}$ (see Fig. 4-10), and $\tilde{\alpha}'_s$ is defined by

$$\sin \tilde{\alpha}'_s = \frac{(h'_s)_{\bar{\lambda}}}{\tilde{R}'} \qquad (4\text{-}32)$$

where $\tilde{R}'$ is the distance $E'_{\bar{\lambda}}\,\bar{Q}'_{\bar{\lambda}}$ in Fig. 4-10. $\delta G'_s$ and $\delta H'_s$ are calculated from

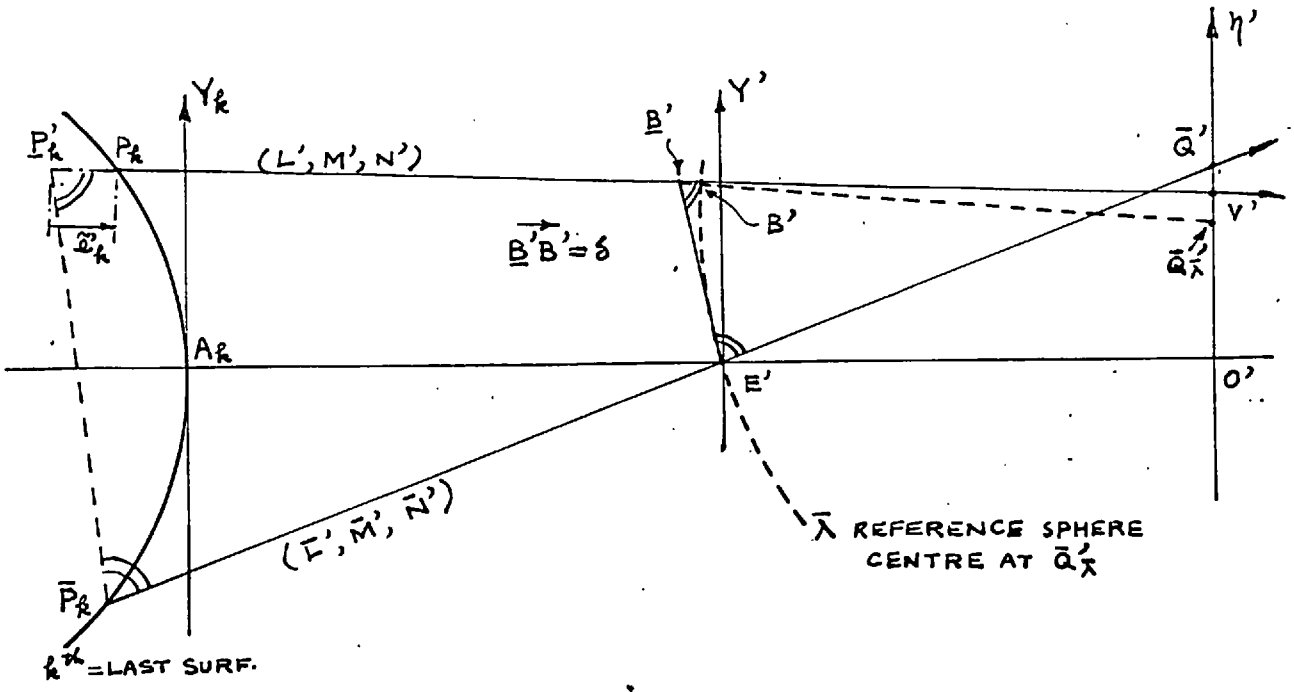


Fig. 4-9 CLOSING FORMULAE:
EXTRA-AXIAL IMAGE POINTS

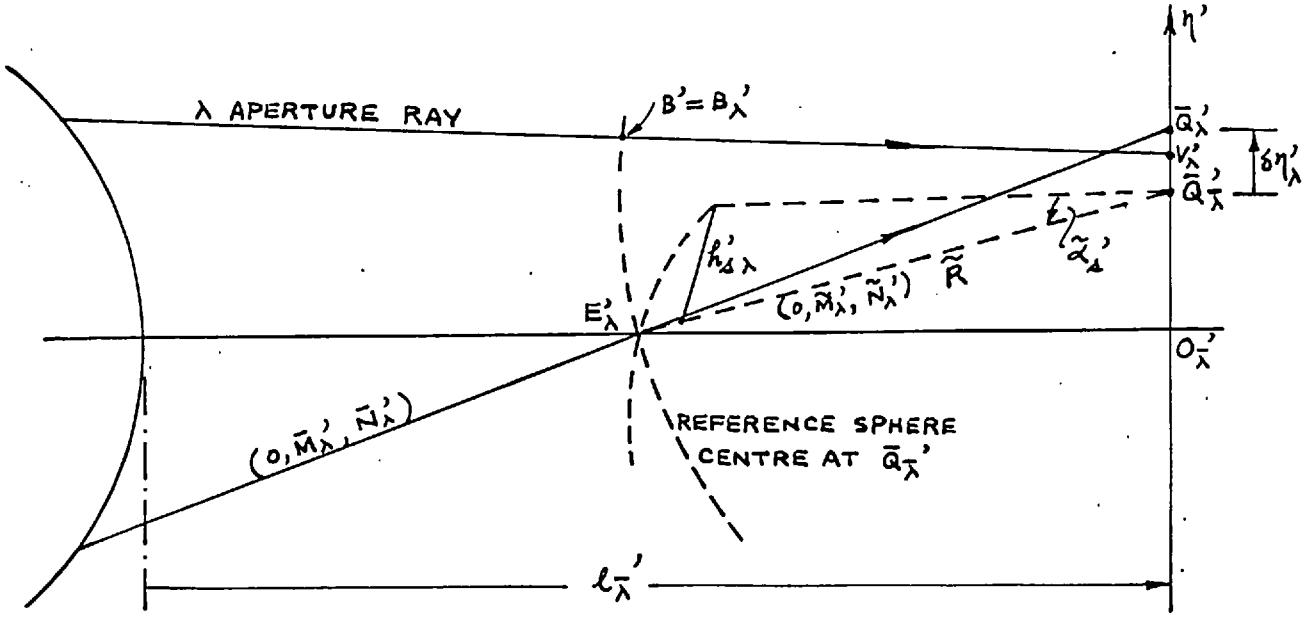


Fig 4-10

$$\delta G'_s = \pi' G'_o$$

$$\delta H'_s = \pi' \left\{ H'_o - (H'_o)_{\bar{\lambda}} \right\} \qquad (4\text{-}33)$$

with $$\pi' = \left\{ 1 + (\bar{M}'_k)^2 \frac{\delta \bar{H}'_0}{\bar{H}'_0} \right\} \frac{\bar{H}_s}{\bar{H}'_0} \qquad (4\text{-}34)$$

$$\delta \bar{H}'_o = \bar{H}'_{o\lambda} - \bar{H}'_{o\bar{\lambda}} \qquad (4\text{-}35)$$

$$\bar{H}'_{o\lambda} = \frac{n'_k}{\bar{N}'_k} \left\{ (\bar{N}' \bar{Y} - \bar{M}' \bar{Z}) u' + \bar{M}' h \right\}_k \qquad (4\text{-}36)$$

The expressions for G'_o and H'_o are given in (4-16)A.

It will be useful to mention here that the aperture normalising factor π' in (4-34) is introduced in the following manner. Starting with the expression for $\delta G'_s$ in (4-31)

$$\delta G'_s = n'_\lambda \sin \tilde{\alpha}'_s \, \delta \xi'_\lambda = \left[\frac{n'_\lambda \sin \tilde{\alpha}'_s}{n'_{\bar{\lambda}} \sin \alpha'_{\bar{\lambda}}} \right] n'_{\bar{\lambda}} u'_{\bar{\lambda}} \delta \xi'_\lambda \qquad (4\text{-}37)$$

since $u'_{\bar{\lambda}} = \sin \alpha'_{\bar{\lambda}}$. If we now define π' as

$$\pi' = \frac{n'_\lambda \sin \tilde{\alpha}'_s}{n'_{\bar{\lambda}} \sin \alpha'_{\bar{\lambda}}} \qquad (4\text{-}38)$$

and using (4-16)A, we have

$$\delta G'_s = \pi' G'_o$$

as in (4-33). It can be shown that the expression (4-38) leads to that given in (4-34)

Finally the relative co-ordinates of B' (Fig. 4-9) are given by

$$x'_s = \underline{x}'_s - L' \frac{\bar{M}' \delta\Omega'}{\bar{H}_s}$$

$$y'_s = \underline{y}'_s - M' \frac{\bar{M}' \delta\Omega'}{\bar{H}_s} \qquad (4\text{-}39)$$

$$z'_s = \underline{z}'_s - N' \frac{\bar{M}' \delta\Omega'}{\bar{H}_s}$$

where $\delta\Omega'$ is that from (4-30), and the value of h'_s

used for normalisation is again obtained from equation (4-22).

C) Calculation of wave aberrations.

For the axial case, the formulae for the monochromatic wave aberration W and chromatic aberration ∂W_λ are as given by (4-10)A and (4-11).

For extra-axial pancils, the formulae are somewhat more complicated. In the ray trace, the quantities e, e' are calculated at each surface, referred to the optical axis, together with $\tilde{e}_1$ and $\tilde{e}'_k$ calculated with reference to the pupil ray. The expression for the wave aberration, referred to the invariant focus in the image space, is

$$\Omega'_k = \sum_{i=1}^{k} \Delta(ne)_i - \sum_{i=1}^{k} \Delta(n\bar{e})_i + n_1 e_1 - n'_k e'_k - n_1 \bar{e}_1 + n'_k \bar{e}'_k + n'_k \tilde{e}'_k - n_1 \tilde{e}_1 \qquad (4\text{-}40)$$

where the quantities $\bar{e}, \bar{e}'$ denote the values of e for the pupil ray calculated relative to the optical axis.

The value of W is calculated relative to the image point $\bar{Q}'_\lambda$, and is given by

$$W = \Omega'_k + \delta\Omega' \qquad (4\text{-}41)$$

where $\delta\Omega'$ is obtained from (4-30).

The Conrady formula for the chromatic aberration for the extra-axial case may be derived from (4-40) by replacing n by δn, i.e.

$$\partial W_\lambda = \sum_{i=1}^{k} \Delta(\partial n.e)_i - \sum_{i=1}^{k} \Delta(\partial n.\bar{e})_i + \partial n_1 e_1 - \partial n'_k e'_k - \partial n_1 \bar{e}_1 + \partial n'_k \bar{e}'_k + \partial n'_k \tilde{e}'_k - \partial n_1 \tilde{e}_1 \qquad (4\text{-}42)$$

(IV-ii) Finite ray trace.

The surface equation used for a 'non-figured' surface is

$$c\left[X^2 + Y^2 + (1 + Q)\, Z^2\right] - 2\, Z = 0 \qquad (4\text{-}43)$$

c = vertex curvature

$Q > 0$ ellipsoid about major axis

$Q = 0$ sphere

$0 > Q > -1$ ellipsoid about minor axis

$Q = -1$ paraboloid

$-1 > Q$ hyperboloid

Equation (4-43) may be solved to give

$$Z = \frac{c(X^2 + Y^2)}{1 + \left\{1 - (1+Q)\, c^2(X^2+Y^2)\right\}^{\frac{1}{2}}} \qquad (4\text{-}44)$$

For a figured surface, we write

$$Z = Z_s(\rho) + A_4 p^2 + A_6 p^3 + \ldots\ldots \qquad (4\text{-}45)$$

where $\rho^2 = X^2 + Y^2$

$$Z_s(\rho) = c\,(X^2+Y^2) \Big/ \left\{1+\left[1-(1+Q)\, c^2(X^2+Y^2)\right]^{\frac{1}{2}}\right\}$$

$$p = \frac{X^2 + Y^2}{\rho_m^2}$$

ρ_m = radius of clear aperture or any other convenient normalising aperture radius; A_4, A_6, ... are figuring coefficients.

Rays are all traced relative to the optical axis. The starting data required are the co-ordinates $(\underline{X},\underline{Y},\underline{Z})$ of a point $\underline{P}$ or $\underline{P}_1$ (see Fig. 4-8) on the ray, together with its direction cosines (L_1, M_1, N_1) as described earlier.

Two ray tracing flow diagrams will be given later: one for pupil exploration (see next section), the other

for wave-aberration and aberration derivative calculations (see Chapter 7). It is convenient here to point out the basic differences between these two ray tracing flow diagrams, resulting from different requirements of the two cases:

(i) The transfer equations (say from surface 1 to 2).

In the pupil exploration case, a direct transfer from the surface point P_1 (see Fig. 4-2) to the next surface point P_2 is made, with the distance $P_1P_2 = D$. However, in the second case, the transfer equations (4-9) are to be preferred, and the distance D to be calculated represents $\underline{P}_2P_2$ $(= e_2)$.

(ii) Image space quantities.

For the pupil exploration case, only $(X,Y,Z)_k$ and $(L',M',N')_k$ are needed. However, in the ray trace flow diagram in Chapter 7, all the quantities described in equations (4-28) to (4-42), together with the special data required for aberration-derivative calculations, need to be computed and stored.

(IV-iii) The pupil exploration.

As a first step in the design process, the effective pupil ray of an extra axial pencil must be found and a number of rim rays are required to be determined. The values of the major and minor semi-axes, a_s and b_s, are then calculated for the best fitting ellipse given by the equation (2-27). These are the main functions of the pupil exploration. The actual procedures involved in a pupil exploration computer programme may be conveniently divided into four parts:

(a) Input stage - concerned with the system data and

the starting data (h_1, u_1), $(\bar{h}_1, \bar{u}_1)$. In addition, the values of the maximum semi-diameter of each surface must also be made available.

(b) The locating of the first ray which would pass through the system, followed by finding the limiting rim ray using an iterative method. This is carried out for different azimuths.

(c) The selection of the effective pupil ray, and the finding of the best-fitting ellipse by using the least-squares method.

(d) Out-put gives the values of a_s, b_s, the co-ordinates of the limiting rim rays and the effective centre of the vignetted pupil.

In the ray tracing procedure employed here, the aberrations of the system and the exit pupil co-ordinates are not calculated. One only requires to know which are the limiting rays that just pass through the system; and, if not, the reason for failure and the corresponding surface number are required to be printed out.

In addition to the usual system data, pupil exploration requires also the special input information listed below: these are accompanied by notes explaining the roles they play in the exploration.

1) ρ max (K) : this is the maximum clear semi-diameter of the K^{TH} surface. It is required for the following purposes:-

(a) for testing the ray height. If the ray co-ordinates at surface K is (x,y) then

$$\rho(K) = +\left[x^2 + y^2\right]^{\frac{1}{2}}$$

and

$$\rho \text{ max } (K) \rangle \rho(K)$$

is the condition for the ray to pass through surface K. When the ray falls outside the aperture radius i.e. $\rho \max \langle \rho$, a code number F is alloted the value "1" (see later).

(b) for calculating the increment of the ray height δr in (4-48).

2) $\underline{\Delta r}$: this is the ray radius increment used to establish the first ray to pass through the system. In this process the gaussian entrance pupil co-ordinates (x,y) are used to determine the rays to be traced. The first ray to be considered is the one which passes through the centre of the gaussian pupil, with x=y=o. If this ray fails the ray x=o, y=Δ r is next traced. A suitable value for Δr is 0.30. If this second ray fails, the co-ordinates to be tried next follow the sequence:

$$x = o,\ y = -\Delta r$$
$$x = o,\ y = 2\,\Delta r$$
$$x = o,\ y = -\,2\,\Delta r$$
$$x = o,\ y = 3\,\Delta r$$

etc. until a ray does pass (see Block I of the flow diagram). In general this ray can be located without any difficulty. In a rare case, failures may occur for all rays tried up to $|y| = 2.0$, and the calculation must then be stopped. The reasons for each failure and the surface numbers would be printed out. In

practice, this means that for $\Delta r = 0.3$, the maximum number of rays to be tried is J, given by

$$\left| \frac{J}{2} \Delta r \right| \leqslant 2 \qquad (4\text{-}46)$$

i.e. $J \leqslant 13$

The next step to follow after locating the first ray at (o, r_o) is to establish a crude estimate of the upper and lower periphery of the pupil. The upper limit is approached by using the pupil co-ordinates (o,y) with $y = r_o + n \Delta r$; rays are then traced in succession with $n = 1,2....,N$, where N corresponds to the ray that fails to pass. The required pupil co-ordinates are then (o,r_1) where

$$r_1 = r_o + (N - 1) \Delta r$$

and the corresponding ray heights at all surfaces $\rho(K)$ are stored as $\rho_1(K)$'s. For the lower limit of the periphery, rays are again traced in succession but now with $n = -1, -2, -3, ..., -N'$, leading to the approximate lower periphery point (o, r_R) where

$$r_R = r_o - (N' - 1)\Delta r$$

and its corresponding $\rho_R(K)$'s stored. The subscript "R" in r_R denotes the R^{th} point on the pupil periphery, corresponding to the azimuth $\varphi = \pi$. The minimum value of R is 3. The values of $r_2, r_3, ..., r_{R-1}$ are to be determined later.

The first estimates of the centre of the pupil c and its semi-diameter b in the meridian section are then

$$c = \frac{r_1 + r_R}{2}$$

$$b = \frac{r_1 - r_R}{2} \qquad (4\text{-}47)$$

and these are shown in Fig. 4-11.

3) Tolerance ξ for calculating the pupil radius.

This defines the accuracy with which the quantities such as r_1 and r_R must be determined. The magnitude of ξ may be taken as 0.01. This order of desired accuracy can be achieved by iterative ray tracing. If r_1 is now taken to be measured from the approximate centre c, then $r_1 = b$ is our starting point for locating the upper periphery of the pupil in the azimuth $\varphi = o$. Since this ray has already been traced with $\rho_1(K)$'s stored as mentioned above, we need now to know the increment for r_1 in order to arrive at a more accurate value for b. First it is necessary to examine at each surface the value of $\delta r_1(K)$ given by

$$\delta r_1(K) = \left[\frac{\rho \text{ max } (K) - \rho_1(K)}{|h(K)|} \right] \qquad (4\text{-}48)$$

where h(K) is the paraxial marginal ray height at the K^{th} surface. There are two cases to be considered:
(a) If h(K) is that of a field lens then $\delta r_1(K)$ is not calculated. Instead, the quantity

$$\Delta \rho (K) = \rho_{max} (K) - \rho (K)$$

is found, and if it is negative then this $\Delta \rho (K)$ will be printed out, indicating that the field lens is of too small diameter. In order to detect the presence of a field lens, the test

$$\left| h(K) \right| > h \text{ min.}$$

is applied to all surfaces. h min. has a small magnitude (say o.1).

(b) The value of (δr_1) min. : if this is the smallest magnitude of $\delta r_1(K)$ for all surfaces, then the following tests are applied:

(i) if $\left| (\delta r_1)_{\text{min.}} \right| \leqslant \xi$, then no further improvement for r_1 is necessary. In this case, we proceed to apply the next test

$$\left| \delta r_1(K) \right| < \xi$$

for all surfaces (excluding the stop and the field lens, if any). The surfaces which satisfy $\left| \delta r_1(K) \right| < \xi$ will have their surface number printed out, indicating that they are also the "limiting" surfaces which define the pupil periphery.

(ii) If $\left| (\delta r_1)_{\text{min.}} \right| > \xi$, then further improvement for r_1 is necessary: a ray is to be traced with

$$r_1 = r_1 + (\delta r_1)_{\text{min.}} \quad \left[\text{computer notation} \right]$$

This procedure is iterated until the limiting ray satisfying $\left| (\delta r_1)_{\text{min}} \right| \leqslant \xi$ is established, and the corresponding r_1 gives the position of the point B_1 on the pupil periphery. In Fig. 4-12, B_1 is shown as the point withray number i = 1.

The value of r_R, corresponding to i = R and azimuth $\varphi = \pi$, is the next to be determined. The method employed here is similar to that described above, and details are presented in Block II of the flow diagram (see later). The improved values of c and b are now given by

$$c = c + \frac{r_1 - r_R}{2}$$

$$b = \frac{r_1 + r_R}{2} \qquad (4\text{-}49)$$

and these are used for locating the other periphery points B_i with i = 2,3,.....,(R-1), corresponding to a set of azimuth angles φ_i given by

$$\varphi_i = (i-1)\,\Delta\varphi$$

$$\Delta\varphi = \frac{\pi}{R-1} \qquad (4\text{-}50)$$

For $1 < i \leq (R-1)$, the starting value of r_i to be used for each azimuth is taken to be the limiting ray value r_{i-1} found for the previous azimuth. (See (iii) of Block II flow diagram and Fig. 4-12).

If the first ray fails to pass with this r_i value, then it must be reduced by an amount δr_o in succession until the ray does pass. This is described by (i) in Block II flow diagram. The value of δr_o is to be provided by the input data, and in general $\delta r_o = 0.1$ is suitable. The value of r_i in each case is improved by (δr) min. similar to that described above for r_1, leading to the periphery points B_i's sought. The pupil co-ordinates (x_i, y_i) of B_i are calculated from

$$x_i = r_i \sin \phi_i$$

$$y_i = c + r_i \cos \phi_i \qquad (4\text{-}51)$$

and may be plotted as shown in Fig. 4-12.

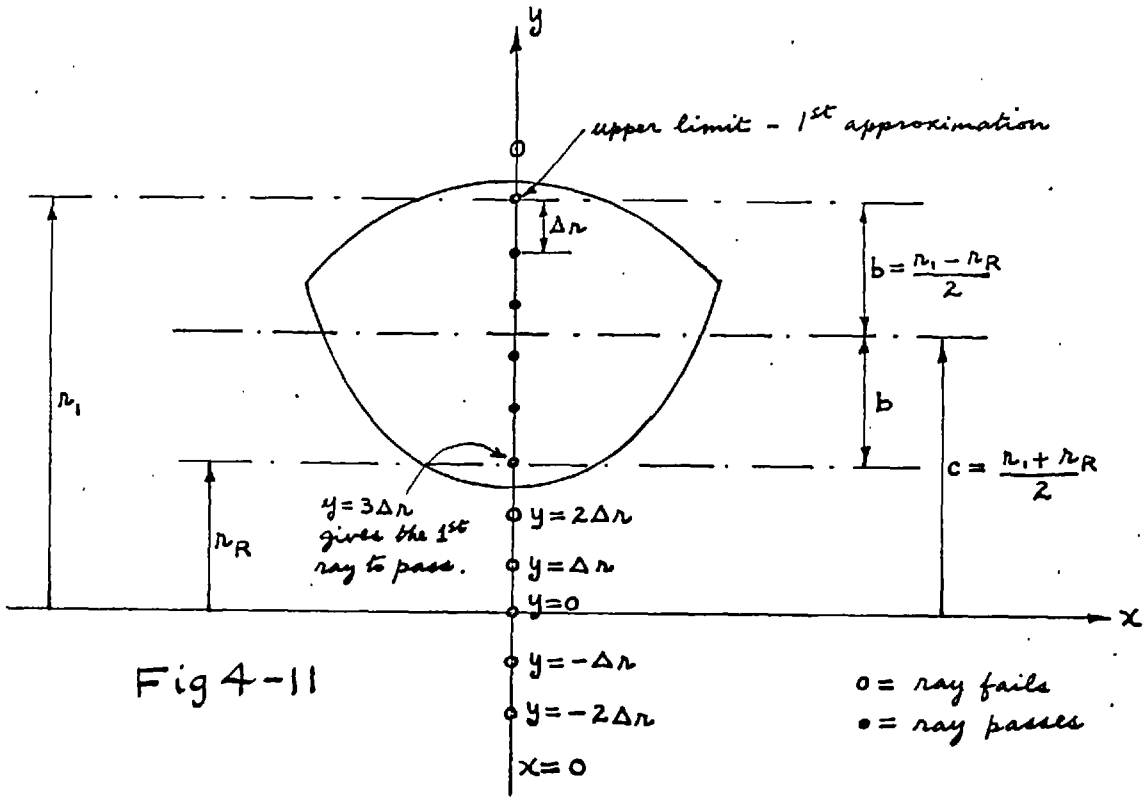


Fig 4-11

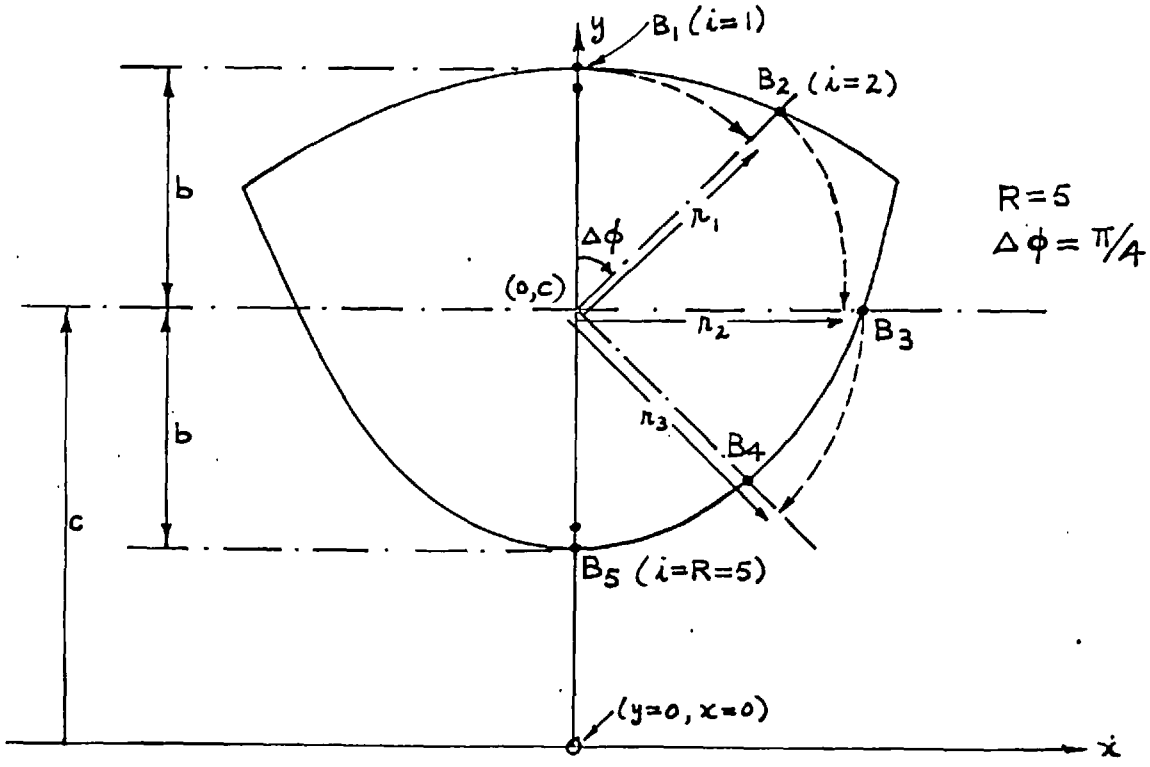


Fig 4-12 VIGNETTED PUPIL SHAPE - PLOT OF B_i's ($i = 5$)

Reasons for Failure.

The pupil exploration programme incorporates many logical tests which are used in the process of establishing the periphery points B_i's. These tests detect not only whether a ray will pass or not, but also analyse the cause of any failure. The flow diagram includes several special cases where the failure of a ray causes the calculation to be stopped. When this happens, the reason will be printed out in code, with F = 3,4 or 5 as the case may be. The significance of the coded values of F are as follows:

F = o : Ray pases.

F = 1 : $(\rho \text{ max} - \rho) < o$ i.e. ray height exceeds the clear aperture radius ρ max.

F = 3 : $\cos^2 I \leqslant o$ i.e. ray misses the surface.

F = 4 : $\cos^2 I' \leqslant o$ i.e. total internal reflection.

F = 5 : (when $\bar{u}_1 = o$) $N_1^2 \leqslant 0$ i.e. when the extra-axial ray in the object space misses the first surface.

F = 2 : this is a special test involved in loop (IV) in Block II flow diagram. Its purpose is to detect any ray for which $\rho < \rho$ max at each surface but which fails to pass because of F = 3 or 4. Thus for a given azimuth φ_i, the ray is allowed to pass loop (i) in Block II flow diagram if F = o or 1 : its value of r_i is imporved via.

$$r_i = r_i + (\delta r_i) \text{ min.} \quad (r_i \text{ is always + ve.})$$

(Fig. 4-13 shows the case under consideration.).

A test, $r_i < Z_o$ is then applied: at this point it is necessary for the ray to pass direct to the ray trace, and this requires $r_i < Z_o$. To facilitate this,

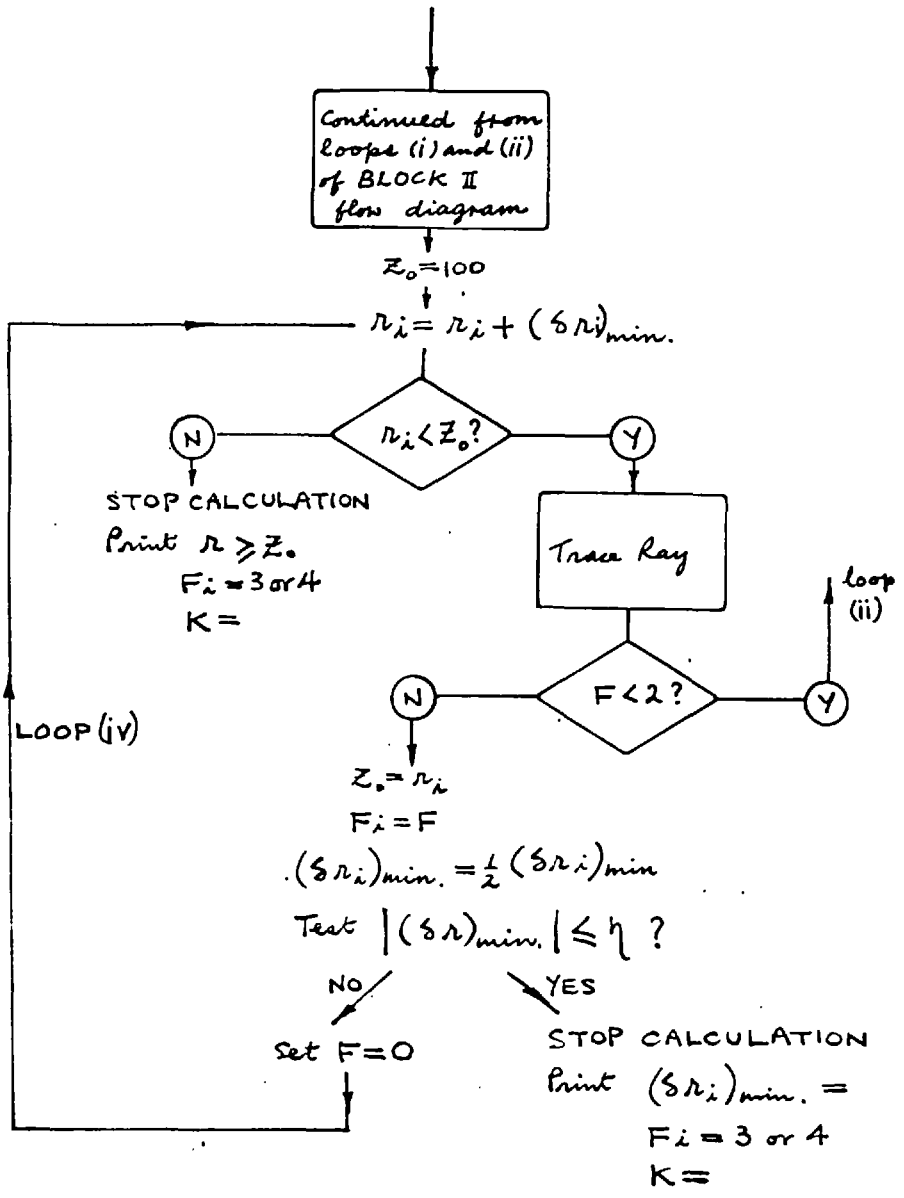


Fig. 4-13 LOOP (iv) OF BLOCK II FLOW DIAGRAM (FOR REFERENCE ONLY)

<u>Note</u> above illustrates the notes on the $F \gtrless 2$ case; loop (iv) will be presented again later in the Block II flow diagram.

Z_o is set to a large value, say 100. Following the ray trace, the value of F is tested : for $F > 2$, Z_o is set to the value of r_i whilst (δr_i) min. is halved, and the quantity $r_i = r_i + (\delta r_i)$ min. is recalculated. If (δr_i) min. is a + ve. quantity, then this means that $\rho < \rho$ max for all surfaces (see equation 4-48) but the ray fails to pass because of $F > 2$, indicating that the limiting factor here is not the lens diameters. This important conclusion is printed out with the appropriate code number for F if the test $r_i \geqslant Z$ is found to be true, and the calculation is stopped. When $r_i < Z_o$, i.e. (δr_i) min. is - ve., the ray is retraced with this new r_i value and the new verdict for F is again tested for $F > 2$, as indicated in Fig. 4-13. Note that the successive reduction of δr min. leads to a lower limit, denoted by η . When $|\delta r \text{ min.}| \leqslant \eta$, the calculation is stopped, and the reason of failure and the surface number, are both printed out.

The vignetted pupil - as approximated by a best-fitting ellipse (see Block III of the flow diagram).

In Fig. 4-12, the centre of the vignetted pupil is approximately given by the point (o,c), where c is the same quantity as that used in equation (4-51). If the vignetted pupil shape is to be approximated by an ellipse of the form

$$\frac{x^2}{a^2} + \frac{(y-c)^2}{b^2} = 1 \qquad (4\text{-}52)$$

then one may take the initial values for the major and minor semi-axes as being given by

$$a = r_{\frac{R+1}{2}} \quad ; \quad b = r_1$$

where $r_{\frac{R+1}{2}}$ is the polar radius of the periphery point $B_{\frac{R+1}{2}}$, in the azimuth $\varphi = \pi/2$, R being odd.

The pupil periphery co-ordinates (x_i, y_i) as given by (4-51) are stored at the end of Block II of the flow diagram for i = 1,2,...,R. If the following changes are now introduced

$$\begin{aligned} a_s &= a + \delta a \\ b_s &= b + \delta b \\ c_s &= c + \delta c \end{aligned} \qquad (4\text{-}53)$$

then to a first approximation

$$\left\{ \frac{2}{a}\left(\frac{x_i}{a}\right)^2 \right\} \delta a + \left\{ \frac{2}{b}\left(\frac{y_i - c}{b}\right)^2 \right\} \delta b + \left\{ \frac{2(y_i - c)}{b^2} \right\} \delta c + \left\{ 1 - \left[\left(\frac{x_i}{a}\right)^2 + \left(\frac{y_i - c}{b}\right)^2 \right] \right\} = 0 \qquad (4\text{-}54)$$

The values of (x_i, y_i) are now used to give R equations of this type, which are solved for (δa, δb, δc) by least-squares, and the best-fitting ellipse is then defined by (a_s, b_s, c_s) which are the corrected values of (a,b,c) in (4-53).

The least-squares method.

The above case may be represented by :

$$\begin{aligned} A_{11}\alpha_1 + A_{12}\alpha_2 + A_{13}\alpha_3 &= D_1 \\ A_{21}\alpha_1 + A_{22}\alpha_2 + A_{23}\alpha_3 &= D_2 \\ \cdots\cdots & \\ A_{R1}\alpha_1 + A_{R2}\alpha_2 + A_{R3}\alpha_3 &= D_R \end{aligned} \qquad (4\text{-}55)$$

The least-squares solutions for these R equations are:-

$$\sum_{n=1}^{3}\left\{\sum_{m=1}^{R} A_{mk}\,A_{mn}\right\}\alpha_n - \sum_{m=1}^{R} A_{mk}\,D_m = o \quad (4\text{-}56)$$

with k = 1,2,3. (4-56) is a special case of equation (1-24). It is convenient to represent (4-56) by:

$$k = 1 : B(1,1)\,\alpha_1 + B(1,2)\,\alpha_2 + B(1,3)\,\alpha_3 = H(1)$$
$$k = 2 : B(2,1)\,\alpha_1 + B(2,2)\,\alpha_2 + B(2,3)\,\alpha_3 = H(2) \quad (4\text{-}57)$$
$$k = 3 : B(3,1)\,\alpha_1 + B(3,2)\,\alpha_2 + B(3,3)\,\alpha_3 = H(3)$$

In the actual calculation, the coefficients A_{11}, A_{12} etc. are stored in a 2-D array i.e.

Aij = A(I,J) with I=1,R, J=1,3 { FORTRAN NOTATION meaning I=1,2,..,R

and D_m is stores as D(M), M=1,R. The typical coefficients B(k,J) and H(k) are computed from

Cycle k = 1,3

H(k) = o

Cycle J = 1,3

B(k,J) = o

Cycle I = 1,R

B(k,J) = B(k,J) + A(I,k) x A(I,J)

↓

J > 1?

Yes → (to Repeat I)

No → ↓

H(k) = H(k) + A(I,k) x D(I)

↓

Repeat I

Repeat J

Repeat k

Finally, the following determinants:

$$G=\begin{vmatrix} H(1) & B(1,2) & B(1,3) \\ H(2) & B(2,2) & B(2,3) \\ H(3) & B(3,2) & B(3,3) \end{vmatrix}$$

$$W=\begin{vmatrix} B(1,1) & H(1) & B(1,3) \\ B(2,1) & H(2) & B(2,3) \\ B(3,1) & H(3) & B(3,3) \end{vmatrix}$$

$$P=\begin{vmatrix} B(1,1) & B(1,2) & H(1) \\ B(2,1) & B(2,2) & H(2) \\ B(3,1) & B(3,2) & H(3) \end{vmatrix}$$

$$F=\begin{vmatrix} B(1,1) & B(1,2) & B(1,3) \\ B(2,1) & B(2,2) & B(2,3) \\ B(3,1) & B(3,2) & B(3,3) \end{vmatrix}$$

are calculated. Then

$$\delta a = \alpha_1 = G/F$$
$$\delta b = \alpha_2 = W/F$$
$$\delta c = \alpha_3 = P/F$$

From (4-53)

$$a_s = a + \delta a$$
$$b_s = b + \delta b$$
$$c_s = c + \delta c$$

Flow diagram.

Ray tracing for pupil exploration.

The values of (h_1, u_1), $(\bar{h}_1, \bar{u}_1)$ will be given.

Only the paraxial marginal ray need to be traced, using the refraction equation

$$n'u' - nu = h(n'-n).c$$

and the transfer equation

$$h' = h - d'u'$$

For finite rays, the following data

$$\bar{N}_o = +\frac{1}{[1 + (\bar{u}_1\tau)^2]^{\frac{1}{2}}}; \quad \bar{M}_o = -\bar{N}_o(\bar{u}_1\tau)$$

$$\bar{Y}_o = \bar{h}_1\tau \qquad (u_s)_o = \bar{N}_o u_1 \qquad (h_s)_o = h_1$$

are calculated and stored. τ = fractional object size $= \eta / \eta$ max. The pupil ray data $(\bar{h}_1, \bar{u}_1)$ are for $\tau = 1$.

A special counter f (=o or 1) is used in the pupil exploration: its value will have been set by the main flow diagram before the ray trace is called. For f = o, the ray fails for F = 1,3, or 4; for f = 1 the ray passes with F = o or 1, and fails for F = 3 or 4.

If the surface is not spherical, then the factor Q in equation (4-43) is not zero. A counter A is used to indicate whether the surface is figured ~~not~~ or not: for the former case A = 1, for the latter case A = o.

The values of c, φ and r will be specified prior to entering the ray trace routine. The following data are then calculated:

$$X = r \sin\phi$$
$$y = c + r\cos\phi$$
$$L_1 = -(\mu_s)_0\, x$$
$$M_1 = \bar{M}_0 - (\mu_s)_0\, y$$
$$N_1^2 = 1 - (L_1^2 + M_1^2)$$

$N_1^2 > 0$?

(N) Print:
$N_1^2 \leq 0$
$F = 5$
STOP CALCULATION

(Y)
$$N_1 = +\sqrt{1-(L_1^2+M_1^2)}$$
$$\underline{X}_1 = h_1\, x$$
$$\underline{Y}_1 = \bar{Y}_0 + h_1\, y$$
$$\underline{Z}_1 = -\left[\frac{L_1 x + (M_1+\bar{M}_0)\, y}{(N_1+\bar{N}_0)}\right] h_1$$

Starting point of the basic ray trace cycle

Cycle K $(K = 1, 2, 3, \cdots, NOSURF+1)$

NOSURF = total no. of glass surfaces.

All quantities below are associated with K, unless specified otherwise

$$G_1 = C_K\left[\underline{X}^2 + \underline{Y}^2 + (1+Q)\,\underline{Z}^2\right] - 2Z$$
$$G_2 = N - C_K\left[L\underline{X} + M\underline{Y} + (1+Q)N\underline{Z}\right]$$
$$J_1 = G_2^2 - (1+QN^2)\, C_K G_1$$

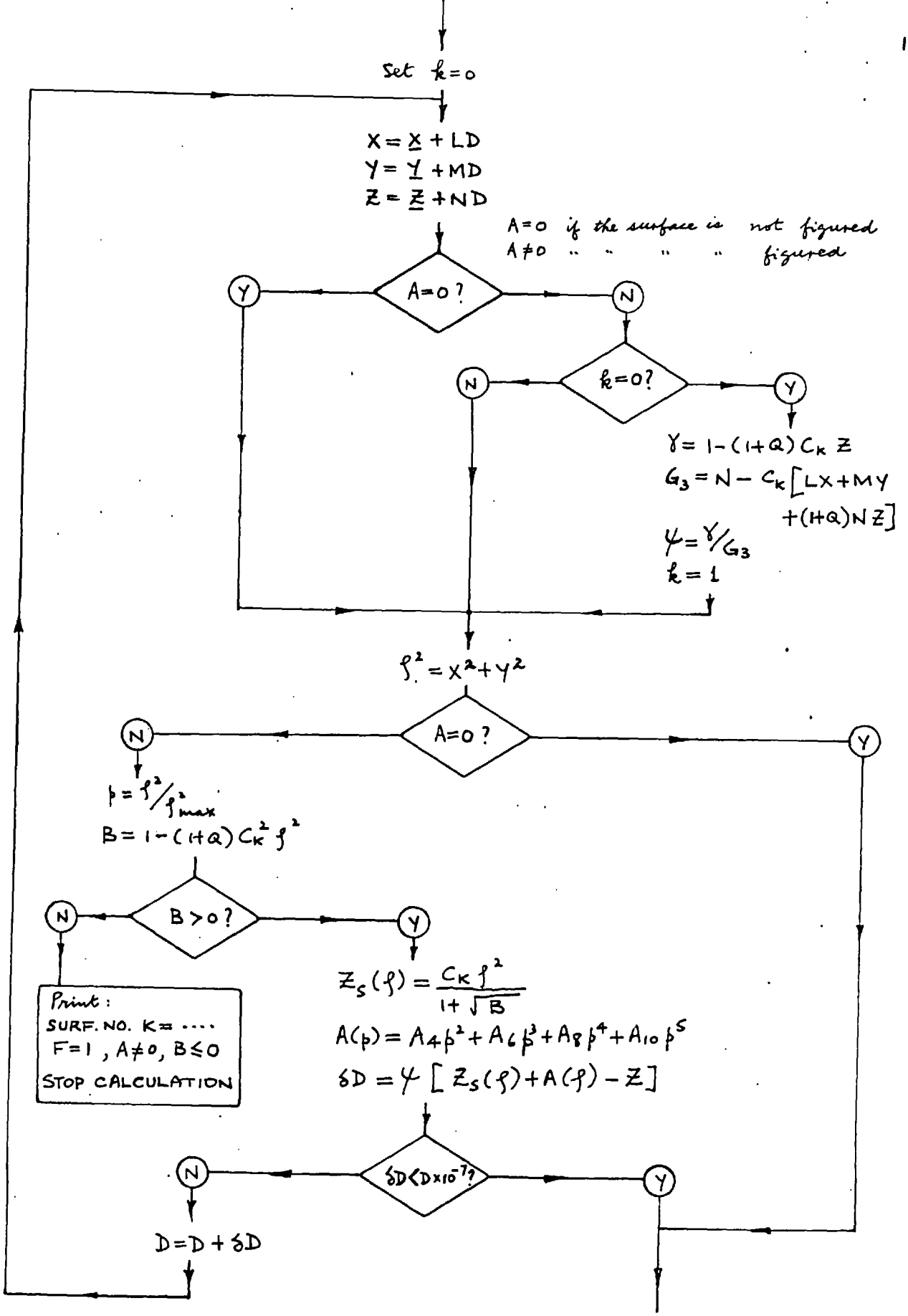
$J_1 > 0$?

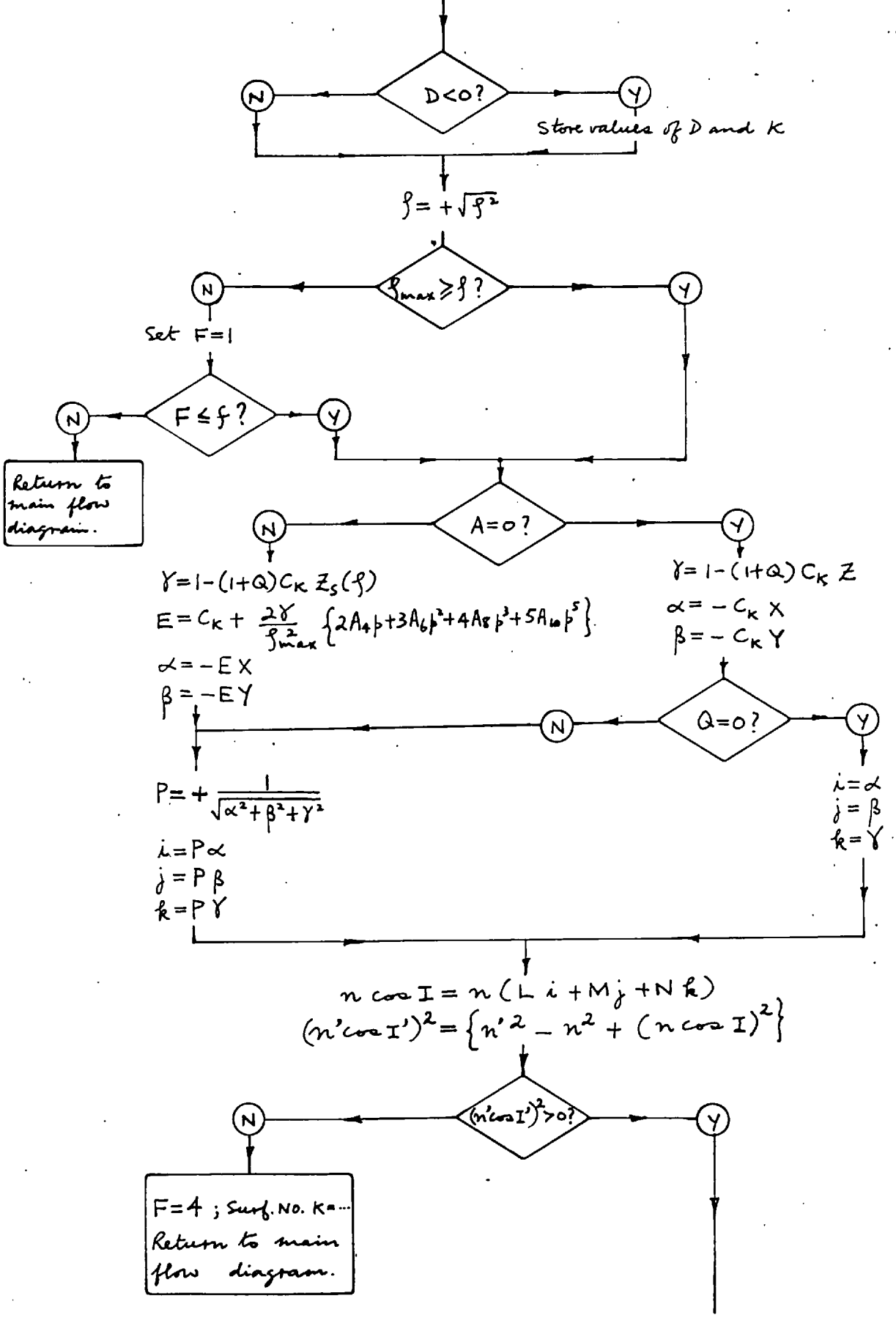
(N) $F = 3$

Return to main flow diagram (see later)

(Y)
$$D = \frac{G_1}{G_2 + \sqrt{J_1}}$$

P.T.O.


Set $k=0$
$X = \underline{X} + LD$
$Y = \underline{Y} + MD$
$Z = \underline{Z} + ND$
A=0 if the surface is not figured
A≠0 " " " " figured
A=0 ?
Y
N
k=0?
N
Y
$\gamma = 1-(1+Q)C_k Z$
$G_3 = N - C_k[LX+MY+(1+Q)NZ]$
$\psi = \gamma/G_3$
$k = 1$
$\rho^2 = X^2+Y^2$
A=0 ?
N
Y
$p = \rho^2/\rho^2_{max}$
$B = 1-(1+Q)C_k^2\rho^2$
B>0 ?
N
Y
Print:
SURF. NO. K= ....
F=1, A≠0, B≤0
STOP CALCULATION
$Z_S(\rho) = \frac{C_k\rho^2}{1+\sqrt{B}}$
$A(p) = A_4p^2 + A_6p^3 + A_8p^4 + A_{10}p^5$
$\delta D = \psi[Z_S(\rho)+A(\rho)-Z]$
$\delta D < D\times10^{-7}$?
N
Y
$D = D + \delta D$


D<0?
N
Y
Store values of D and K
$\rho = +\sqrt{\rho^2}$
$\rho_{max} \geq \rho$?
Set F=1
F ≤ f ?
Return to main flow diagram.
A=0?
$\gamma = 1-(1+Q)C_K Z_S(\rho)$
$E = C_K + \frac{2\gamma}{\rho^2_{max}}\{2A_4\rho+3A_6\rho^2+4A_8\rho^3+5A_{10}\rho^5\}$
$\alpha = -EX$
$\beta = -EY$
$\gamma = 1-(1+Q)C_K Z$
$\alpha = -C_K X$
$\beta = -C_K Y$
Q=0?
$P = +\frac{1}{\sqrt{\alpha^2+\beta^2+\gamma^2}}$
$i = P\alpha$
$j = P\beta$
$k = P\gamma$
$i = \alpha$
$j = \beta$
$k = \gamma$
$n\cos I = n(Li + Mj + Nk)$
$(n'\cos I')^2 = \{n'^2 - n^2 + (n\cos I)^2\}$
$(n'\cos I')^2 > 0$?
F=4 ; Surf. No. K=...
Return to main flow diagram.

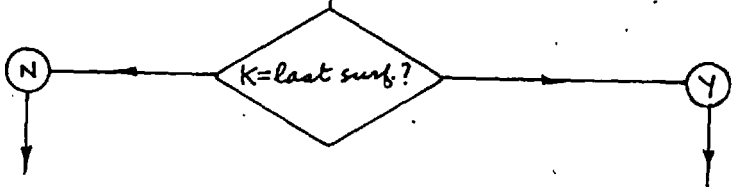
$$n'\cos I' = +\sqrt{(n'\cos I')^2}$$

$$\Delta(n\cos I) = (n'\cos I' - n\cos I)$$

$$L'_k = \left(\frac{n}{n'}\right)L + \frac{\Delta(n\cos I)}{n'} \cdot i$$

$$M'_k = \left(\frac{n}{n'}\right)M + \frac{\Delta(n\cos I)}{n'} \cdot j$$

$$N'_k = \left(\frac{n}{n'}\right)N + \frac{\Delta(n\cos I)}{n'} \cdot k$$



Transfer equations:

$$L_{K+1} = L'_K$$

$$M_{K+1} = M'_K$$

$$N_{K+1} = N'_K$$

$$\underline{X}_{K+1} = X$$

$$\underline{Y}_{K+1} = Y$$

$$\underline{Z}_{K+1} = Z - d_{K+1}$$

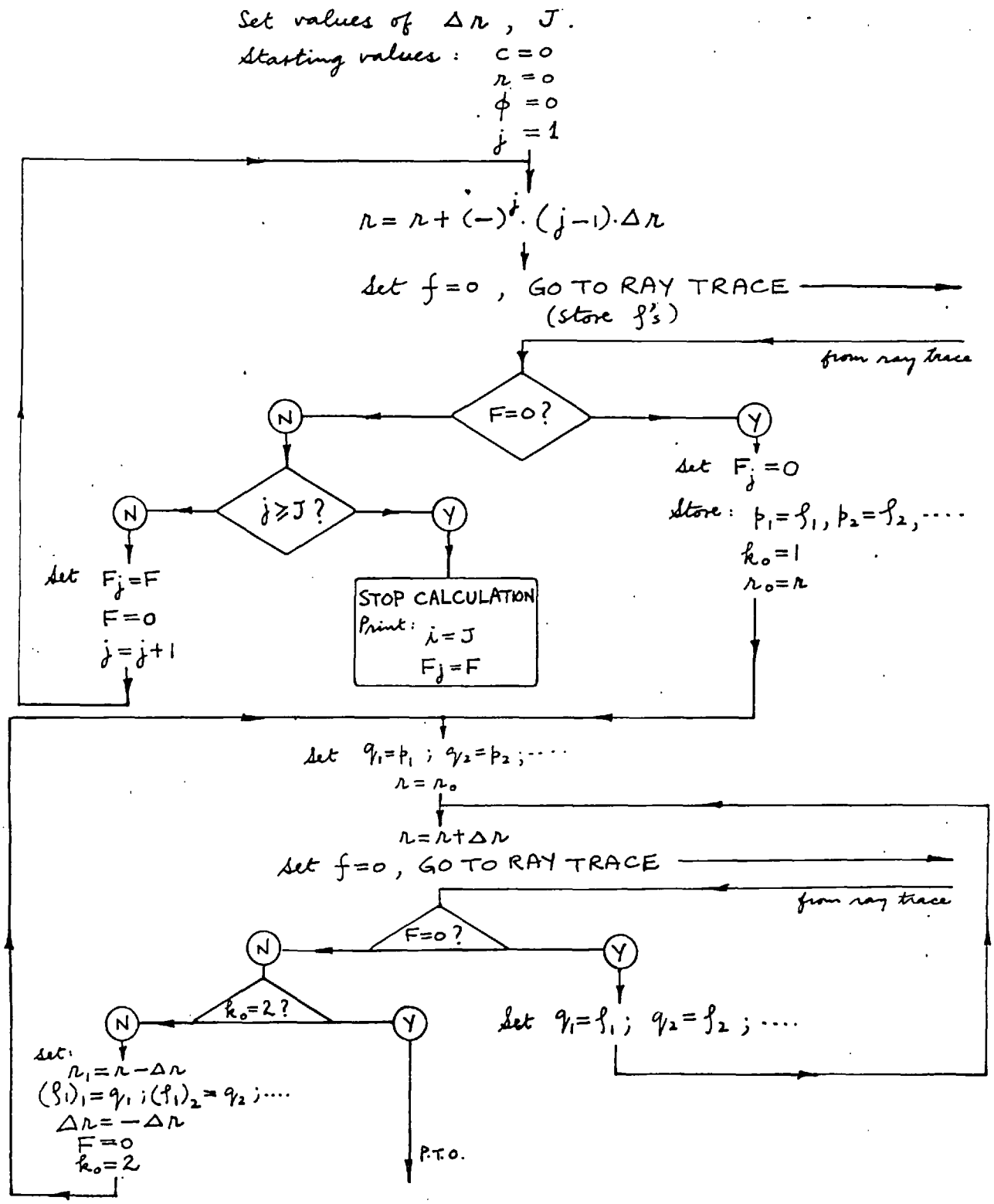
REPEAT K (Return to the starting point of the ray trace cycle: ready for the next K value, i.e. the next surface)

Return to main flow diagram.

PUPIL EXPLORATION

– Main flow diagram

BLOCK I



$r_R = r - \Delta r$

$(\rho_R)_1 = \rho_1 \; ; \; (\rho_R)_2 = \rho_2 \; ; \cdots$

$F = 0$

$c = \frac{1}{2}(r_1 + r_R)$

$b = \frac{1}{2}(r_1 - r_R)$

Block II

PUPIL EXPLORATION

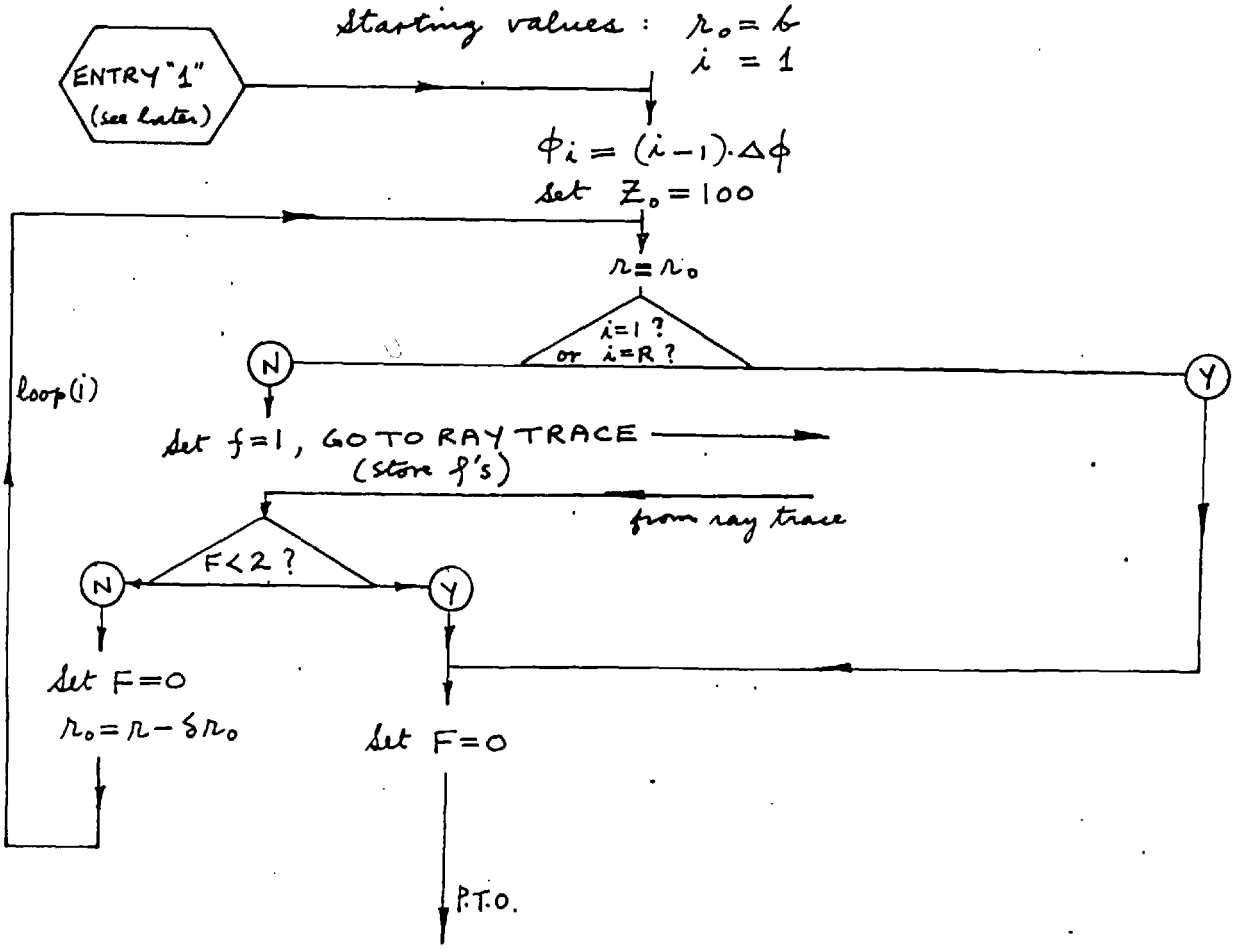
- Main flow diagram

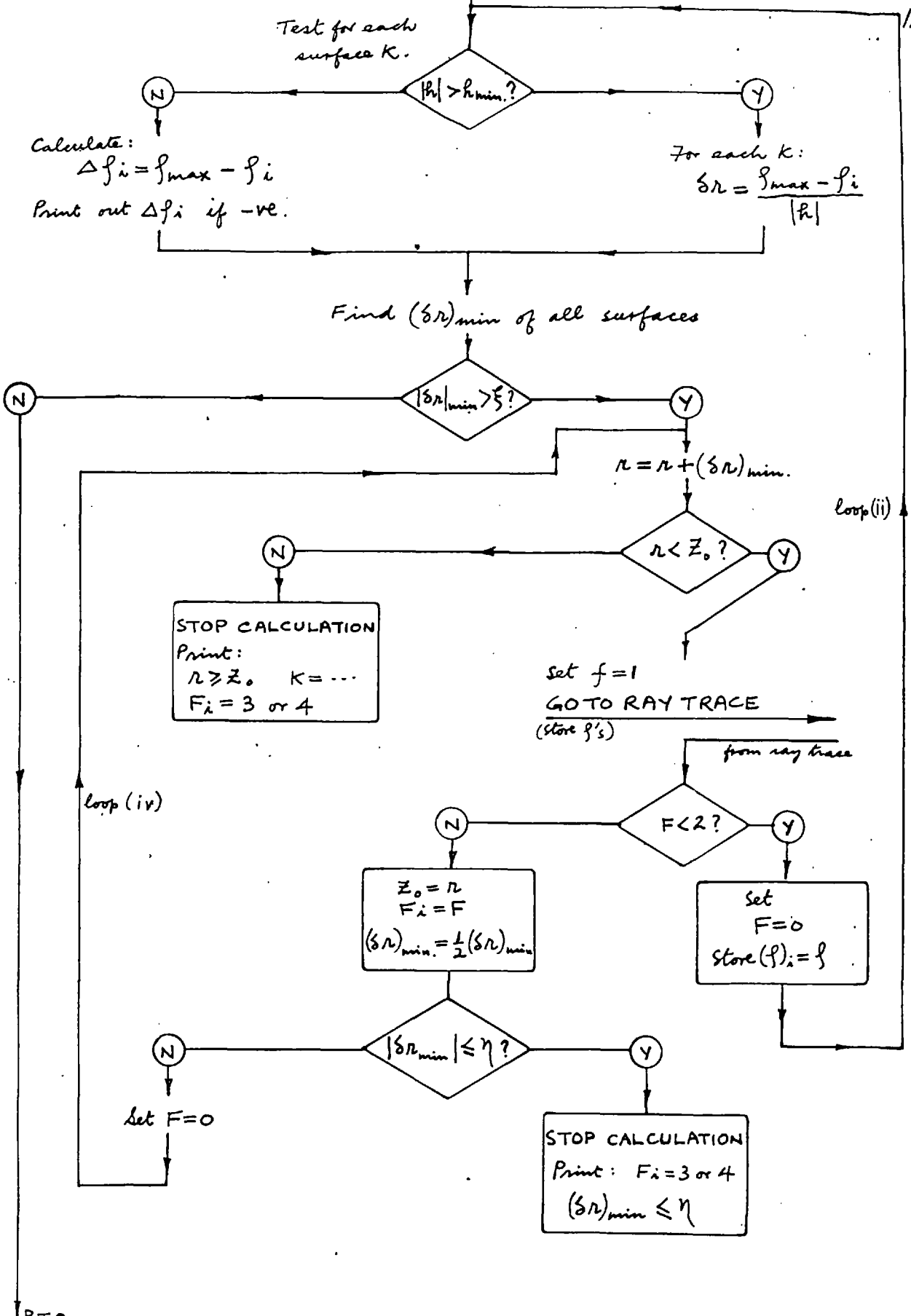
Block II

Set values of $\Delta\phi \; (= \frac{\pi}{R-1})$, $(h)_{min.}$, ξ, δr_0, η.

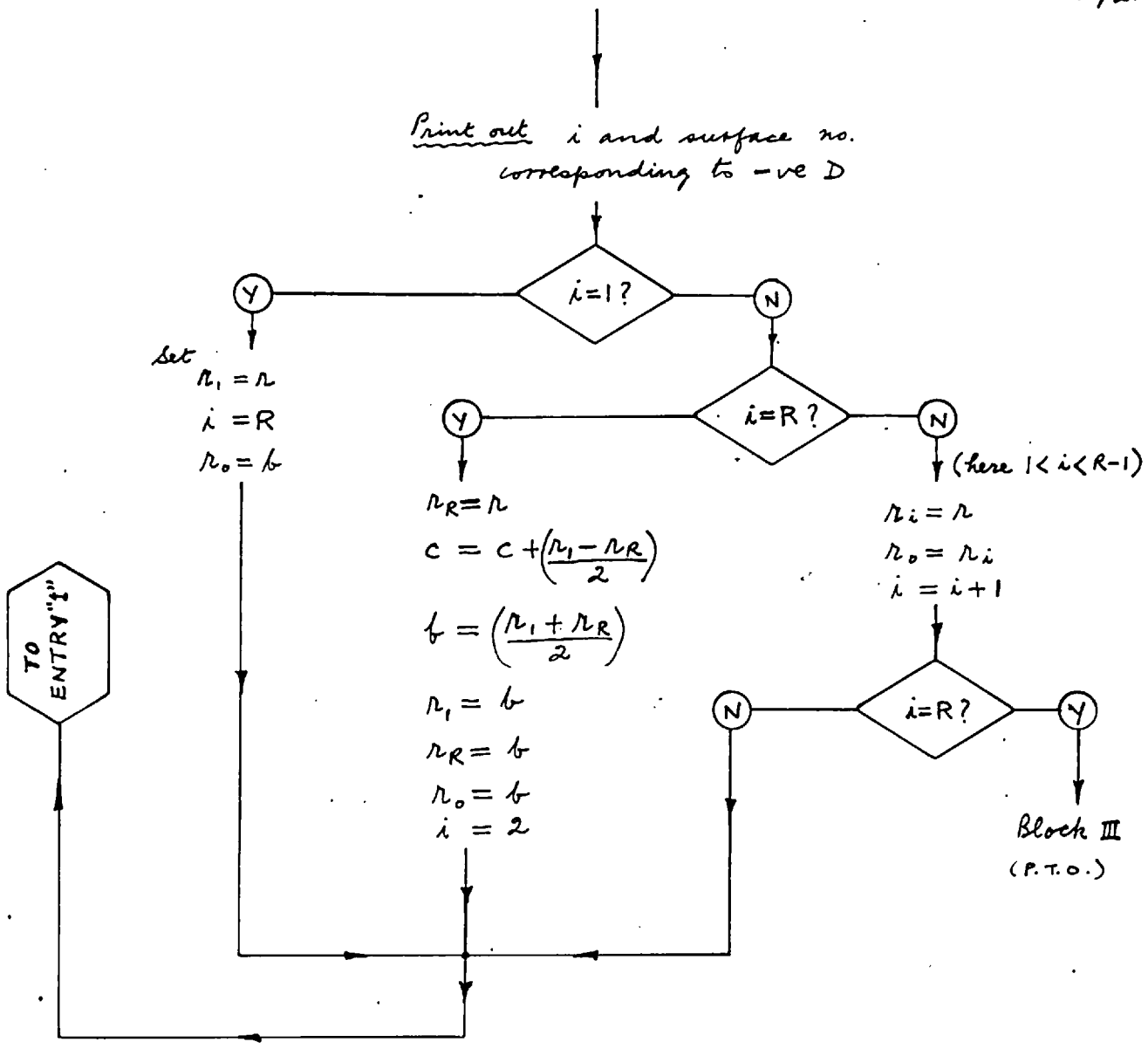
Starting values: $r_0 = b$

$i = 1$




Test for each surface K.
$|h| > h_{min.}$?
N
Y
Calculate:
$\Delta\rho_i = \rho_{max} - \rho_i$
Print out $\Delta\rho_i$ if -ve.
For each K:
$\delta r = \frac{\rho_{max} - \rho_i}{|h|}$
Find $(\delta r)_{min}$ of all surfaces
$|\delta r|_{min} > \xi$?
$r = r + (\delta r)_{min.}$
loop (ii)
$r < Z_o$?
STOP CALCULATION
Print:
$r \geq Z_o$ K = ...
F_i = 3 or 4
Set $f = 1$
GO TO RAY TRACE
(store ρ's)
from ray trace
loop (iv)
F<2?
$Z_o = r$
$F_i = F$
$(\delta r)_{min.} = \frac{1}{2}(\delta r)_{min}$
Set F=0
Store $(\rho)_i = \rho$
$|\delta r_{min}| \leq \eta$?
Set F=0
STOP CALCULATION
Print: F_i = 3 or 4
$(\delta r)_{min} \leq \eta$

P.T.O.


Print out i and surface no. corresponding to -ve D
i=1?
Y
N
Set
$r_1 = r$
$i = R$
$r_0 = b$
i=R?
Y
N
(here 1< i <R-1)
$r_R = r$
$c = c + \left(\frac{r_1 - r_R}{2}\right)$
$b = \left(\frac{r_1 + r_R}{2}\right)$
$r_1 = b$
$r_R = b$
$r_0 = b$
$i = 2$
$r_i = r$
$r_0 = r_i$
$i = i + 1$
i=R?
N
Y
Block III
(P.T.O.)
TO ENTRY "1"

PUPIL EXPLORATION

– Main flow diagram.

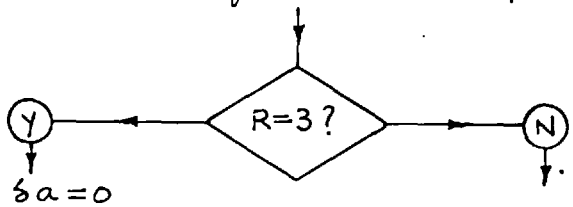
BLOCK III

$a = r_{(\frac{R+1}{2})}$

$b = r_1$

$c = c$

$x_i = r_i \sin\phi_i$
$y_i = c + r_i \cos\phi_i$
} Calculate and store for $i = 1, 2, \cdots, R$ $(R \geq 3)$.



$\delta a = 0$

$\delta b = 0$

$\delta c = 0$

Substitute (x_i, y_i), for $i = 1, 2, \cdots, R$, into equations (4-54).
Solve these R equations by least-squares for δa, δb and δc.
(see equations 4-55, 4-56, 4-57)

$a_s = a + \delta a$

$b_s = b + \delta b$

$c_s = c + \delta c$

Print out:

a_s, b_s, c_s

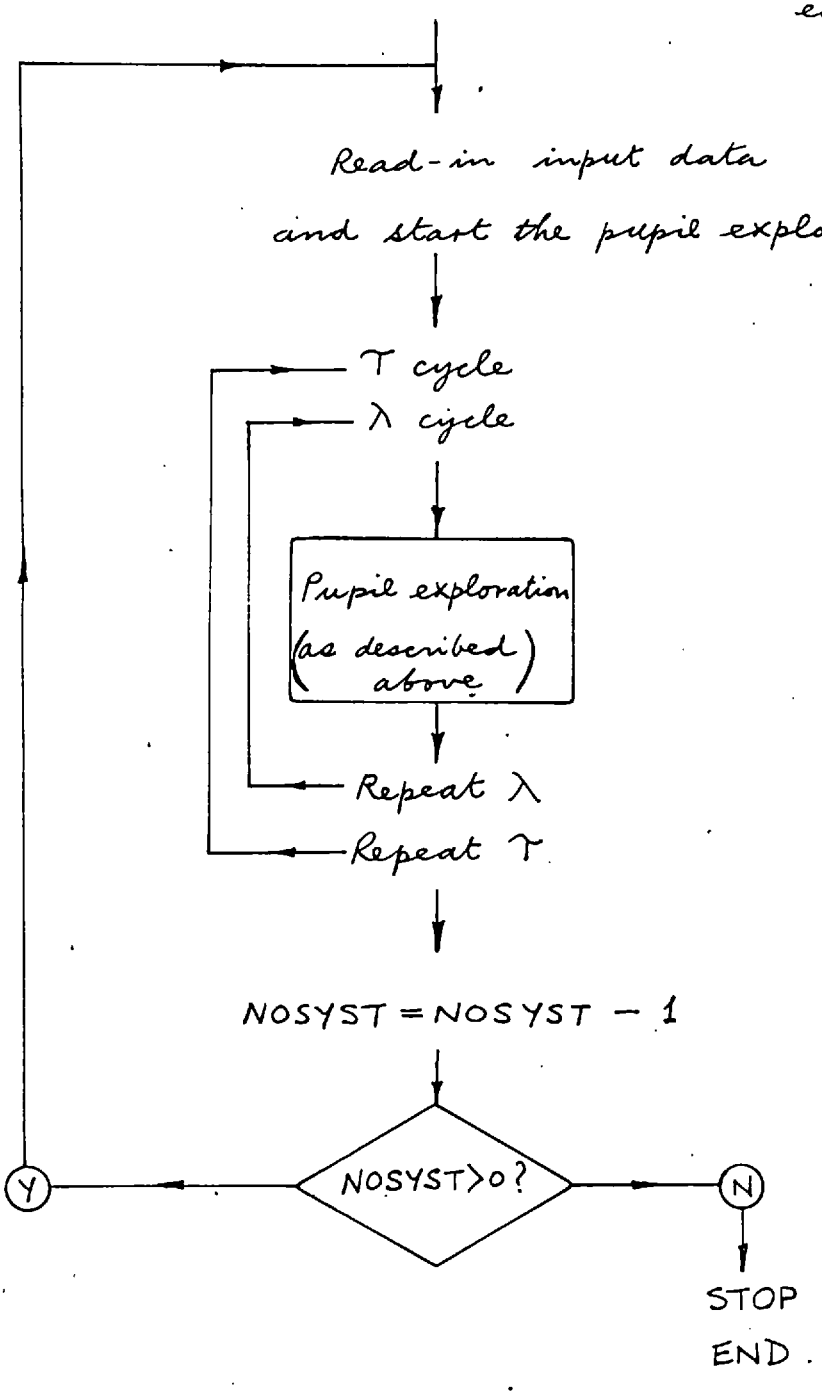
(x_i, y_i) ; $(x_i/a_s, y_i/b_s)$, $i = 1, 2, \cdots, R$.

Pupil Exploration - Discussion and Conclusions.

In order to have a complete check on the numerical results, two independent computer programmes for the pupil exploration have been written by the author. The flow diagram given above represents the first programme, which was written for the Elliott 803A computer using the Elliott Auto-code. The second pupil programme is more versatile (the modifications involved are given below) and was written for the IBM 7094 computer using the Fortran code. The results yielded by the two independent programmes are in complete agreement. The machine time required for the exploration depends obviously on the number of rim rays (R) required, as well as the complexity of the system: for the example given in Fig. 4-14, the machine time on IBM 7094 is of the order of a few seconds.

The modifications made in the IBM programme include the introduction of the counter "NOSYST", which is used when more than one system are to be explored, and cycles are introduced to enable pupil exploration to be carried out in succession for different fields (τ's) and different wavelengths when required. The modified flow diagram is now:

<u>At the start</u>: read-in the value of "NOSYST" (= the total no. of systems to be explored)



For each field, two sets of co-ordinates are printed out in the out-put. These are (x_i, y_i) and $\left(x_i/a_s\ ,\ y_i/b_s\right)$, i = 1,2,....,R. A plot of the co-ordinates (x_i, y_i) gives the vignetted pupil shape, and an example of this is shown in Fig. 4-14. (also P.127A) The scaled co-ordinates $\left(x_i/a_s,\ y_i/b_s\right)$ may also be plotted (Fig. 4-15) and this illustrates how accurately the pupil may be represented by the equation:

$$\left(\frac{x_i}{a_s}\right)^2 + \left(\frac{y_i - c_s}{b_s}\right)^2 = 1$$

For a microscope objective, or other systems of small field angle, three rim rays (R = 3) will usually suffice to give the extra axial pupil shape. For a wider field system, such as a photographic objective, at least five rim rays are usually needed to describe the pupil, and as many as fifteen rim rays may be needed to get a really accurate description of the pupil shape.

In practice, $(a_s,\ b_s,\ c_s)$ are calculated for all the τ values required. Thus for each τ value (each field), the effective pupil ray is chosen to pass through the centre of the vignetted pupil i.e. through (o, c_s) for that field. The meridian rays to be traced are then symmetrically distributed in this section even when there is heavily asymmetrical vignetting in the system. It will be described in Chapter 7 (under the "ASRAYS" flow diagram) how $(a_s,\ b_s,\ c_s)$ values, obtained from the pupil exploration, are used to calculate the ray co-ordinates (x,y)'s for the starting data in skew ray trace.

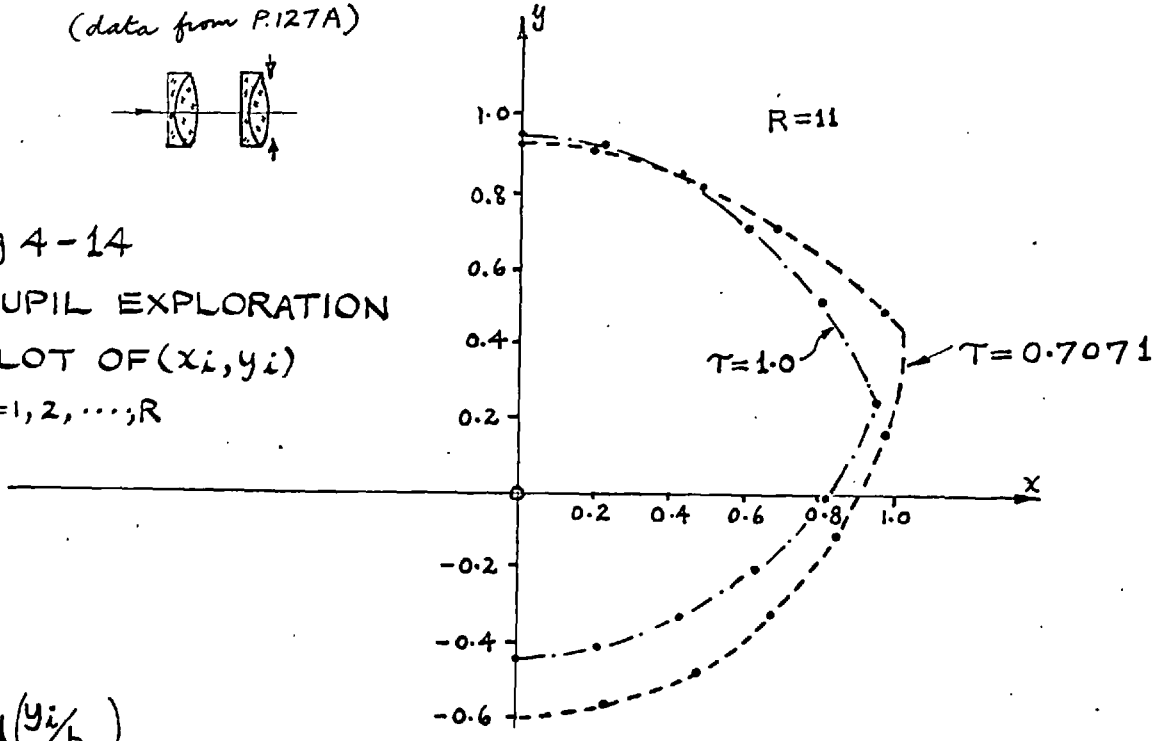
(data from P.127A)

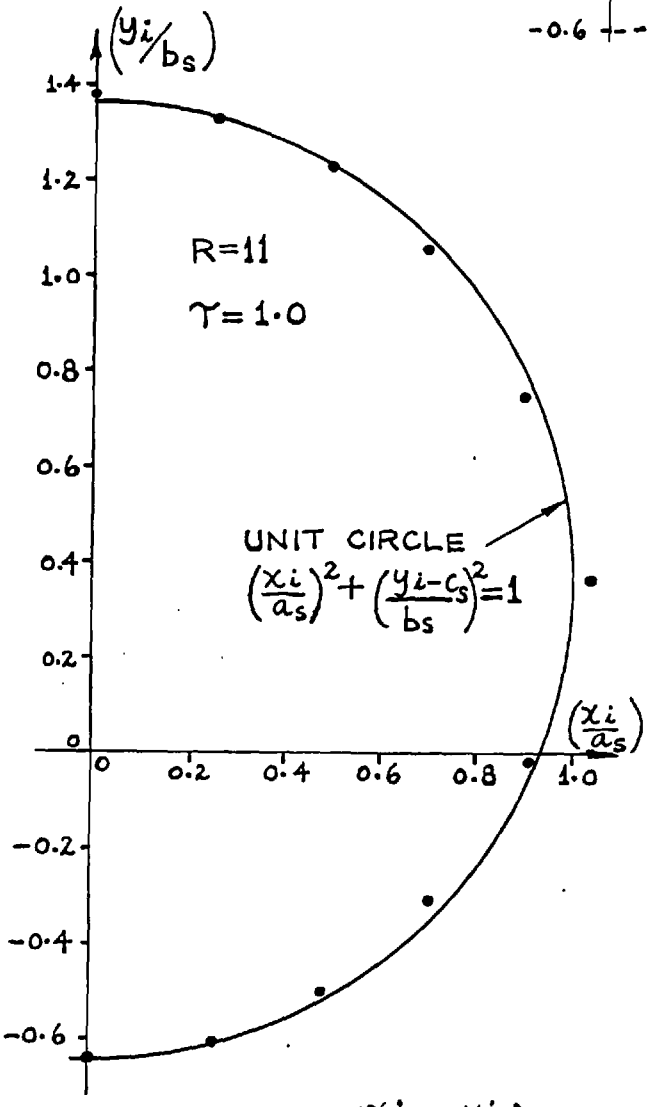
Fig 4-14

PUPIL EXPLORATION

PLOT OF (x_i, y_i)

$i = 1, 2, \cdots, R$





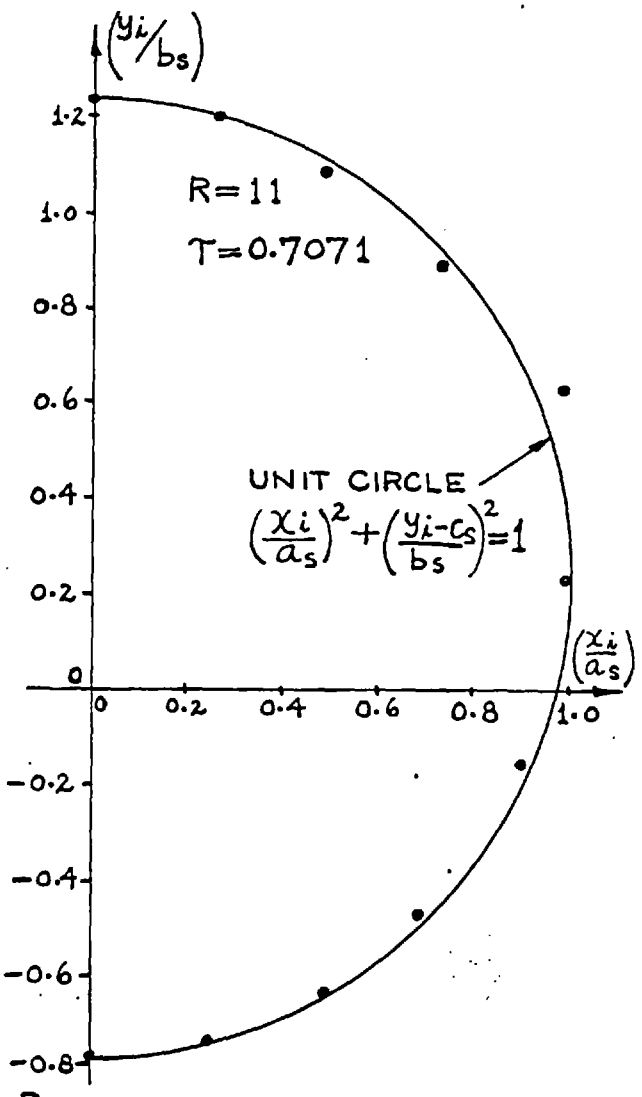


Fig 4-15 PLOT OF $\left(\frac{x_i}{a_s}, \frac{y_i}{b_s}\right)$ $i = 1, 2, \cdots, R$

PUPIL EXPLORATION

TOTAL NO. OF SURFACES = 6

K = 1 C = 0. MAX. SEMI-APERTURE = 1.786000 Q = 0.010 P =1.00

K = 2 C = 0.15932500 MAX. SEMI-APERTURE = 2.055000 Q = 0. P =0.

K = 3 C = -0.15747999 MAX. SEMI-APERTURE = 2.397000 Q = 0.020 P =1.00

K = 4 C = 0. MAX. SEMI-APERTURE = 4.151000 Q = 0. P =0.

K = 5 C = 0.08425599 MAX. SEMI-APERTURE = 4.329000 Q = 0.030 P =1.00

K = 6 C = -0.07792300 MAX. SEMI-APERTURE = 4.510000 Q = 0.040 P =1.00

FIGURING COEFS. AT K = 1
A_4 = 0.05000000
A_6 = 0.09000000
A_8 = 0.13000000
A_{10} = 0.16999999

FIGURING HT. MAX. = 2.309000

FIGURING COEFS. AT K = 3
0.15999999
0.09999999
0.14000000
0.18000000

FIGURING HT. MAX. = 2.800000

FIGURING COEFS. AT K = 5
0.07000000
0.11000000
0.15000000
0.19000000

FIGURING HT. MAX. = 4.800000

FIGURING COEFS. AT K = 6
0.08000000
0.12000000
0.16000000
0.20000000

FIGURING HT. MAX. = 5.300000
D = 1.2000000 K = 2
D = 2.4000000 K = 3
D = 14.0590000 K = 4
D = 1.6000000 K = 5
D = 3.3000000 K = 6

H(1) = 1.71416000 U(1) = -0.28000000 TOLERANCE FOR DELTA D (FIGURED SURF.)= 0.0000010 (SPHERE)

DELTA PHI (DEGREES) = 18.000000 MIN. +VE. HT. AT FIELD LENS = 0.100000 DELTA RC (FOR F .GT. 2) =0.09999999

TOLERANCE FOR LIMITING RAY =0.01000000 ETA (=0S) = 0.0400000 NO. OF RIM RAYS(=R) = 11 J = 13

NO. OF FIELDS = 3 NO. OF WAVELENGTHS = 1 UBAR(1) = 0. H(1)BAR =1.023305 DELTA R (1ST RAY) =0.300000

WAVELENGTH = 5461.00 A
E = 1.0000000 K = 1
E = 1.6799700 K = 2
E = 1.5210400 K = 3
E = 1.0000000 K = 4
E = 1.6299000 K = 5
E = 1.5210400 K = 6
E = 1.0000000 K = 7

THE FOLLOWING RAY DATA ARE FOR FIELD = 0. WAVELENGTH = 5461.00 A

MOD. DELTA R LESS THAN 0.010000 T =0. I = 1 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 11 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 2 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 3 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 4 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 5 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 6 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 7 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 8 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 9 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0. I = 10 AT SURFACE
K = 1

ELLIPTICAL PUPIL

FIELD = 0. WAVELENGTH = 5461.00 A R = 11 A = 0.985894 B = 0.988894 C = 0.0000000

RAY 1	X = 0.	Y = 0.985894	X/A = 0.	Y/B = 1.000000	
RAY 2	X = 0.304658	Y = 0.937641	X/A = 0.309017	Y/B = 0.951057	
RAY 3	X = 0.579494	Y = 0.797605	X/A = 0.587785	Y/B = 0.809017	
RAY 4	X = 0.797605	Y = 0.579494	X/A = 0.809017	Y/B = 0.587785	
RAY 5	X = 0.937641	Y = 0.304658	X/A = 0.951057	Y/B = 0.309017	
RAY 6	X = 0.985894	Y = 0.000000	X/A = 1.000000	Y/B = 0.000000	
RAY 7	X = 0.937641	Y =-0.304658	X/A = 0.951057	Y/B =-0.309017	
RAY 8	X = 0.797605	Y =-0.579494	X/A = 0.809017	Y/B =-0.587785	
RAY 9	X = 0.579494	Y =-0.797605	X/A = 0.587785	Y/B =-0.809017	
RAY 10	X = 0.304658	Y =-0.937641	X/A = 0.309017	Y/B =-0.951057	
RAY 11	X = 0.000000	Y =-0.985894	X/A = 0.000000	Y/B =-1.000000	

THE FOLLOWING RAY DATA ARE FOR FIELD = 0.70710700 WAVELENGTH = 5461.00 A

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 1 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 11 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 2 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 3 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 4 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 5 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 6 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 7 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 8 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 9 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =0.707107 I = 10 AT SURFACE
K = 1

ELLIPTICAL PUPIL

FIELD = 0.707107 WAVELENGTH = 5461.00 A R = 11 A = 0.970475 B = 0.754428 C = 0.1719984

RAY 1	X = 0.	Y = 0.932743	X/A = 0.	Y/B = 1.236357	
RAY 2	X = 0.241241	Y = 0.908836	X/A = 0.248581	Y/B = 1.204669	
RAY 3	X = 0.469153	Y = 0.812105	X/A = 0.483426	Y/B = 1.076451	
RAY 4	X = 0.665583	Y = 0.649946	X/A = 0.685832	Y/B = 0.861508	
RAY 5	X = 0.975288	Y = 0.483262	X/A = 1.004960	Y/B = 0.640567	
RAY 6	X = 0.965131	Y = 0.166372	X/A = 0.994494	Y/B = 0.220527	
RAY 7	X = 0.857400	Y =-0.112215	X/A = 0.883485	Y/B =-0.148741	
RAY 8	X = 0.677493	Y =-0.325836	X/A = 0.698105	Y/B =-0.431924	
RAY 9	X = 0.470491	Y =-0.481203	X/A = 0.484805	Y/B =-0.637838	
RAY 10	X = 0.239883	Y =-0.572220	X/A = 0.247284	Y/B =-0.758481	
RAY 11	X = 0.000000	Y =-0.600000	X/A = 0.000000	Y/B =-0.795304	

THE FOLLOWING RAY DATA ARE FOR FIELD = 1.00000000 WAVELENGTH = 5461.00 A

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 1 AT SURFACE
K = 6

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 11 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 2 AT SURFACE
K = 6

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 3 AT SURFACE
K = 6

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 4 AT SURFACE
K = 6

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 5 AT SURFACE
K = 6

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 6 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 7 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 8 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 9 AT SURFACE
K = 1

MOD. DELTA R LESS THAN 0.010000 T =1.000000 I = 10 AT SURFACE
K = 1

ELLIPTICAL PUPIL

FIELD = 1.000000 WAVELENGTH = 5461.00 A R = 11 A = 0.892866 B = 0.683230 C = 0.2341464

RAY 1	X = 0.	Y = 0.948407	X/A = 0.	Y/B = 1.388123	
RAY 2	X = 0.219058	Y = 0.927112	X/A = 0.245342	Y/B = 1.356955	
RAY 3	X = 0.431847	Y = 0.847308	X/A = 0.483864	Y/B = 1.240151	
RAY 4	X = 0.619843	Y = 0.703264	X/A = 0.694217	Y/B = 1.029323	
RAY 5	X = 0.810231	Y = 0.516102	X/A = 0.907451	Y/B = 0.755903	
RAY 6	X = 0.947965	Y = 0.252622	X/A = 1.061711	Y/B = 0.370186	
RAY 7	X = 0.815031	Y =-0.011698	X/A = 0.912826	Y/B =-0.017414	
RAY 8	X = 0.629099	Y =-0.204145	X/A = 0.704585	Y/B =-0.298794	
RAY 9	X = 0.428206	Y =-0.336453	X/A = 0.479586	Y/B =-0.492444	
RAY 10	X = 0.218736	Y =-0.411043	X/A = 0.241622	Y/B =-0.601618	
RAY 11	X = 0.000000	Y =-0.442563	X/A = 0.000000	Y/B =-0.647751	

As described in the pupil exploration flow diagram the intermediate output such as

(i) the surface number (s) which is responsible for the vignetted pupil shape

(ii) the values of i and surface number corresponding to - ve. D (see loop (ii), Block II)

(iii) -ve. $\Delta \rho_i$'s (if any) — from loop (ii), Block II.

(iv) the reason for failure of the ray (if any),

etc. are obviously useful information to the lens designer, and these may lead to a modified system at an early stage of the design.

REFERENCES

(1) H.H. Hopkins, Proc. Phys. Soc., B vol. LXV, P934 (1952).
(2) H.H. Hopkins, Optical Calculation I, Imperial College (1963).
(3) H.H. Hopkins, Nature, 159, P573 (1947).
(4) H.H. Hopkins, Summer School Notes, Imperial College (1965).

CHAPTER 5

The coefficients K(n,q;j) introduced in equation (2.32) in Chapter 2 depend on the specification of rays as well as the aberration polynomials used. The derivation of formulae for these, together with the P(i,j) coefficients in (2.33), will now be described.

(V-i) Specification of pupil co-ordinates in ray trace.

The rays to be traced are specified by giving the values of the two quantities A and S; A is the number of aperture rays traced in the azimuth $\phi=o$ of the pupil, and S denotes the number of skew rays to be traced (S being always an odd number). The total number of rays traced from each extra-axial object point will be

$$J = 2A+S \qquad (5.1)$$

of which there are A in the azimuth $\phi=o$, A in the azimuth $\phi=\pi$, and a single ray in each of the skew azimuths.

$$\phi_i = i\left(\frac{\pi}{s+1}\right) \quad , \; i=1,2,\ldots S \qquad (5.2)$$

The major and minor semi-axes (a_s and b_s) of the elliptical pupil are given, and the scaled pupil co-ordinates $(x_s,y_s)_j$ for ray tracing are found from

$$(x_s)_j = a_s x_j \quad ; \quad (y_s)_j = b_s y_j \qquad (5.3)$$

j denotes the j^{th} ray

where $(x,y)_j$ are given as for a circular pupil, i.e. with $a_s=b_s=1$.

In the azimuth $\phi=0$, the rays are numbered

$$j=1,2,\ldots A$$

and have co-ordinates

$$x_j=o, \quad y_j = \pm\sqrt{\frac{(A+1)-j}{A}} \qquad (5.4)$$

The forms of the wave-aberration polynomials (see later) argue against having a uniform distribution of rays over the whole pupil area. With the above convention the rays are equally spaced in the aperture radius squared, and j=1 is the marginal ray.

For the axial object point, S is set to zero and A aperture rays are traced in the azimuth $\phi=o$ of the pupil.

For an extra-axial object point, S is set to the specified value, and A aperture rays are traced, using (5.4). In addition, the rays $j=A+1,A+2,\ldots,(A+\frac{S+1}{2})$ are traced respectively in the azimuth ϕ_i (see 5.2), with $i=1, 2,\ldots \frac{S+1}{2}$. The polar radii, r_i, of these rays are :-

$$S=1, \quad i=1, \quad r_1=1$$

$$S=3, \quad i=1,2 \qquad \begin{cases} i=1, & r_1 = y_2 = \pm\sqrt{\frac{A-1}{A}} \\ i=2, & r_2 = 1 \end{cases} \qquad (5.5)$$

and the ray co-ordinates for these are :

$$\begin{aligned} x_i &= r_i \sin \phi_i \\ y_i &= r_i \cos \phi_i \end{aligned} \qquad (5.6)$$

The rays j=1,2,...,A,A+1,...N, with

$$N = A + \frac{S+1}{2} \tag{5.7}$$

are thus traced in the upper half of the pupil (see Fig.5.1), with the $j=N^{th}$ ray in the azimuth $\phi = \pi/2$. The rays in the lower half of the pupil with

$$j=N+1,\ N+2,\ \ldots,\ N+A+\frac{S-1}{2} \tag{5.8}$$

are traced with

$$\begin{aligned} x_j &= x_{j-N} \\ y_j &= -y_{j-N} \end{aligned} \tag{5.9}$$

so that for $j<N$, the rays j and j+N are symmetrical about the x axis, leading to simplified calculations of odd and even aberrations from the ray trace. With the above conventions, it is possible to specify the rays to be traced for any system by means of the two numbers A and S.

Note It will be seen in the following section that the presence of wave-aberration terms of the form

$$W_{31}r^3\cos\phi + W_{33}r^3\cos^3\phi$$

requires $S \geqslant 3$ to be chosen. This is because that the two types of aberration

$$W_{31}r_i^3\cos\phi_i \quad , \quad W_{33}r_i^3\cos^3\phi_i$$

can not be differentiated when rays are only traced in the $\phi=o$, $\phi=\pi/2$ and $\phi=\pi$ sections (corresponding to S=1).

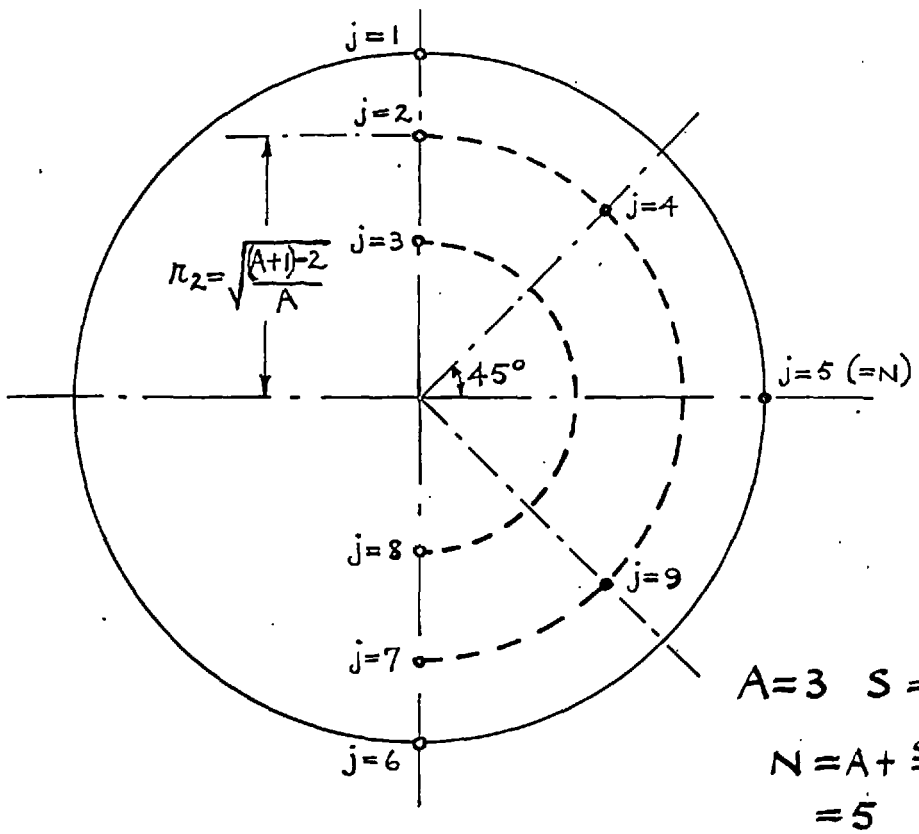


Fig 5-1 SPECIFICATION OF RAYS

CASE NO.	EVEN PART OF ABERRATION (U_j)	ODD PART OF ABERRATION (V_j)	AXIAL IMAGES $T=0$ (W_j)
(1) A=3, S=3	$q=0$; $n=2,4,6$ $q=2$; $n=2,4$	$q=1$; $n=1,3,5$ $q=3$; $n=3$	A=3, S=0 } $q=0$; $n=2,4,6$
(2) A=3, S=1	$q=0$; $n=2,4,6$ $q=2$; $n=2$	$q=1$; $n=1,3,5$	A=3, S=0 } $q=0$; $n=2,4,6$
(3) A=4, S=1	$q=0$; $n=2,4,6,8$ $q=2$; $n=2$	$q=1$; $n=1,3,5,7$	A=4, S=0 } $q=0$; $n=2,4,6,8$
(4) A=5, S=1	$q=0$; $n=2,4,6,8,10$ $q=2$; $n=2$	$q=1$; $n=1,3,5,7,9$	A=5, S=0 } $q=0$; $n=2,4,6,8,10$
(5) A=2, S=1	$q=0$; $n=2,4$ $q=2$; $n=2$	$q=1$; $n=1,3$	A=2, S=0 } $q=0$; $n=2,4$

TABLE 5-1

(V-ii) Wave-aberration polynomials.

The aberration polynomial is written as a function of the equivalent circular entrance pupil co-ordinates (x,y). There will be (A-1) orders of spherical aberration, (A-1) orders of circular coma, a transverse focal shift, a sagittal curvature coefficient and an astigmatism term. This gives a total of (2A+1) coefficients. The total number of rays traced (=2A+S, see (5-1)) depends on the coefficients actually required. In any practical case, the values of A and S are kept as small as is consistent with an accurate representation of the wave-aberration. Suitable values of A and S for five different types of optical systems are :

Case (1) Photographic objective: A=3, S=3.

Case (2) Microscope objective (low power): A=3,S=1.

Case (3) " " (medium power): A=4, S=1.

Case (4) " " (high power): A=5, S=1.

Case (5) Eye piece : A=2, S=1.

The value S=3 is used for case (1) due to its larger angular field coverage. In this case, the number of aberration coefficients required is (2A+3), two more than that which has been described above. These two coefficients may conveniently be chosen to be

$$W_{33}\ y^3\ ,\ W_{42}\ (x^2+y^2)\ y^2$$

Thus the aberration polynomial for A=3, S=3 (case (1))

will be

$$W(x,y)= W_{11}y+W_{20}(x^2+y^2)+W_{22}y^2+W_{31}(x^2+y^2)y + W_{33}y^3+W_{40}(x^2+y^2)^2+W_{42}(x^2+y^2)y^2 + W_{51}(x^2+y^2)^2y + W_{60}(x^2+y^2)^3 \qquad (5.10)$$

For the j^{th} ray, $W(x,y)$ may be written as W_j. As a consequence of using the ray co-ordinates defined by (5.9), the even and odd parts of $W(x,y)$ may be separated out. Thus for the extra-axial ($\tau \neq o$) case,

$$(W_j)_{even} = U_j = \frac{W_j+W_{j+N}}{2} \quad , \quad j=1,2,\ldots,N-1 \qquad (5.11)$$

$$U_j = W_j \quad , \quad j=N$$

From (5.10), U_j may be written as

$$U_j = W_{20}(x^2+y^2)_j+W_{40}(x^2+y^2)_j^2 +W_{60}(x^2+y^2)^3_j + W_{22}\, y_j^2 + W_{42}\,(x^2+y^2)y_j^2 \qquad (5.12)$$

and

$$(W_j)_{odd} = V_j = \frac{W_j - W_{j+N}}{2} \quad , \quad j = 1,2,\ldots,N-1 \qquad (5.13)$$

From (5.10), V_j may be written as

$$V_j = W_{11}y_j + W_{31}(x^2+y^2)_j y_j+W_{51}(x^2+y^2)_j^2\, y_j + W_{33}\, y_j^3 \qquad (5.14)$$

Thus for A=3, S=3, $N=A+\frac{S+1}{2}=5$, we have five equations for U, four equations for V and these can be solved

for the nine aberration coefficients. By substituting the values of (x_j, y_j) as specified by (5.4) and (5.5) into (5.12), (5.14) above, it can be shown that the co-ordinates of the ray pattern lead to well conditioned equations for the aberration coefficients.

For the axial object point ($\tau=0$) the aberration polynomial for A=3, S=0 reduces to

$$W_j = W_{20}\, y_j^2 + W_{40}\, y_j^4 + W_{60}\, y_j^6 \qquad (5.15)$$

with $J = 1, 2, 3(=A)$.

The aberration polynomials for the remaining four cases are chosen as follows :

<u>For case (2), A=3, S=1</u>

$\underline{\tau=0}$ $W(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{60}(x^2+y^2)^3$

$\underline{\tau\neq 0}$ $U(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{60}(x^2+y^2)^3+W_{22}y^2$

$$V(x,y)=W_{11}y+W_{31}(x^2+y^2)y+W_{51}(x^2+y^2)^2y$$

<u>For case (3), A=4, S=1</u>

$\underline{\tau=0}$ $W(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{60}(x^2+y^2)^3$
$+W_{80}(x^2+y^2)^4$

$\underline{\tau \neq o}$ $U(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{60}(x^2+y^2)^3$

$+W_{80}(x^2+y^2)^4+W_{22}y^2$

$V(x,y)=W_{11}y+W_{31}(x^2+y^2)y+W_{51}(x^2+y^2)^2y$

$+W_{71}(x^2+y^2)^3y$

For case (4), A=5,S=1

$\underline{\tau=o}$ $W(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{60}(x^2+y^2)^3$

$+W_{80}(x^2+y^2)^4 + W_{100}(x^2+y^2)^5$

$\underline{\tau \neq o}$ $U(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{60}(x^2+y^2)^3$

$+W_{80}(x^2+y^2)^4+W_{100}(x^2+y^2)^5+W_{22}y^2$

$V(x,y)=W_{11}y+W_{31}(x^2+y^2)y+W_{51}(x^2+y^2)^2y$

$+W_{71}(x^2+y^2)^3y+W_{91}(x^2+y^2)^4y$

For case (5),A=2,S=1

$\underline{\tau=o}$ $W(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2$

$\underline{\tau \neq o}$ $U(x,y)=W_{20}(x^2+y^2)+W_{40}(x^2+y^2)^2+W_{22}y^2$

$V(x,y)=W_{11}y + W_{31}(x^2+y^2)y$

The general expression for the above cases may be written as follows

(a) <u>for the even part of the aberration ($\tau \neq o$)</u>

$$U_j = \sum_n \sum_q W(n,q).(x_j^2+y_j^2)^{(n-q)/2} . y_j^q \qquad (5.16)$$

with n,q both even, j=1,2,...N. These N equations may be inverted to give

$$W(n,q) = \sum_{j=1}^{N} K(n,q;j)U_j \qquad (5.17)$$

(b) <u>for the odd part of the aberration ($\tau \neq o$)</u>

$$V_j = \sum_n \sum_q W(n,q) \ (x_j^2+ y_j^2)^{(n-q)/2} \ y_j^q \qquad (5.18)$$

with n,q both odd, j=1,2,... N-1. These (N-1) equations may be inverted to give

$$W(n,q)= \sum_{j=1}^{N-1} K(n,q;j) \ V_j \qquad (5.19)$$

(c) <u>for the axial image (S=o)</u>

$$W_j = \sum_n W(n,o) \ y_j^n \qquad (5.20)$$

n being even, and j=1,2,...,A. These A equations may be inverted to give

$$W(n,o) = \sum_{j=1}^{A} K(n,o,j)W_j \qquad (5.21)$$

The actual permissible values of n and q depend on the values of A and S, and these are tabulated in Table 5-1 for reference.

(V.iii) Evaluation of the variance E of the wave-aberration

The aberration polynomials given above are functions of the co-ordinates (x,y) of the equivalent circular pupil. The aberrations W_j calculated in the ray trace depend on the scaled pupil co-ordinates $(x_s,y_s)_j$ - see expression (5.3). In consequence the aberration coefficients in the polynomial have no direct significance in analytical aberration theory (except when τ=o), and are only used as intermediate auxiliary quantities for the present purpose. The advantage of the procedure lies in reducing all cases to the problem of a circular aperture, and the resulting possibility of calculating the variance E (of both the axial and extra-axial images) for any system very simply once the aberrations W_j's are known. This simplification follows directly from the use of universal coefficients P(i,j), Q(i,j) and R(i,j) which are introduced in the following way :

Referring to the expression for the variance of W(x,y) as given by (1.5), i.e.

$$E = \int\int_{\text{Pupil}} [W(x,y)]^2 \, dA - \left[\int\int_{\text{Pupil}} W(x,y)dA\right]^2 \qquad (5.22)$$

if we assume the equivalent pupil to be $x^2+y^2=1$, then

$dA = (dx.dy)/\pi$. It is more convenient to use polar co-ordinates to evaluate (5.22), so we write

$$x_i = r_i \cos \phi_i$$

$$y_i = r_i \sin \phi_i$$

The general expression for the wave-aberration is then

$$W = \sum_n \sum_q W(n,q)\ r^n \cos^q \phi \qquad (5.23)$$

The even and odd part of W contribute separately to E, so that

$$E = E_{even} + E_{odd} \qquad (5.24)$$

where

$$E_{even} = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} U^2.rdrd\phi - \left[\frac{1}{\pi} \int_0^1 \int_0^{2\pi} U.rdrd\phi \right]^2 \qquad (5.25)$$

$$E_{odd} = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} V^2.rdrd\phi \qquad (5.26)$$

the integration of the odd function $V(r,\phi)$ over the pupil being zero. If E is evaluated for the general aberration function (5.23) we have

$$E = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} \left[\sum_n \sum_q W(n,q) r^n \cos^q \phi \right]^2 r.drd\phi$$

$$- \left[\frac{1}{\pi} \int_0^1 \int_0^{2\pi} \sum_n \sum_q W(n,q) r^n \cos^q \phi .rdr.d\phi \right]^2$$

i.e.

$$E = \sum_m \sum_p \sum_n \sum_q W(m,p)W(n,q) \int_o^1 r^{m+n+1}\, dr . \frac{1}{\pi} \int_o^{2\pi} \cos^{p+q}\phi d\phi$$

$$- \left[\sum_n \sum_q W(n,q) \int_o^1 r^{n+1} dr . \frac{1}{\pi} \int_o^{2\pi} \cos^q \phi d\phi \right]^2 \qquad (5.27)$$

Consider now the first part of (5.27):

$$\frac{1}{\pi} \int_o^{2\pi} \cos^{p+q}\phi d\phi = 2 . \frac{1.3.5 \ldots (p+q-1)}{2.4.6 \ldots (p+q)}$$

if $(p+q) \geqslant 2$; $(p+q)$ being always even since p,q are either both even or both odd.

The value of the second part of (5.27) depend on the value of q, since

$$\frac{1}{\pi} \int_o^{2\pi} \cos^q \phi . d\phi = o \quad \text{if } q = \text{odd}$$

$$= 2 . \frac{1.3 \ldots (q-1)}{2.4 \ldots q} \quad \text{if } q = \text{even}$$

Substituting these results in (5.27) lead to the following results :-

(a) <u>For even aberrations (p,q both even), $\tau \neq o$.</u>

$$E_{even} = \sum_m \sum_n \sum_p \sum_q W(m,p)W(n,q) \left\{ \left(\frac{2}{m+n+2} \right) \left[\frac{1.3 \ldots (p+q-1)}{2.4 \ldots (p+q)} \right] \right.$$

$$\left. - \frac{4}{(m+2)(n+2)} \left[\frac{1.3 \ldots (p-1)}{2.4 \ldots p} \right] . \left[\frac{1.3 \ldots (q-1)}{2.4 \ldots q} \right] \right\} \qquad (5.28)$$

(b) For the odd part of the aberrations (p,q both odd)

$$E_{odd} = \sum_m \sum_n \sum_p \sum_q W(m,p)W(n,q) \left\{ \frac{2}{(m+n+2)} \left[\frac{1.3\ldots(p+q-1)}{2.4\ldots(p+q)} \right] \right\} \tag{5.29}$$

The terms in the square brackets in (5.28) and (5.29) have the value unity if (p+q)=o or p=q=o respectively.

(c) For the axial image case, r=o, (p,q both zero)

$$E_{axial} = \sum_m \sum_n W(m,o)W(n,o) \left\{ \frac{2}{(m+n+2)} - \frac{4}{(m+2)(n+2)} \right\} \tag{5.30}$$

The aberration coefficients from (5.17), (5.19) and (5.21) may now be substituted in (5.28), (5.29) and (5.30) respectively. The final forms of E_{even}, E_{odd} and E_{axial} are

$$E_{even} = \sum_{i=1}^{N} \sum_{j=1}^{N} Q(i,j)\, U_i U_j \tag{5.31}$$

$$E_{odd} = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} R(i,j) V_i V_j \tag{5.32}$$

$$E_{axial} = \sum_{i=1}^{A} \sum_{j=1}^{A} P(i,j) W_i W_j \tag{5.33}$$

where

$$Q(i,j)=\sum_m\sum_n\sum_p\sum_q\left\{\frac{2}{(m+n+2)}\left[\frac{1.3\ .....(p+q-1)}{2.4\ .....(p+q)}\right]\right.$$

$$\left.-\frac{4}{(m+2)(n+2)}\left[\frac{1.3....(p-1)}{2.4.....(p)}\right].\left[\frac{1.3....(q-1)}{2.4....(q)}\right]\right\}K(m,p;i)$$

$$\times K(n,q;j) \qquad (5.34)$$

$$R(i,j)=\sum_m\sum_n\sum_p\sum_q\left\{\frac{2}{(m+n+2)}\left[\frac{1.3...(p+q-1)}{2.4...(p+q)}\right]\right\}K(m,p;i)$$

$$\times K(n,q;j) \qquad (5.35)$$

$$P(i,j)=\sum_m\sum_n\frac{2mn}{(m+2)(n+2)(m+n+2)}K(m,o;i)K(n,o;j) \qquad (5.36)$$

The values of n,q (also m,p) are those given in Table 5.1 corresponding to the particular values of A and S. The coefficients K(n,q;j), K(m,p;i) as defined earlier are known constants once the values of A and S are given, since the ray trace pattern will then always be the same in the equivalent circular pupil. Thus the above coefficients Q(i,j), R(i,j) and P(i,j) may be calculated once for all for any given values of (A,S).

(V-iv) The computation of the coefficients P(i,j),Q(i,j) and R(i,j)

Referring to Table 5.1, there are a total of five cases to be considered; and for each case the inverted matrix coefficients K(n,q;j) associated with W_j, U_j and V_j must be found (see equations 5.21, 5.17, 5.19).

(V-iv)A Case (1), A=3,S=3

The total number of rays to be considered are N $(=A+\frac{S+1}{2})$;for the present case, N=5, with the rays numbered j=1,2,...5.

The A(=3) aperture-ray co-ordinates (see equation 5.4) are :

$$\left.\begin{aligned} x_j &= o \\ y_j &= \pm\sqrt{\frac{(A+1)-j}{A}} = +\sqrt{\frac{4-j}{3}} \end{aligned}\right\} \quad j = 1,2,3$$

From (5.2), the azimuth angles ϕ_i, i=1,2 are

$$\phi_1 = \pi/4 \quad , \quad \phi_2 = \pi/2$$

and using (5.5), (5.6), the ray co-ordinates for j=4,5 are :

$$\left.\begin{aligned} x_4 &= \pm\sqrt{\frac{A-1}{A}} \cdot \frac{1}{\sqrt{2}} = \pm\frac{1}{\sqrt{3}} \\ y_4 &= +\frac{1}{\sqrt{3}} \end{aligned}\right\} \quad j=4$$

$$\left.\begin{aligned} x_5 &= 1 \\ y_5 &= o \end{aligned}\right\} \quad j = 5$$

The ray pattern is as shown in Fig.5.1. The aberration W_j

associated with the j^{th} ray is given by

$$W_j = \sum_n \sum_q W(n,q).(x_j^2 + y_j^2)^{(n-q)/2}\, y_j^q \qquad (5.37)$$

For the even part of the aberration

Table 5.1 gives for A=3, S=3 the permissible values of n,q corresponding to the even part of the aberration ($W_j = U_j$) as

$$\left.\begin{array}{ll} q=o\ ; & n = 2,4,6 \\ q=2\ ; & n = 2,4 \end{array}\right\} \text{ for equation (5.37).}$$

Let $A(j,k)=(x_j^2+y_j^2)^{(n-q)/2}\, y_j^q$; $j=1,2,\ldots,5$

$k=1,2,\ldots,5$

with k=1 corresponding to q=o, n=2

k=2 " " q=o, n=4

k=3 " " q=o, n=6 see flow diagram.

k=4 " " q=2, n=2

k=5 " " q=2, n=4

Thus the coefficients A(j,k) may be calculated for j= 1,2,...,5 ; k=1,2,...,5, and stored as a 5x5 matrix, i.e.

$$\begin{array}{llll} A(1,1) & A(1,2) & \ldots\ldots & A(1,5) \\ A(2,1) & A(2,2) & \ldots\ldots & A(2,5) \\ \ldots\ldots & & & \\ A(5,1) & A(5,2) & \ldots\ldots & A(5,5) \end{array}$$

These may then be inverted to give the values of a(k,j) corresponding to the coefficients K(n,q;j) in equation (5.17) for the even part of the aberration.

For the odd part of the aberration

Table 5.1 gives for A=3, S=3 the permissible values of n,q corresponding to the odd part of the aberration ($W_j=V_j$) as

$$\left.\begin{array}{l} q=1,\ n=1,3,5 \\ q=3,\ n=3 \end{array}\right\} \text{ for equation (5.37)}$$

Let $B(j,k) = (x_j^2+y_j^2)^{(n-q)/2}\ y_j^q$

$$\begin{array}{l} j=1,2,3,4(=N-1) \\ k=1,2,3,4 \end{array}$$

with k=1 corresponding to q=1, n=1

k=2 " " q=1, n=3 see flow diagram

k=3 " " q=1, n=5

k=4 " " q=3, n=3

Thus the coefficients B(j,k) may be calculated for j= 1,2,3,4; k = 1,2,3,4, and stored as a 4x4 matrix, i.e.

$$\begin{array}{llll} B(1,1) & B(1,2) & \ldots & B(1,4) \\ B(2,1) & \ldots\ldots & & B(2,4) \\ \ldots\ldots & & & \\ B(4,1) & \ldots\ldots & & B(4,4) \end{array}$$

These may then be inverted to give coefficients b(k,j) corresponding to the coefficients K(n,q;j) in equation (5.19) for the odd part of the aberration.

Axial image case.

For the axial image case (τ=o), A=3, S=o, Table 5-1 gives the permissible values of n,q as

q=o; n=2,4,6 for equation (5.37)

Let $C(j,k) = (x_j^2+y_j^2)^{(n-q)/2}\, y_j^q$; j=1,2,3(=A)

k=1,2,3.

with k=1 corresponding to n=2
k=2 " " n=4
k=3 " " n=6
} q=o (see flow diagram)

Thus the coefficients C(j,k), j=1,2,3; k=1,2,3 may be calculated and stored as a 3x3 matrix, i.e.

C(1,1) C(1,2) C(1,3)
C(2,1) C(2,2) C(2,3)
C(3,1) C(3,2) C(3,3)

These may then be inverted to give coefficients c(k,j) corresponding to K(n,o;j) in equation (5.21) for the axial case.

Referring to equation (5.34), the permissible values of n,q (and therefore m,p) for the summations are given by Table 5.1. Thus for the even aberration case, A=3, S=3, we have the following combinations:

q=o ; n = 2,4,6

p=o ; m = 2,4,6

q=2 ; n = 2,4

p=2 ; m = 2,4

Substituting these in (5.34) and simplifying we have the following expression for Q(i,j)

$$Q(i,j)= \sum_{\substack{m=2,4,6\\ n=2,4,6}} \sum \frac{2mn}{(m+n+2)(m+2)(n+2)} K(m,o;i)K(n,o;j)$$

$$+ \sum_{\substack{m=2,4,6\\ n=2,4}} \sum \frac{mn}{(m+n+2)(m+2)(n+2)} K(m,o;i)K(n,2;j)$$

$$+ \sum_{\substack{m=2,4\\ n=2,4,6}} \sum \frac{mn}{(m+n+2)(m+2)(n+2)} K(m,2;i)K(n,o;j)$$

$$+ \sum_{\substack{m=2,4\\ n=2,4}} \sum \frac{3mn+2(m+n+2)}{4(m+n+2)(m+2)(n+2)} K(m,2;i)K(n,2;j) \qquad (5.38)$$

with i=1,2,...,5 ; j = 1,2,...,5

Referring to equation (3.35), the permissible values n,q (and therefore m,p) for the summations are given by Table 5.1 as for the odd aberration case. Thus for A=3, S=3, R(i,j) may be written as

$$R(i,j) = \sum_{\substack{m \\ m=1,3,5 \\ n=1,3,5}} \sum_{n} \frac{1}{(m+n+2)} K(m,1;i)\, K(n,1;j)$$

$$+ \sum_{\substack{m \\ m=1,3,5 \\ n=3}} \sum_{n} \frac{0.75}{(m+n+2)} K(m,1;i) K(n,3;j)$$

$$+ \sum_{\substack{m \\ m=3 \\ n=1,3,5}} \sum_{n} \frac{0.75}{(m+n+2)} K(m,3;i) K(n,1;j)$$

$$+ \sum_{\substack{m \\ m=3 \\ n=3}} \sum_{n} \frac{5}{8(m+n+2)} K(m,3;i) K(m,3;j) \qquad (5.39)$$

with i=1,2,3,4; j=1,2,3,4. The above formula is expressed in a standard form for easy comparison with later formulae.

Referring to equation (5.36), the permissible values of n (and therefore m) are obtained from Table 5.1 as for A=3, S=0 case. The expression for P(i.j) may now be written as

$$P(i,j)= \sum_{m}\sum_{n} \frac{2mn}{(m+n+2)(m+2)(n+2)} K(m,o;i)K(n,o;j)$$

$$\begin{pmatrix} m=2,4,6 \\ n=2,4,6 \end{pmatrix} \qquad (5.39)A$$

with i=1,2,3(=A); j= 1,2,3.

The flow diagrams for P(i,j), Q(i,j) and R(i,j) computations are given on the following pages.

FLOW DIAGRAMS (FOR $P(i,j)$, $Q(i,j)$ AND $R(i,j)$ COMPUTATIONS)
A=3, S=3 AND 0 (CASE 1).

EVEN-ABERRATIONS ($T \neq 0$)

STORE $A(j,k)$ AS 5×5 MATRIX.

CYCLE j ($j=1,2,3$)

SET $r = \sqrt{\frac{4-j}{3}}$

$A(j,1) = r^2$
$A(j,2) = r^4$
$A(j,3) = r^6$
$A(j,4) = r^2$
$A(j,5) = r^4$

REPEAT j

SET $r = \sqrt{2/3}$

$A(4,1) = r^2$
$A(4,2) = r^4$
$A(4,3) = r^6$
$A(4,4) = r^2/2$
$A(4,5) = r^4/4$
$A(5,1) = 1$
$A(5,2) = 1$
$A(5,3) = 1$
$A(5,4) = 0$
$A(5,5) = 0$

ODD-ABERRATIONS

STORE $B(j,k)$ AS 4×4 MATRIX.

CYCLE j ($j=1,2,3$)

SET $r = \sqrt{\frac{4-j}{3}}$

$B(j,1) = r$
$B(j,2) = r^3$
$B(j,3) = r^5$
$B(j,4) = r^3$

REPEAT j

SET $s = 1/\sqrt{2}$
$r = \sqrt{2/3}$

$B(4,1) = r \cdot s$
$B(4,2) = r^3 \cdot s$
$B(4,3) = r^5 \cdot s$
$B(4,4) = r^3 \cdot s^3$

W AXIAL CASE ($T=0$)

STORE $C(j,k)$ AS 3×3 MATRIX.

CYCLE j ($j=1,2,3$)

SET $r = \sqrt{\frac{4-j}{3}}$

$C(j,1) = r^2$
$C(j,2) = r^4$
$C(j,3) = r^6$

REPEAT j

SUBROUTINE: INVERT THE MATRICES A, B, C [USING I.B.M. LIBRARY ROUTINE "MATINV"]

i.e. $A(j,k) \rightarrow a(k,j)$
STORE $a(k,j)$ in $K(n,q;j)$

CYCLE k (k=1, 2, 3)
  CYCLE j (j=1, 2, 3, 4, 5)
    $K(2k, 0; j) = a(k,j)$
  REPEAT j
REPEAT k

CYCLE k (k=4, 5)
  CYCLE j (j=1, 2, 3, 4, 5)
    $K(2k-6, 2; j) = a(k,j)$
  REPEAT j
REPEAT k

PRINT OUT:
INVERTED MATRIX
$K(n,q;j)$
A=3, S=3 W=EVEN

$K(2,0;1)$ $K(2,0;2)$ ..... $K(2,0;5)$
$K(4,0;1)$ $K(4,0;2)$ ..... $K(4,0;5)$
$K(6,0;1)$ $K(6,0;2)$ ..... $K(6,0;5)$
$K(2,2;1)$ $K(2,2;2)$ ..... $K(2,2;5)$
$K(4,2;1)$ $K(4,2;2)$ ..... $K(4,2;5)$

i.e. $B(j,k) \rightarrow b(k,j)$
STORE $b(k,j)$ in $K(n,q;j)$

CYCLE k (k=1, 2, 3)
  CYCLE j (j=1, 2, 3, 4)
    $K(2k-1, 1; j) = b(k,j)$
  REPEAT j
REPEAT k

CYCLE j (j=1, 2, 3, 4)
  $K(3, 3; j) = b(4,j)$
REPEAT j

PRINT OUT:
INVERTED MATRIX
$K(n,q;j)$
A=3, S=3 W=ODD

$K(1,1;1)$ $K(1,1;2)$ ..... $K(1,1;4)$
$K(3,1;1)$ $K(3,1;2)$ ..... $K(3,1;4)$
$K(5,1;1)$ $K(5,1;2)$ ..... $K(5,1;4)$
$K(3,3;1)$ $K(3,3;2)$ ..... $K(3,3;4)$

i.e. $C(j,k) \rightarrow c(k,j)$
STORE $c(k,j)$ in $K(n,q;j)$

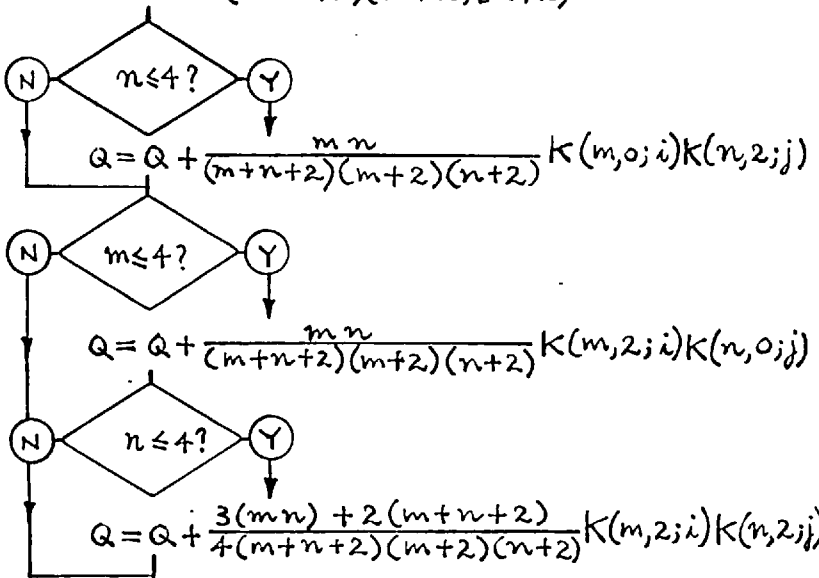
CYCLE k (k=1, 2, 3)
  CYCLE j (j=1, 2, 3)
    $K(2k, 0; j) = c(k,j)$
  REPEAT j
REPEAT k

PRINT OUT:
INVERTED MATRIX
$K(n,q;j)$
A=3, S=0 (T=0)

$K(2,0;1)$ $K(2,0;2)$ $K(2,0;3)$
$K(4,0;1)$ $K(4,0;2)$ $K(4,0;3)$
$K(6,0;1)$ $K(6,0;2)$ $K(6,0;3)$

CYCLE i ($i=1,2,\cdots,5$)
CYCLE j ($j=1,2,\cdots,5$)
SET $Q=0.0$
CYCLE m ($m=2,4,6$)
CYCLE n ($n=2,4,6$)

$$Q = Q + \frac{2mn}{(m+n+2)(m+2)(n+2)} K(m,0;i)K(n,0;j)$$



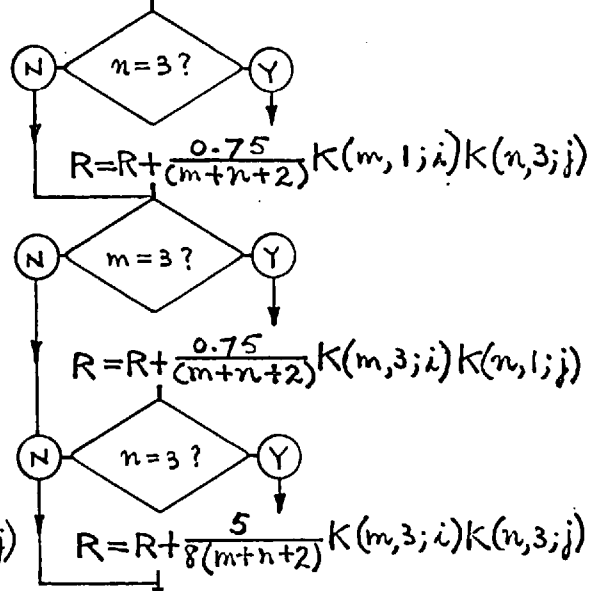
REPEAT n
REPEAT m
$Q(i,j) = Q$ [SEE EQUN 5-38]
REPEAT j
REPEAT i
PRINT OUT: $Q(i,j)\begin{cases} i=1,2,\cdots,5 \\ j=1,2,\cdots,5 \end{cases}$

$A=3$, $S=3$

$Q(1,1)$ $Q(1,2)$ _____ $Q(1,5)$
$Q(2,1)$ $Q(2,2)$ _____ $Q(2,5)$
_____
$Q(5,1)$ $Q(5,2)$ _____ $Q(5,5)$

CYCLE i ($i=1,2,3,4$)
CYCLE j ($j=1,2,3,4$)
SET $R=0.0$
CYCLE m ($m=1,3,5$)
CYCLE n ($n=1,3,5$)

$$R = R + \frac{1}{(m+n+2)} K(m,1;i)K(n,1;j)$$



REPEAT n
REPEAT m
$R(i,j) = R$ [SEE EQUN 5-39]
REPEAT j
REPEAT i
PRINT OUT: $R(i,j)\begin{cases} i=1,2,3,4 \\ j=1,2,3,4 \end{cases}$

$A=3$, $S=3$

$R(1,1)$ $R(1,2)$ $R(1,3)$ $R(1,4)$
$R(2,1)$ _____ $R(2,4)$
_____
$R(4,1)$ _____ $R(4,4)$

CYCLE i ($i=1,2,3$)
CYCLE j ($j=1,2,3$)
SET $P=0.0$
CYCLE m ($m=2,4,6$)
CYCLE n ($n=2,4,6$)

$$P = P + \frac{2mn}{(m+n+2)(m+2)(n+2)} \{1\}$$

$K(m,0;i)K(n,0;j)$

REPEAT n
REPEAT m
$P(i,j) = P$ [SEE EQUN 5-39A]
REPEAT j
REPEAT i
PRINT OUT:
$P(i,j)$ $i=1,2,3$ $j=1,2,3$

$A=3$, $S=0$

$P(1,1)$ $P(1,2)$ $P(1,3)$
$P(2,1)$ $P(2,2)$ $P(2,3)$
$P(3,1)$ $P(3,2)$ $P(3,3)$

(V-iv)B Cases(2) to (5), A=2,3,4,5; S=0,1

There are four different values of A to be considered. For each case, S=1 (S=0 for axial object), and the total number of rays to be considered is N(=A+1).

The A aperture-ray co-ordinates (see equation 5.4) are :

$$\left.\begin{aligned} x_j &= 0 \\ y_j &= \pm\sqrt{\frac{(A+1)-j}{A}} \end{aligned}\right\} \qquad j=1,2,\ldots A$$

The $(A+1)^{th}$ ray has azimuth angle $\phi = \pi/2$ and from (5.5) and (5.6), the pupil co-ordinates for this ray is

$$\left.\begin{aligned} x_{(A+1)} &= 1 \\ y_{(A+1)} &= o \end{aligned}\right\} \quad \text{corresponding to } j=(A+1)$$

Referring to the aberration polynomial (5.37) and Table 5.1, the permissible value of n,q corresponding to the even part of the aberration ($W_j=U_j$) are

$$\begin{aligned} q &= o, \quad n = 2,4,6,\ldots,2A \\ q &= 2, \quad n = 2 \end{aligned} \qquad (5.40)$$

As in the section (V-iv)A, we let

$$A(j,k)=(x_j^2+y_j^2)^{(n-q)/2}\, y_j^q \; ; \; j=1,2,\ldots,(A+1), \quad k=1,2,\ldots,(A+1)$$

with k=1,2,3,...,A corresponding to n=2,4,6,...2A respectively, q being zero, and

k=(A+1) corresponding to n=2, q=2.

Thus the coefficients A(j,k) may be calculated for j= 1,2,...(A+1); k=1,2,....(A+1), and these may be inverted to give a(k,j)'s, corresponding to K(n,q;j) coefficients in equation (5.17) for the even part of the aberration, as described in (V-iv)A.

For the odd part of the aberration, we again refer to the aberration polynomial (5.37) and Table 5.1, the permissible values of n,q being

$$q=1;\ n=1,3,5,\ldots,(2A-1) \qquad (5.41)$$

As before, we let

$$B(j,k)=(x_j^2+y_j^2)^{(n-q)/2}\ y_j^q;\ j=1,2,\ldots,A \quad k=1,2,\ldots,A \qquad (5.42)$$

and corresponding to each k we have n=2k-1 and q=1 in (5.42). Thus the coefficients B(j,k) may be calculated for j=1,2,...,A; k=1,2,...,A and these coefficients may be inverted to give coefficients b(k,j) corresponding K(n,q;j) in equation (5.19) for the odd part of the aberration (see flow diagram).

Lastly, for the axial image case, the permissible values of n,q are

$$q=o;\quad n=2,4,6,\ldots,2A \qquad (5.43)$$

leading to values of C(j,k) as described in (V-iv)A. The inverted coefficients, c(k,j), give the values of K(n,o;j) required (see flow diagram).

P(i,j),Q(i,j) and R(i,j) formulae for S=o and 1.

Referring to (5.34), the permissible values of n,q (and therefore m,p) are given by (5.40). Substituting these values in (5.34) and simplifying we have the following expression for Q(i,j), with S=1, A=2,3,4 or 5.

$$Q(i,j) = \sum_{m}\sum_{n} \frac{2mn}{(m+n+2)(m+2)(n+2)} K(m,o;i)K(n,o;j)$$

$$\begin{pmatrix} m=2,4,6,\ldots,2A \\ n=2,4,6,\ldots,2A \end{pmatrix}$$

$$+ \sum_{m} \frac{m}{2(m+2)(m+4)} K(m,o;i)K(2,2;j)$$

$$(m=2,4,6,\ldots,2A)$$

$$+ \sum_{n} \frac{n}{2(n+2)(n+4)} K(m,2;i)K(n,o;j)$$

$$(n=2,4,6,\ldots,2A)$$

$$+ \frac{1}{16} K(2,2;i)K(2,2;j) \qquad (5.44)$$

with i=1,2....,(A+1); j=1,2,...,(A+1)

Referring to equation (5.35) for the odd aberration case, the permissible values of n,q (and there m,p) for the summations are given by (5.41). Using these values we have the following expression for R(i,j) with S=1,A=2,3,4 or 5.

$$R(i,j) = \sum_{m}\sum_{n} \frac{1}{(m+n+2)} .K(m,1;i)K(n,1;j) \qquad (5.45)$$

$$\begin{pmatrix} m=1,3,5,\ldots,2A-1 \\ n=1,3,5,\ldots,2A-1 \end{pmatrix}$$

with $i=1,2,3,\ldots A$; $j=1,2,\ldots A$.

Finally, for the axial image case, the $P(i,j)$ coefficient in equation (5.36) take the values of q,n (and therefore p,m) as given by (5.43). Using these values we have the following expression for $P(i,j)$ with $S=o, A=2, 4$, or 5.

$$P(i,j) = \sum_{m}\sum_{n} \frac{2mn}{(m+n+2)(m+2)(n+2)} K(m,o;i)K(n,o;j)$$

$$\begin{pmatrix} m=2,4,6,\ldots 2A \\ n=2,4,6,\ldots 2A \end{pmatrix} \qquad (5.46)$$

with $i=1,2,\ldots,A$; $j=1,2,\ldots,A$.

Note the expression for $P(i,j)$ with $A=3$, $S=o$ is given by (5.39)A.

The flow diagrams for $P(i.j)$, $Q(i,j)$ and $R(i,j)$ computations will be found on the following pages for $A=2,3,4$ or 5, S being equal to unity for Q and R coefficients and zero for P coefficients.

FLOW DIAGRAMS (FOR $P(i,j)$, $Q(i,j)$ AND $R(i,j)$ COMPUTATIONS)

$A = 2, 3, 4$ OR 5 ; $S = 0$ OR 1

Read-in the value of "A"
Set $\alpha = A$

EVEN ABERRATIONS ($\tau \neq 0$)

Store $A(j,k) \rightarrow (\alpha+1) \times (\alpha+1)$ matrix

CYCLE j ($j = 1, 2, 3, \cdots, \alpha$)

$r = \sqrt{\frac{(A+1)-j}{A}}$

$A(j, \alpha+1) = r^2$

$A(j, 1) = r^2$

$A(j, 2) = r^4$

$A(j, 3) = r^6$

.....

$A(j, \alpha) = r^{2\alpha}$

REPEAT j

$A(\alpha+1, \alpha+1) = 0$

$A(\alpha+1, 1) = 1$

$A(\alpha+1, 2) = 1$

$A(\alpha+1, 3) = 1$

.....

$A(\alpha+1, \alpha) = 1$

ODD ABERRATIONS

Store $B(j,k) \rightarrow \alpha \times \alpha$ matrix

CYCLE j ($j = 1, 2, 3, \cdots, \alpha$)

$r = \sqrt{\frac{(A+1)-j}{A}}$

$B(j, 1) = r$

$B(j, 2) = r^3$

$B(j, 3) = r^5$

.....

$B(j, \alpha) = r^{(2\alpha-1)}$

REPEAT j

W AXIAL CASE ($T = 0$)

Store $C(j,k) \rightarrow \alpha \times \alpha$ matrix

CYCLE j ($j = 1, 2, 3, \cdots, \alpha$)

$r = \sqrt{\frac{(A+1)-j}{A}}$

$C(j, 1) = r^2$

$C(j, 2) = r^4$

.....

$C(j, \alpha) = r^{2\alpha}$

REPEAT j

SUBROUTINE : Invert the matrices A, B and C [using library routine MATINV (I.B.M.)]

i.e. $A(j,k) \rightarrow a(k,j)$

Store $a(k,j)$ in $K(n,q;j)$

CYCLE k ($k = 1, 2, \cdots, \alpha$)
  CYCLE j ($j = 1, 2, \cdots, \alpha+1$)
    $K(2k, 0; j) = a(k,j)$
  REPEAT j
REPEAT k

CYCLE j ($j = 1, 2, \cdots, \alpha+1$)
  $K(2, 2; j) = a(\alpha+1, j)$
REPEAT j

PRINT OUT :

INVERTED MATRIX
W=EVEN, A=α, S=1

$K(n, 0; j)$ $n = 2, 4, 6, \cdots, 2\alpha$ $\Big\}$ $j = 1, 2, \cdots, \alpha+1$
$K(2, 2; j)$

$K(2,0;1)\ K(2,0;2) \ldots\ldots K(2,0;\alpha+1)$
$K(4,0;1)\ K(4,0;2) \ldots\ldots K(4,0;\alpha+1)$
$K(6,0;1)\ K(6,0;2) \ldots\ldots K(6,0;\alpha+1)$
$\ldots\ldots\ldots$
$K(2\alpha,0;1)\ K(2\alpha,0;2) \ldots\ldots K(2\alpha,0;\alpha+1)$
$K(2,2;1)\ K(2,2;2) \ldots\ldots K(2,0;\alpha+1)$

i.e. $B(j,k) \rightarrow b(k,j)$

Store $b(k,j)$ in $K(n,q;j)$

CYCLE k ($k = 1, 2, \cdots, \alpha$)
  CYCLE j ($j = 1, 2, \cdots, \alpha$)
    $K(2k-1, 1; j) = b(k,j)$
  REPEAT j
REPEAT k

PRINT OUT :

INVERTED MATRIX
W=ODD, A=α, S=1

$K(n, 1; j)$; $n = 1, 3, 5, \cdots (2\alpha-1)$
$j = 1, 2, 3, \cdots, \alpha$

$K(1,1;1)\ K(1,1;2) \ldots\ldots K(1,1;\alpha)$
$K(3,1;1)\ K(3,1;2) \ldots\ldots K(3,1;\alpha)$
$\ldots\ldots\ldots$
$K(2\alpha-1,1;1)\ K(2\alpha-1,1;2) \ldots K(2\alpha-1,1;\alpha)$

i.e. $C(j,k) \rightarrow c(k,j)$

Store $c(k,j)$ in $K(n,q;j)$

CYCLE k ($k = 1, 2, \cdots, \alpha$)
  CYCLE j ($j = 1, 2, \cdots, \alpha$)
    $K(2k, 0; j) = c(k,j)$
  REPEAT j
REPEAT k

PRINT OUT :

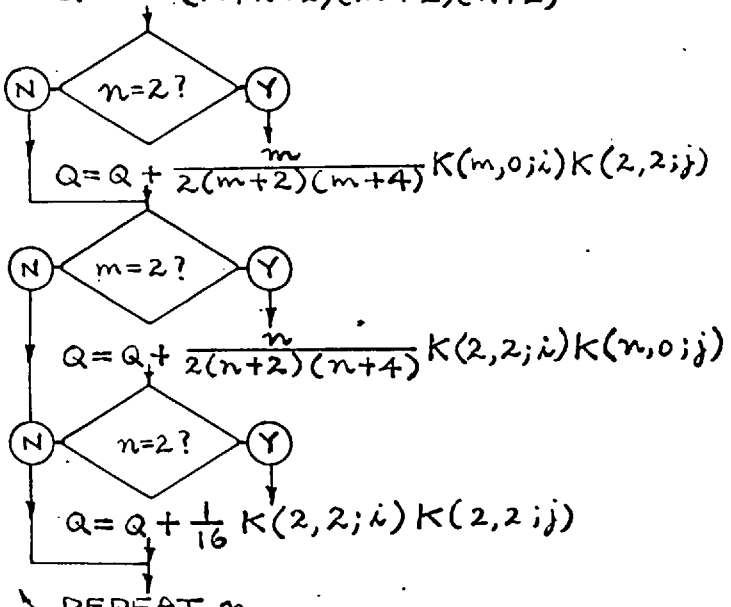
INVERTED MATRIX
AXIAL IMAGE CASE.
A=α, S=0

$K(n, 0; j)$; $n = 2, 4, 6, \cdots, 2\alpha$
$j = 1, 2, \cdots, \alpha$

$K(2,0;1)\ K(2,0;2) \ldots K(2,0;\alpha)$
$K(4,0;1)\ K(4,0;2) \ldots K(4,0;\alpha)$
$\ldots\ldots\ldots$
$K(2\alpha,0;1)\ K(2\alpha,0;2) \ldots K(2\alpha,0;\alpha)$

CYCLE i ($i=1,2,\cdots,\alpha+1$)
CYCLE j ($j=1,2,\cdots,\alpha+1$)
Set $Q=0.0$
CYCLE m ($m=2,4,6,\cdots,2\alpha$)
CYCLE n ($n=2,4,6,\cdots,2\alpha$)

$$Q=Q+\frac{2mn}{(m+n+2)(m+2)(n+2)}K(m,0;i)K(n,0;j)$$



REPEAT n
REPEAT m
$Q(i,j)=Q$ [SEE EQUN 5-44]
REPEAT j
REPEAT i
PRINT OUT: $Q(i,j)\begin{cases}i=1,2,\cdots,\alpha+1\\ j=1,2,\cdots,\alpha+1\end{cases}$
$A=\alpha$, $S=1$

$Q(1,1)\ Q(1,2)\ \cdots\cdots\ Q(1,\alpha+1)$
$Q(2,1)\ Q(2,2)\ \cdots\cdots\ Q(2,\alpha+1)$
$\cdots\cdots\cdots\cdots$
$Q(\alpha+1,1)\ Q(\alpha+1,2)\ \cdots\cdots\ Q(\alpha+1,\alpha+1)$

CYCLE i ($i=1,2,\cdots,\alpha$)
CYCLE j ($j=1,2,\cdots,\alpha$)
Set $R=0.0$
CYCLE m ($m=1,3,5,\cdots,2\alpha-1$)
CYCLE n ($n=1,3,5,\cdots,2\alpha-1$)

$$R=R+\frac{1}{(m+n+2)}K(m,1;i)K(n,1;j)$$

REPEAT n
REPEAT m
$R(i,j)=R$ [SEE EQUN 5-45]
REPEAT j
REPEAT i

PRINT OUT: $R(i,j)\begin{cases}i=1,2,\cdots,\alpha\\ j=1,2,\cdots,\alpha\end{cases}$
$A=\alpha$, $S=1$

$R(1,1)\ R(1,2)\ \cdots\ R(1,\alpha)$
$R(2,1)\ R(2,2)\ \cdots\ R(2,\alpha)$
$\cdots\cdots\cdots$
$R(\alpha,1)\ R(\alpha,2)\ \cdots\ R(\alpha,\alpha)$

CYCLE i ($i=1,2,\cdots,\alpha$)
CYCLE j ($j=1,2,\cdots,\alpha$)
Set $P=0.0$
CYCLE m ($m=2,4,6,\cdots,2\alpha$)
CYCLE n ($n=2,4,6,\cdots,2\alpha$)

$$P=P+\frac{2mn}{(m+n+2)(m+2)(n+2)}\times K(m,0;i)K(n,0;j)$$

REPEAT n
REPEAT m
$P(i,j)=P$
REPEAT j
REPEAT i

PRINT OUT: $P(i,j)\begin{cases}i=1,2,\cdots,\alpha\\ j=1,2,\cdots,\alpha\end{cases}$
$A=\alpha$, $S=0$

$P(1,1)\ P(1,2)\ \cdots\ P(1,\alpha)$
$P(2,1)\ P(2,2)\ \cdots\ P(2,\alpha)$
$\cdots\cdots\cdots$
$P(\alpha,1)\ P(\alpha,2)\ \cdots\ P(\alpha,\alpha)$

Discussion and Conclusion.

The computations of the P(i,j),Q(i,j) and R(i,j) coefficients were carried out on the Elliot 803A computer and also independently on the IBM7094 computer. The flow diagrams given above represent the version for 7094 using Fortran code, whilst the Elliot version is slightly more complicated due to its lack of facility for handling 3-D arrays (e.g. K(n,q;j)). The output from the two computers agreed within the accuracies specified (to six decimal places) with the exception of A=5 case (see below). Further spot checks were also attempted on a desk calculator to ensure that no common errors were made in the two independently written computer programmes.

From the formulae (5.34), (5.35) and (5.36), it was expected the square matrices Q(i,j), R(i,j) and P(i,j) to have diagonal symmetries about elements Q(i,i), R(i,i) and P(i,i) respectively; these provided further checks on the numerical accuracy of the output. Initially, it was found that for the special case A=5, S=1, the diagonal symmetries were not strictly observed; further detailed investigation showed that the lack of precision was introduced by the matrix inversion library routine used. Double precision arithmetics were then specified for this special case and the above mentioned inaccuracies were thereby rectified. Tabulated values of K(n,q;j), P(i,j), Q(i,j) and R(i,j) coefficients are presented on the following pages.

Further comments concerning the applications of the tabulated coefficients.

1) The calculation of the variance E.

It is appropriate here to emphasize again that the P(i,j),Q(i,j) and R(i,j) coefficients have only to be calculated once for all. For a given pair of A,S values, the variance E for an extra axial image case may be calculated from $E = E_{(even)} + E_{(odd)}$, as described earlier, with $E_{(even)}$ and $E_{(odd)}$ as given by the expressions (5.31) and (5.32) respectively, and $E=E_{(axial)}$ for $\tau=o$ as given by (5.33). In each case the tabulated values of P(i,j), Q(i,j) and R(i,j) corresponding to the particular A,S values are employed. The values of the wave aberrations U_j, V_j and W_j required in the variance formulae will be provided directly by the ray-tracing, to be described in Chapters 7 and 8. When the system has reached a high degree of correction, its Strehl Intensity ratio I may be calculated from Maréchal's approximate expression given by (1.4), using the value of E calculated above. In Chapter 8, further applications of the P, Q and R coefficients in the optimization procedure will be described.

2) The calculation of aberration coefficients W(n,q).

It has been seen in Chapter 3 that the numerical integration method employed in the Strehl Intensity calculation requires the values of W(n,q). These may be calculated for a given pair of A,S values using equations (5.17), (5.19) and

(5.21), and the permissible values of n,q are obtained from Table 5.1. The coefficients K(n,q;j) required are those tabulated on the following pages, and the wave aberrations W_j, U_j and V_j will be obtained from the ray-trace.

Taking the case A=3, S=1 as an example : Table 5.1 gives the permissible values of n,q for the even aberration U_j as

$q=o;\quad n= 2,4,6$

$q=2;\quad n= 2$

Recalling equation 5.17

$$W(n,q) = \sum_{j=1}^{N} K(n,q;j)\, U_j \qquad (5.47)$$

$$N = A + \frac{S+1}{2} = 4$$

The aberration coefficients (n,q both even) are thus :

$W(2,o),\quad W(4,o),\quad W(6,o),\quad W(2,2)$

where W(6,o), for example, is given by n=6, q=o in (5.47), i.e.

$$W(6,0)= K(6,o;1)U_1+K(6,o;2)U_2+K(6,o;3)U_3+K(6,o;4)U_4 \qquad (5.48)$$

From K(n,q;j) tables, corresponding to A=3, S=1, W=EVEN, we have

$K(6,o;1) = 4.50$

$K(6,o;2) = -13.50$

$K(6,o;3) = 13.50$

$K(6,o;4) = 0.0$

substituting these in (5.48) we have

$$W(6,o) = 4.5\ U_1 - 13.5\ U_2 + 13.5U_3$$

where U_1, U_2 and U_3 are obtained from the ray trace, corresponding to the j=1,2,3 rays in the ray pattern defined by A=3, S=1.

For the odd aberration V_j's, Table 5.1 gives for A=3, S=1 the following permissible n,q's

$$q=1 \ ; \quad n= 1,3,5$$

and the aberration coefficients W(n,q) to be considered are :

$$W(1,1), \quad W(3,1) \text{ and } W(5,1)$$

These may be calculated from equation (5.19), using the K(n,q;j) coefficients tabulated under A=3, S=1, W=ODD.

For the axial image case, τ=o and S=o. Table 5.1 gives for A=3 the permissible n,q's as

$$q=o \ ; \quad n = 2,4,6$$

The permissible aberration coefficients W(n,q) are W(2,o), W(4,o) and W(6,o).

These may be calculated from equation (5.21), using the K(n,o;j) coefficients tabulated under A=3, S=o, AXIAL CASE.

Referring to the general aberration polynomial given in (3.3), for the particular case A=3, S=1, the aberration coefficients W(8,0), W(10,0), W(2,2), W(4,2), W(4,4), W(3,3), W(7,1) and W(9,1) are not required; all these will be set to zero for the input when the Strehl Intensity programme is being used.

As pointed out earlier in Chapter 3, the main function of the Strehl Intensity programme is to investigate the relative influences of W(n,q)'s on the intensity ratio I. In optical design of high quality systems, the Maréchal's approximate expression for I, using the variance values described above, is preferable due to its simplicity.

COEFFICIENTS. -see flow diagram for reference.

A=2, S=0 AXIAL CASE.

INVERTED MATRIX K(n,q; j) :

$$\begin{pmatrix} -1 & 4 \\ 2 & -4 \end{pmatrix}$$

P(i,j) i=1,2 j=1,2

0.105556	-0.044444
-0.044444	0.088889

A=2, S=1 W=EVEN CASE

INVERTED MATRIX K(n,q; j) :

$$\begin{pmatrix} -2 & 4 & 1 \\ 2 & -4 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

Q(i,j) i=1,2,3 j=1,2,3

0.084722	-0.044444	-0.020834
-0.044444	0.088889	0.000000
-0.020834	0.000000	0.062500

A=2, S=1 W=ODD CASE.

INVERTED MATRIX K(n,q; j)

2.828427	-1.000000
2.000000	-2.828427

R(i,j) i=1,2 j=1,2

0.083333	0.000000
0.000000	0.333333

A=3 , S=1 (SEE FLOW DIAGRAM)

W=EVEN CASE

INVERTED MATRIX $K(n,q;j)$

-.0000000	-4.500000	9.000000	1.000000
-4.500000	18.00000	-22.50000	.0000000
4.500000	-13.50000	13.50000	.0000000
1.000000	.0000000	.0000000	-1.000000

Q(I,J) COEFS. I=1,2,3,4 ; J=1,2,3,4

.0772321	-.0441965	-.0120535	-.0395833
-.0441964	.2450894	-.1888393	.0562500
-.0120535	-.1888393	.2450894	-.0562500
-.0395833	.0562500	-.0562500	.0625000

AXIAL CASE A=3, S=0

$K(n,q;j)$ COEFS.

1.000000	-4.500000	9.000000
-4.500000	18.00000	-22.50000
4.500000	-13.50000	13.50000

P(I,J) COEFS. I=1,2,3 ; J=1,2,3

.0605655	.0120535	-.0683035
.0120535	.2450894	-.1888393
-.0683035	-.1888393	.2450894

W=ODD CASE A=3, S=1

$K(n,q;j)$ COEFS.

1.000000	-3.674235	5.196152
-4.500000	14.69694	-12.99038
4.500000	-11.02270	7.794229

R(I,J) COEFS. I=1,2,3 ; J=1,2,3

.0437500	.0153093	.0108253
.0153093	.2250000	-.0795495
.0108253	-.0795495	.2812500

A=3, S=3 W=EVEN CASE

INVERTED MATRIX $K(n,q;j)$ [SEE FLOW DIAGRAM]

3.00000000	-13.5000000	8.99999997	9.00000000	-2.00000000
-7.50000000	27.0000000	-22.4999999	-9.00000003	3.00000001
4.50000001	-13.5000000	13.5000000	.000000040	-.000000013
-1.99999999	8.99999994	.000000060	-8.99999997	1.99999999
2.99999999	-8.99999994	-.000000060	8.99999997	-2.99999999

Q(I,J) COEFS. I=5 J=5 MAX.

.027232148	-.025446370	.006696507	-.000000001	-.008333327
-.025446385	.582589328	-.245089203	-.393749941	.093749997
.006696478	-.245089352	.245089471	.056250055	-.075000003
-.000000004	-.393749951	.056250005	.450000017	-.056249989
-.008333332	.093749992	-.074999987	-.056249992	.049999999

A=4, S=0 AXIAL CASE

INVERTED MATRIX $K(n,q;j)$ — SEE FLOW DIAGRAM

-1.000000	5.333333	-12.00000	16.00000
7.333333	-37.33333	76.00000	-69.33333
-16.00000	74.66667	-128.0000	96.00000
10.66667	-42.66667	64.00000	-42.66667

P(I,J) COEFS. I=1,2,3,4 ; J=1,2,3,4

.0454498	.0245503	-.0410582	-.0177783
.0245503	.1896291	-.1151314	-.0812688
-.0410581	-.1151314	.3123837	-.1151352
-.0177780	-.0812683	-.1151338	.1896319

A=3, S=3 W=ODD CASE

INVERTED MATRIX $K(n,q;j)$ — SEE FLOW DIAGRAM

1.00000001	-3.67423469	5.19615249	-.000000006
-4.50000001	11.0227040	-12.9903812	5.19615243
4.50000003	-11.0227040	7.79422879	-.000000008
-.000000035	3.67423471	-.000000076	-5.19615242

R(I,J) COEFS. I=4, J=4 MAX.

.043750002	-.034445949	.010825312	.070364560
-.034445942	.210937524	-.099436878	-.139211651
.010825320	-.099436861	.281250015	.028124989
.070364572	-.139211643	.028124993	.721875000

A=4, S=1 ; W = ODD CASE

INVERTED MATRIX $K(n,q;j)$ — SEE FLOW DIAGRAM

-1.000000	4.618802	-8.485281	8.000000
7.333333	-32.33161	53.74011	-34.66667
-16.00000	64.66323	-90.50967	48.00000
10.66667	-36.95042	45.25483	-21.33333

R(I,J) COEFS. I = 1,2,3,4 ; J = 1,2,3,4

.0283069	.0183285	-.0000002	-.0105821
.0183286	.1777778	-.0777617	.0733145
-.0000001	-.0777620	.2571439	-.1346863
-.0105820	.0733141	-.1346872	.2624340

A=4, S=1 ; W = EVEN CASE

INVERTED MATRIX $K(n,q;j)$ — SEE FLOW DIAGRAM

-2.000000	5.333333	-12.00000	16.00000	1.000000
7.333333	-37.33333	76.00000	-69.33333	.0000000
-16.00000	74.66667	-128.0000	96.00000	.0000000
10.66667	-42.66667	64.00000	-42.66667	.0000000
1.000000	.0000000	.0000000	.0000000	-1.000000

Q(I,J) COEFS. I = 1,2,···,5 ; J = 1,2,···,5

.0690609	-.0198941	-.0410582	.0266668	-.0430556
-.0198941	.1896291	-.1151314	-.0812688	.0444445
-.0410583	-.1151314	.3123837	-.1151352	-.0000000
.0266665	-.0812683	-.1151338	.1896319	-.0444444
-.0430556	.0444445	-.0000000	-.0444444	.0625000

A=5 , S=0 AXIAL CASE

COEFFICIENTS (IN I.B.M. OUTPUT NOTATION) (SEE FLOW DIAGRAM)

INVERTED MATRIX K(N,Q; J)

0.99999928D 00	-0.62499966D 01	0.16666660D 02	-0.24999995D 02	0.24999998D 02
-0.10416659D 02	0.63541629D 02	-0.16249993D 03	0.22291660D 03	-0.16041664D 03
0.36458308D 02	-0.21354154D 03	0.51041642D 03	-0.61458310D 03	0.36979156D 03
-0.52083301D 02	0.28645818D 03	-0.62499969D 03	0.67708303D 03	-0.36458319D 03
0.26041653D 02	-0.13020827D 03	0.26041653D 03	-0.26041654D 03	0.13020827D 03

P(I,J) COEFS.

0.33867168D-01	0.29369678D-01	-0.48749827D-01	0.99973376D-02	-0.22910924D-01
0.29369678D-01	0.21182724D-00	-0.13774729D-00	0.26656900D-01	-0.10719558D-00
-0.48749827D-01	-0.13774729D-00	0.28238392D-00	-0.13254110D-00	0.26656929D-01
0.99973376D-02	0.26656900D-01	-0.13254110D-00	0.28238414D-00	-0.13774744D-00
-0.22910924D-01	-0.10719558D-00	0.26656929D-01	-0.13774744D-00	0.21182734D-00

A=5 , S=1 W=ODD CASE

COEFFICIENTS (IN I.B.M. OUTPUT NOTATION) (SEE FLOW DIAGRAM)

INVERTED MATRIX K(N,P, J)

0.99999952D 00	-0.55901679D 01	0.12909941D 02	-0.15811386D 02	0.11180339D 02
-0.10416662D 02	0.56833372D 02	-0.12587192D 03	0.14098485D 03	-0.71740507D 02
0.36458316D 02	-0.19099740D 03	0.39536693D 03	-0.38869653D 03	0.16537583D 03
-0.52083312D 02	0.25621603D 03	-0.48412276D 03	0.42822498D 03	-0.16304658D 03
0.26041658D 02	-0.11646183D 03	0.20171781D 03	-0.16470191D 03	0.58230919D 02

R(I,J) COEFS.

0.20499202D-01	0.18254938D-01	-0.61369261D-02	-0.50107712D-02	0.91274658D-02
0.18254938D-01	0.15414872D-00	-0.86511780D-01	0.91340425D-01	-0.68032111D-01
-0.61369261D-02	-0.86511780D-01	0.26351671D-00	-0.19406634D-00	0.15512467D-00
-0.50107712D-02	0.91340425D-01	-0.19406634D-00	0.34274411D-00	-0.19729541D-00
0.91274658D-02	-0.68032111D-01	0.15512467D-00	-0.19729541D-00	0.25619696D-00

A=5 , S=1 W = EVEN CASE

COEFFICIENTS (IN I.B.M. OUTPUT NOTATION) (SEE FLOW DIAGRAM)

INVERTED MATRIX K(N,Q; J)

-0.71942586D-06	-0.62499966D 01	0.16666660D 02	-0.24999995D 02	0.24999998D 02	1.00000000D 00
-0.10416659D 02	0.63541629D 02	-0.16249993D 03	0.22291660D 03	-0.16041664D 03	0.74882195D-14
0.36458308D 02	-0.21354154D 03	0.51041642D 03	-0.61458310D 03	0.36979156D 03	-0.23780156D-13
-0.52083301D 02	0.28645818D 03	-0.62499969D 03	0.77708303D 03	-0.36458319D 03	0.29634844D-13
0.26041653D 02	-0.13020827D 03	0.26041653D 03	-0.26041654D 03	0.13020827D 03	-0.12649019D-13
0.10000000D 01	0.	-0.	0.	-0.	-0.10000000D 01

Q(I,J) COEFS.

0.68837404D-01	-0.23333691D-01	-0.30148645D-01	-0.86038454D-02	0.29792446D-01	-0.48735118D-01
-0.23333691D-01	0.21182724D-00	-0.13774729D-00	0.26656900D-01	-0.10719558D-00	0.52703368D-01
-0.30148645D-01	-0.13774729D-00	0.28238392D-00	-0.13254110D-00	0.26656929D-01	-0.18601182D-01
-0.86038454D-02	0.26656900D-01	-0.13254110D-00	0.28238414D-00	-0.13774744D-00	0.18601183D-01
0.29792446D-01	-0.10719558D-00	0.26656929D-01	-0.13774744D-00	0.21182734D-00	-0.52703371D-01
-0.48735118D-01	0.52703368D-01	-0.18601182D-01	0.18601183D-01	-0.52703371D-01	0.62500000D-01

CHAPTER 6

The wave-aberration derivatives.

It has been mentioned in Chapter 1 that, in optical design, it is advantageous to employ modern techniques of image assessment to give the criterion of image quality. This requires that the derivatives of the wave-aberration should be used. The formulae given below for $\partial W/\partial p$ were first published by Hopkins in an IBM Research Report[1] under the heading "Differential Correction Techniques Based on Optical Path Method", in which detailed proofs of the formulae were also given. Due to space limitation, no attempt will be made here to repeat these proofs, although brief accounts will be given for formulae of special importance. This chapter forms an introduction to the materials in the next two chapters.

(VI-i) The change in aberration with respect to the original reference sphere: surface change.

For any wavelength, not merely the mean wavelength $\bar{\lambda}$, the centre of the reference sphere in the image space is taken to lie in the paraxial image plane for $\bar{\lambda}$ light, and at the point where this plane is intersected by the pupil ray for this mean wavelength. This point will be denoted by $(\bar{Q}_o')_{\bar{\lambda}}$.

In Fig. 6.1, $\bar{Q}\,\bar{P}_1\bar{P}_s\bar{P}_k E'_\lambda$ is the pupil ray, for the wavelength λ, from the extra axial object point $\bar{Q}$, and passes through the exit pupil E'_λ for λ light. $E'_\lambda P'$ is a reference sphere centred at the point $(\bar{Q}_o')_{\bar{\lambda}}$. The aperture

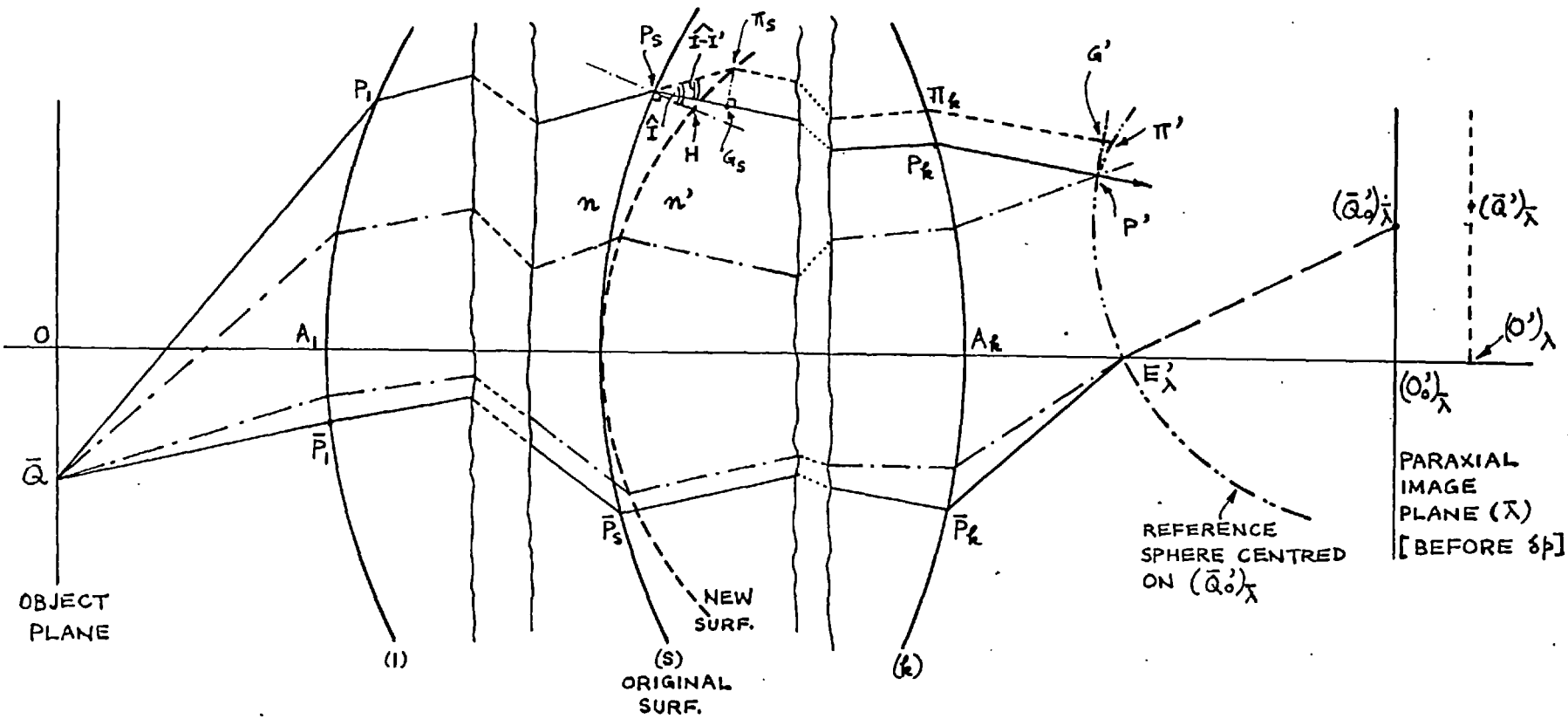

P_s
$\hat{I}-\hat{I}'$
π_s
P_1
$\hat{I}$
H
G_s
π_k
G'
π'
P_k
n
n'
P'
$(\bar{Q}_o')_{\bar{\lambda}}$
$(\bar{Q}')_{\bar{\lambda}}$
O
A_1
A_k
$E'_{\bar{\lambda}}$
$(O')_{\bar{\lambda}}$
$(O_o')_{\bar{\lambda}}$
$\bar{Q}$
$\bar{P}_1$
$\bar{P}_s$
$\bar{P}_k$
PARAXIAL IMAGE PLANE ($\bar{\lambda}$) [BEFORE δp]
REFERENCE SPHERE CENTRED ON $(\bar{Q}_o')_{\bar{\lambda}}$
OBJECT PLANE
NEW SURF.
(1)
(S) ORIGINAL SURF.
(k)

Fig. 6-1

ray $\bar{Q}P_1 \ldots P_k P'$ cuts the reference sphere at P', and the wavefront aberration along this ray is given by

$$W = [\bar{Q}\bar{P}_1 \ldots \bar{P}_k E'_\lambda] - [\bar{Q}P_1 \ldots P_k P'] \qquad (6.1)$$

where [ ] denotes optical path length. In (6.1) the aberration for any wavelength is referred to its own reference sphere, having its centre at $(\bar{Q}'_o)_\lambda$ but passing through E'_λ, the exit pupil point for the wavelength in question.

At any surface, a parameter change δp can be either a change in the vertex curvature δc, a change in eccentricity (if an aspheric surface) δQ or δA_{2m}, or a displacement δZ of the surface along the optical axis. Fig. 6.1 shows the case with $\delta p = \delta c$ at the s^{th} surface. The form of the new surface is defined, at any given point P_s by the distance $P_s H$:

$$P_s H = f.\delta p \qquad (6.2)$$

measured along the normal to the original surface at P_s. The factor f is simply the coefficient of proportionality in any given case.

In the modified system, the aperture ray follows the new path $P_s \pi_s \ldots . \pi_k \pi'$, cutting the original reference sphere at π', as shown by the path – – – –. The resulting change in optical path length along this ray relative to the original aperture ray is given by

$$\delta_o = [\bar{Q}P_1 \ldots P_s \pi_s \ldots \pi'] - [\bar{Q}P_1 \ldots P_s \ldots P']$$

which may be written as

$$\delta_o = [P_s\pi_s]+[\pi_s..G']+[G'\pi']-[P_sG_s]-[G_s...P'] \qquad (6.3)$$

where the point G' is such that G'P' is the perpendicular at P' to the original ray, and G_s is similarly the foot of the perpendicular from π_s on the original ray.

By Fermat's Theorem

$$[\pi_s....G'] = [G_s...P'] + \underline{0}\ (\delta p^2) \qquad (6.4)$$

and, if p' has pupil co-ordinates (x'_s, y'_s) and π' has co-ordinates $(x'_s+\delta x'_s, y'_s+\delta y'_s)$,

$$[G'\pi'] = -\left[\frac{\partial W}{\partial x'_s}.\delta x'_s + \frac{\partial W}{\partial y'_s}.\delta y'_s\right] \qquad (6.5)$$

Now the reduced transverse ray aberration formulae (2.26) are

$$\delta G'_s = -\frac{\partial W}{\partial x'_s} \qquad \delta H'_s = -\frac{\partial W}{\partial y'_s}$$

so that (6.5) may be written as

$$[G'\pi'] = +(\delta G'_s.\delta x'_s + \delta H'_s.\delta y'_s) \qquad (6.6)$$

Using this and (6.4),(6.3) becomes

$$\delta_o = [P_s\pi_s]-[P_sG_s]+(\delta G'_s.\delta x'_s + \delta H'_s.\delta y'_s)+\underline{0}(\delta p^2) \qquad (6.7)$$

From Fig. 6.1, for small values of P_sH we have

$$[P_s\pi_s] = n\left(\frac{P_sH}{\cos I}\right) = \frac{n\ f.\delta p}{\cos I} \qquad (6.8)$$

$$[P_sG_s] = n'(P_s\pi_s)\cos(I-I')$$

where I and I' are the angles of incidence and refraction of the original aperture ray at the original surface. Substituting (6.8) in (6.7) and simplifying, we have

$$\delta_o = -f.\delta p\Delta\,(n\cos I)+(\delta G'_s.\delta x'_s+\delta H'_s.\delta y'_s) + \underline{0}(\delta p^2) \qquad (6.9)$$

and, similarly for the pupil ray,

$$\bar{\delta}_o = -\bar{f}.\delta p\ \Delta(n\cos\bar{I})+\delta\bar{H}'_s\delta\bar{y}'_s \qquad (6.10)$$

since the transverse aberration $\delta\bar{G}'_s$ for the pupil ray is zero. For the mean wavelength, $\bar{\lambda}$, also $\delta\bar{H}'_s = 0$. For any other wavelength, $\delta\bar{H}'_s$ is the (reduced) transverse aberration of the pupil ray image for this wavelength relative to that for $\bar{\lambda}$; and $\delta\bar{y}'_s$ is the value of $\delta y'_s$ for $y'_s = 0$.

From (6.1), it can be seen that the increment in aberration δW_o due to δp is

$$\delta W_o = \bar{\delta}_o - \delta_o \qquad (6.11)$$

and from (6.9) and (6.10) we have

$$\delta W_o = \{f\Delta(n\cos I)-\bar{f}\Delta(n\cos\bar{I})\}\delta p$$

$$- \{\delta G'_s.\delta x'_s+\delta H'_s.\delta y'_s\}+\delta\bar{H}'_s.\delta\bar{y}'_s \qquad (6.12)$$

the quantity $\mathcal{O}(\delta p^2)$ in (6.9) being neglected. The derivative of the wave aberration is, from (6.12),

$$\frac{\partial W_o}{\partial p} = f.\Delta(n\ \cos I) - \bar{f}\Delta(n\ \cos \bar{I}) - \{\delta G'_s . \frac{\partial x'_s}{\partial p} + \delta H'_s . \frac{\partial y'_s}{\partial p} + \delta \bar{H}'_s \frac{\partial \bar{y}'_s}{\partial p} \qquad (6.13)$$

where $\frac{\partial \bar{y}'_s}{\partial p}$ is the value of $\frac{\partial y'_s}{\partial p}$ for the pupil ray, that is when $y'_s = o$. For $\lambda = \bar{\lambda}$, $\delta \bar{H}'_s = o$ and the final term is absent from (6.13).

The formula (6.13) has been derived by assuming the 'entrance pupil associated ray' is the correct choice, i.e. the rays considered above are taken to be unmodified in the object space when the system undergoes a parameter change. In certain cases, it is necessary to consider the ray corresponding to the original ray to be that ray, which after passing through the modified system, intersects the original exit pupil reference sphere at P', as shown by -·-·- in Fig.6.1. This 'exit pupil associated ray' will therefore lead to a formula $\frac{\partial W_i}{\partial p}$ similar to (6.13) but with the terms $\delta G'_s . \frac{\partial x'_s}{\partial p} + \delta H'_s . \frac{\partial y'_s}{\partial p}$ absent, so that

$$\frac{\partial W_i}{\partial p} = f\Delta(n\cos I) - \bar{f}\Delta(n\cos \bar{I}) \qquad (6.14)$$

By substituting (6.14) in (6.13), the expression for $\frac{\partial W_o}{\partial p}$ may be written as

$$\frac{\partial W_o}{\partial p} = \frac{\partial W_i}{\partial p} - \left(\delta G'_s \cdot \frac{\partial x'_s}{\partial p} + \delta H'_s \cdot \frac{\partial y'_s}{\partial p} \right) + \delta \bar{H}'_s \frac{\partial \bar{y}'_s}{\partial p} \qquad (6.15)$$

It may be shown that the formulae for the parameter derivative of ray position at the exit pupil may be given the forms

$$\frac{\partial x'_s}{\partial p} = \left(\frac{\bar{N}'^2}{\bar{H}'_T} \right) \frac{\partial W_{T|}}{\partial p} \cdot x'_s \qquad (6.16)$$

$$\frac{\partial y'_s}{\partial p} = \left(\frac{\bar{H}'_T}{\bar{H}_s H_T} \right) \cdot \frac{\partial \bar{W}_{T|}}{\partial p} + \left(\frac{\bar{N}'^2}{\bar{H}'_T} \right) \frac{\partial W_{T|}}{\partial p} \; y'_s$$

where $\bar{H}'_T$ is defined to be

$$\bar{H}'_T = \bar{N}'_k \, n'_k \, h'_T \, \bar{M}'_k \qquad (6.17)$$

h'_T being the incident height along the reference sphere at E' of a traced close tangential ray (T-ray) along the reference ray, and

$$H_T = n u_T \eta_T = n' u'_T \eta'_T \qquad (6.18)$$

is an invariant quantity for this T-ray. The formulae for $\frac{\partial W_{T|}}{\partial p}$ and $\frac{\partial \bar{W}_{T|}}{\partial p}$ for different parameter, p, will be given later.

The change in any constructional parameter, keeping the entrance pupil fixed, will result in a shift in the exit pupil E'_λ and a change in the pupil magnification. It follows that the entrance - pupil associated ray passing through the new system will intersect the original reference sphere at a different point from the original ray passing through the original system. The formula for $\partial y'_s / \partial p$ in (6.16) has the first term giving the shift of exit pupil in its exact form; the second term assumes the scale change in the T-section to be the same as in the S-section.

(VI-ii) Focal shift derivatives

For any wavelength the parameter derivatives of the wave aberration are found with the aberration referred to the original reference sphere centred at $(\bar{Q}'_o)_{\bar{\lambda}}$. However the effect of a parameter change will also be to shift the paraxial image plane for $\bar{\lambda}$, and to change the pupil ray image height for this wavelength (the new image point is now at $(\bar{Q}')_{\bar{\lambda}}$ in Fig. 6.1). The focal shift terms for $\bar{\lambda}$, both longitudinal and transverse, are regarded as effectively constructional parameters. If these parameters are specified by $(\delta_o W_{20})_{\bar{\lambda}}$ for the shift of the image plane, and by $(\delta W_{11})_{\bar{\lambda}}$ for each extra-axial image point, then the change in aberration for a single parameter change referred to the current paraxial focal plane and pupil ray image height for $\lambda=\bar{\lambda}$ is given by

$$\left(\frac{\partial W}{\partial p}\right)_{\lambda} = \frac{\partial W_o}{\partial p} + \left(\frac{\partial W}{\partial({}_oW_{20})_{\bar{\lambda}}}\right)\cdot\left(\frac{\partial {}_oW_{20}}{\partial p}\right)_{\bar{\lambda}} + \left(\frac{\partial W}{\partial(W_{11})_{\bar{\lambda}}}\right)\cdot\left(\frac{\partial W_{11}}{\partial p}\right)_{\bar{\lambda}} \qquad (6.19)$$

and for exit pupil associated ray, $\frac{\partial W_i}{\partial p}$ replaces $\frac{\partial W_o}{\partial p}$ in (6.19).

The effect of a set of parameter changes will be to change the aberration along a given ray by an amount

$$\delta W = \sum\left(\frac{\partial W_o}{\partial p}\cdot\delta p\right) + \left(\frac{\partial W}{\partial({}_oW_{20})_{\bar{\lambda}}}\right)\cdot(\delta {}_oW_{20})_{\bar{\lambda}} + \left(\frac{\partial W}{\partial(W_{11})_{\bar{\lambda}}}\right)\cdot(\delta W_{11})_{\bar{\lambda}} \qquad (6.20)$$

and this holds for any wavelength for which the various derivative are found. The formulae for $\partial W/\partial({}_oW_{20})_{\bar{\lambda}}$ and $\partial W/\partial(W_{11})_{\bar{\lambda}}$ will be given in Chapter 7, and very briefly, they are obtained in the following way :

The effect of shifting the centre of the reference sphere from $(\bar{Q}'_o)_{\bar{\lambda}}$ to $(\bar{Q}')_{\bar{\lambda}}$ (see Fig. 6.1) is to introduce an aberration change δW_{λ}. Referring to Fig. 6.2, the exit pupil points for λ and $\bar{\lambda}$ are at E'_{λ} and $E'_{\bar{\lambda}}$ respectively. δW_{λ} may be regarded as resulting from a longitudinal focal

shift from $(\bar{Q}'_o)_{\bar{\lambda}}$ to $\bar{Q}'$, and then a transverse focal shift from $\bar{Q}'$ to $(\bar{Q}')_{\bar{\lambda}}$, and these may easily be calculated. By expressing the results in terms of the longitudinal focal shift for $\bar{\lambda}$, i.e. from $(0'_o)_{\bar{\lambda}}$ to $(0')_{\bar{\lambda}}$ as measured by the coefficient $\delta({}_0W_{20})_{\bar{\lambda}}$, the formula for $\partial W/\partial({}_0W_{20})_{\bar{\lambda}}$ may be obtained.

For the formula $\partial W/\partial(W_{11})_{\bar{\lambda}}$, one considers the centre of reference sphere $E'_\lambda P'$ (Fig. 6.3) shifting from $(\bar{Q}'_o)_{\bar{\lambda}}$ to a new position $(\bar{Q}')_{\bar{\lambda}}$ in the same image plane, and the effect of this is to introduce a change in aberration δW_λ for the aperture ray through P'. δW_λ may be calculated in terms of the transverse focal shift $\delta(W_{11})_{\bar{\lambda}}$ for $\bar{\lambda}$, leading to the expression for $\partial W/\partial(W_{11})_{\bar{\lambda}}$.

(VI-iii) Parameter derivatives for curvature change.

It will be seen that the six quantities :

$$\frac{\partial W_i}{\partial p}\ ,\quad \frac{\partial_o W_{20}}{\partial p}\ ,\quad \frac{\partial_o \bar{W}_{20}}{\partial p}\ ,\quad \frac{\partial \bar{W}_{T1}}{\partial p}\ ,\quad \frac{\partial_1 W_{11}}{\partial p}\ ,\quad \frac{\partial \bar{W}_{T1}}{\partial p}$$

are those required for a complete account of the differential properties of a system. The formulae for these will be given in Chapter 7. Examples of how the above six quantities are calculated for the case $\delta p = \delta c$ are now given below :

Suppose a ray is incident at a point

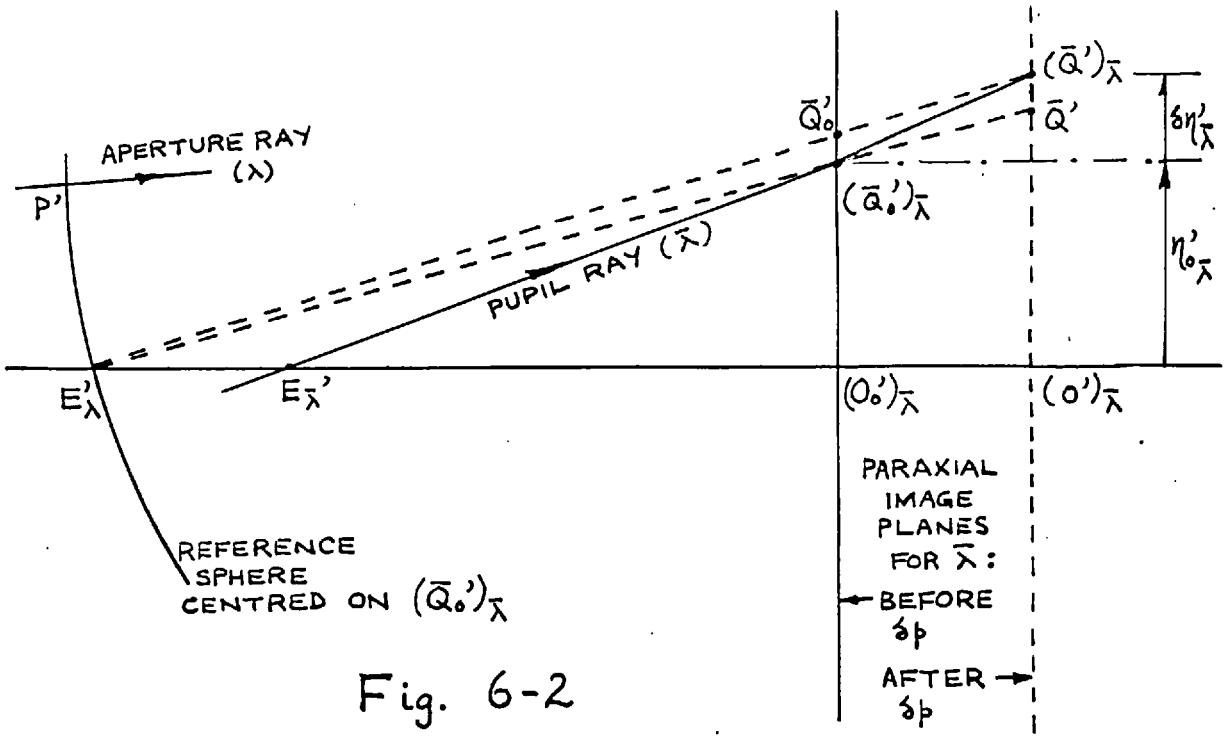

APERTURE RAY
(λ)
P'
PUPIL RAY (λ̄)
$\bar{Q}_o'$
$(\bar{Q}')_{\bar{\lambda}}$
$\bar{Q}'$
$\delta\eta'_{\bar{\lambda}}$
$(\bar{Q}_o')_{\bar{\lambda}}$
$\eta'_{o\bar{\lambda}}$
E'_λ
$E_{\bar{\lambda}}'$
$(O_o')_{\bar{\lambda}}$
$(o')_{\bar{\lambda}}$
PARAXIAL IMAGE PLANES FOR λ̄:
BEFORE δp
AFTER δp
REFERENCE SPHERE CENTRED ON $(\bar{Q}_o')_{\bar{\lambda}}$

Fig. 6-2

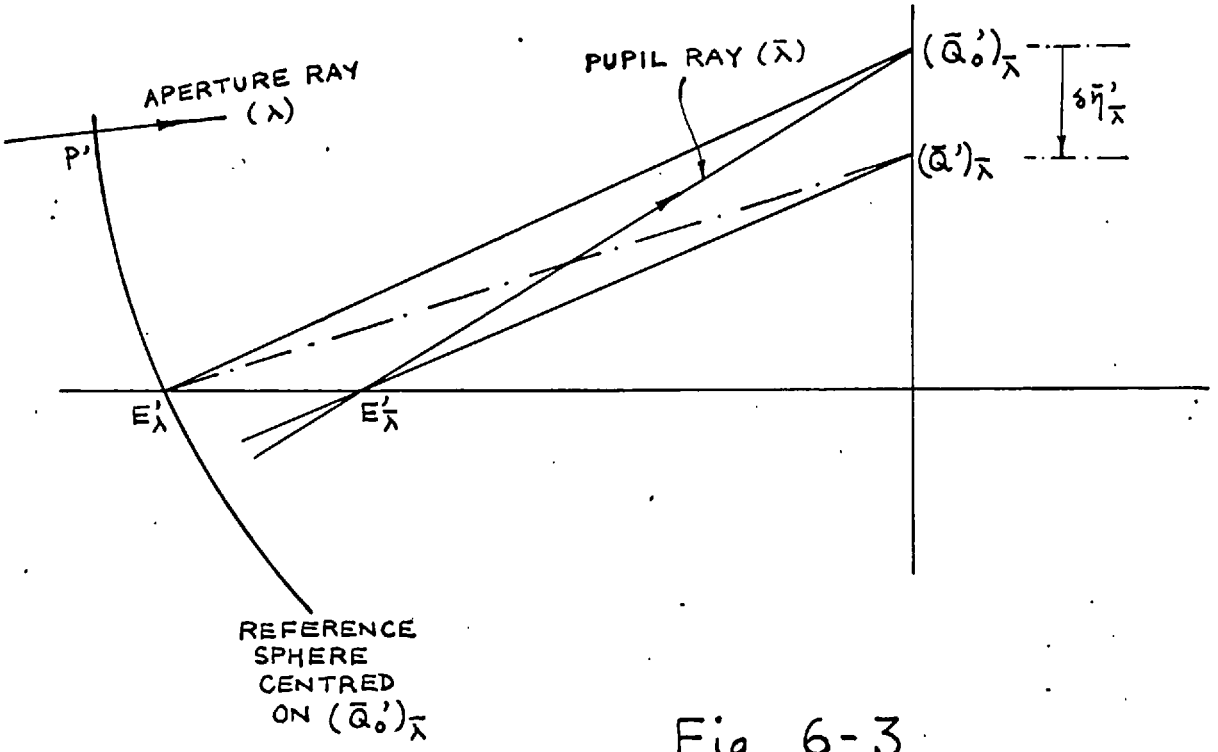

APERTURE RAY
(λ)
P'
PUPIL RAY (λ̄)
$(\bar{Q}_o')_{\bar{\lambda}}$
$\delta\bar{\eta}'_{\bar{\lambda}}$
$(\bar{Q}')_{\bar{\lambda}}$
E'_λ
$E'_{\bar{\lambda}}$
REFERENCE SPHERE CENTRED ON $(\bar{Q}_o')_{\bar{\lambda}}$

Fig. 6-3

$P_s(x,y,z)$ at a spherical interface AP_s (Fig. 6.4) of centre c_o and curvature $c = 1/AC_o$. If the curvature of the surface is changed to $(c+\delta c)= 1/AC_1$, the intercept P_sH (see Fig. 6.4 and Fig. 6.1) is given by

$$(P_sH) = \tfrac{1}{2}(x^2+y^2+z^2).\delta c + \underline{0}(\delta c^2)$$

and by the definition (6.2), we have

$$\tfrac{1}{2}(x^2+y^2+z^2).\delta c = f.\ \delta c$$

or, cancelling δc,

$$f = \tfrac{1}{2}(x^2+y^2+z^2) \qquad (6.21)$$

Hence by (6.14),

$$\frac{\partial W_i}{\partial c} = \tfrac{1}{2}(x^2+y^2+z^2).\Delta(n\cos I)-\tfrac{1}{2}(\bar{x}^2+\bar{y}^2+\bar{z}^2)\Delta(n\cos\bar{I}) \qquad (6.22)$$

and $\bar{x} = 0$ for pupil ray.

For the axial pencil $\bar{y}=\bar{z}=o$, and also $x=o$, so that (6.22) reduces to

$$\frac{\partial W_i}{\partial c} = \tfrac{1}{2}(y^2+z^2).\Delta(n\cos I) \qquad (6.23)$$

whose paraxial form gives

$$\frac{\partial_o W_{20}}{\partial c} = \tfrac{1}{2}.h^2.\Delta(n) \qquad (6.24)$$

where h is the paraxial ray incidence height at the surface in question. Similarly

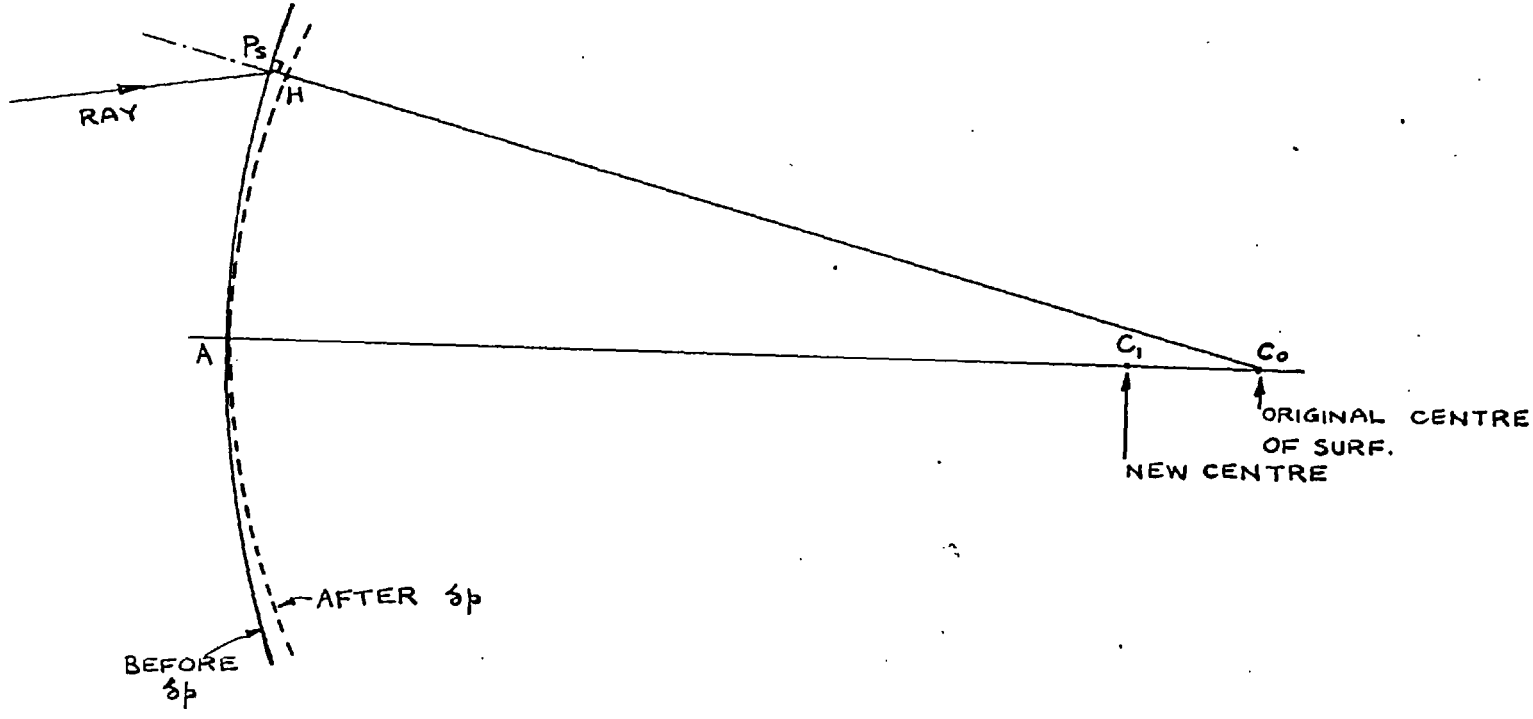

P_s
H
RAY
A
C_1
C_o
ORIGINAL CENTRE
OF SURF.
NEW CENTRE
AFTER δp
BEFORE
δp

Fig. 6-4

$$\frac{\partial_o \bar{W}_{20}}{\partial c} = \tfrac{1}{2} \cdot \bar{h}^2 \cdot \Delta(n) \qquad (6.25)$$

where $\bar{h}$ is the paraxial pupil ray height.

To find $\partial W_{T1}/\partial c$, a limitingly close tangential ray is traced. If, for such a ray, $f = f+\delta f$ and $I = \bar{I}+\delta I$, (6.22) gives

$$\frac{\partial W_{T1}}{\partial c} = \delta f \Delta (n\cos\bar{I}) - \bar{f} (n\sin\bar{I})\Delta(\delta I) \qquad (6.26)$$

and this may be shown to give

$$\frac{\partial W_{T1}}{\partial c} = \bar{y}h_T\Delta(n\cos\bar{I})+\tfrac{1}{2}(\bar{y}^2+ \bar{z}^2)(n\sin\bar{I})\Delta(u_T) \qquad (6.27)$$

where (h_T,u_T) are obtained from the same paraxial tangential ray whose aperture variables are used in the calculation of $\bar{H}'_T$ in (6.17).

The formula for $\partial\bar{W}_{T1}/\partial c$ is obtained from (6.27) by replacing (h_T,u_T) by $(\bar{h}_T,\bar{u}_T)$, the latter are obtained from a $\bar{T}$-ray which is a paraxial tangential ray traced backwards from the exit pupil point E'_λ (see Chapter 7).

The paraxial form of (6.27) is

$$\frac{\partial_1 W_{11}}{\partial c} = \bar{h}h \cdot \Delta(n) \cdot \tau \qquad (6.28)$$

since the paraxial limit gives $\bar{y}=\bar{h}.\tau$, $h_T=h$, $\Delta(n\cos\bar{I}) =\Delta(n)$ and $\bar{z} = o.\ \tau$ is the fractional object height as defined earlier.

The formula for $\partial_1 W_{11}/\partial c$ is symmetrical in h and $\bar{h}$, so that

$$\frac{\partial_1 \bar{W}_{11}}{\partial c} = \frac{\partial_1 W_{11}}{\partial c} \qquad (6.29)$$

and this is found to be so for any parameter. Accordingly, the derivative $\frac{\partial_1 \bar{W}_{11}}{\partial p}$ never needs to be calculated separately.

(IV-iv) Distortion and variation in Gaussian image height. (required for $\bar{\lambda}$ only)

To find the change in distortion resulting from a parameter change, it is convenient to find the change in the ratio of $\bar{\eta}'/\eta'$, where $\bar{\eta}'$ is the finite ray image height and η' is the gaussian value of the image height. This follows from :

$$\delta\left(\frac{\delta\bar{\eta}'}{\eta'}\right) = \left(\frac{\bar{\eta}'-\eta'}{\eta'}\right) - \left(\frac{\bar{\eta}_o'-\eta_o'}{\eta_o'}\right) = \delta\left(\frac{\bar{\eta}'}{\eta'}\right)$$

the subscript zero refers to the original system. When a parameter change occurs, the pupil ray in the image space will have its direction cosines changed by the amounts $(\delta\bar{M},\delta\bar{N})$. By relating these quantities to the change in image height, one obtains the parameter derivative of the fractional distortion as :

$$\frac{\partial}{\partial p}\left(\frac{\delta\bar{\eta}'}{\eta'}\right)_{\bar{\lambda}} = \frac{\bar{H}_o'}{\bar{H}_o}\left\{-\frac{1}{\bar{H}_T'}\cdot\frac{\partial \bar{W}_{T1}}{\partial p} + \frac{1}{\bar{H}}\frac{\partial_1 \bar{W}_{11}}{\partial p}\right\} \qquad (6.30)$$

where $\bar{H}_T'$ is that given in (6.17), and $\bar{H}_o$, $\bar{H}_o'$ and H are obtained from (4.19) and (4.20).

Finally, an expression for the change in gaussian image height, $\delta\eta_{\bar{\lambda}}'$ (Fig. 6.2), due to δp is required. Thus by considering $\delta\bar{\eta}_{\bar{\lambda}}'$ as a result of the longitudinal focal shift

from $(\bar{Q}'_o)_{\bar{\lambda}}$ to $\bar{Q}'$, together with the transverse focal shift from $\bar{Q}'$ to $(\bar{Q}')_{\bar{\lambda}}$, and using paraxial approximation, the derivative of the gaussian image height w.r.t. any parameter p may be shown to be given by

$$\frac{1}{\bar{\eta}'} \frac{\partial \eta'}{\partial p} = \frac{2}{H} \left(\frac{\bar{u}k'}{u'_k} \right) \frac{\partial_o W_{20}}{\partial p} - \frac{1}{H} \frac{\partial_1 W_{11}}{\partial p} \qquad (6.31)$$

For small values of u'_k, the variation in angular size of the image is used instead of $\delta\eta'_{\bar{\lambda}}$. In this case, (6.31) is replaced by

$$\frac{1}{\bar{u}'} \cdot \frac{\partial \bar{u}'}{\partial p} = - \frac{1}{H} \frac{\partial_1 W_{11}}{\partial p} \qquad (6.32)$$

Note that (6.30), (6.31) and (6.32) are only required for $\lambda = \bar{\lambda}$.

REFERENCES

(1) H.H. Hopkins, Optical Calculations V, I.B.M. Research Report (1963)

H.H. Hopkins, Summer School Notes, Imperial College, (1965)

CHAPTER 7

THE CALCULATION OF THE WAVE-ABERRATIONS AND THE ABERRATION DERIVATIVES

General Introduction

A computer programme for computing both the wave aberration, W, and its parameter derivatives, $\partial W/\partial p$, has been written for the I.B.M. 7094 computer. The programme consists of a total of eight subroutine-subprogrammes. These are :

1. Input subroutine - this is concerned with the system data, the ray data and the pupil data.
2. Gaussian Optics subroutine - this deals with the paraxial aperture and pupil ray traces, the primary aberrations and the computations of the derivatives

$$\partial_1 W_{11} / \partial p \ , \quad \partial_o W_{20} / \partial p \ , \quad \partial_o \bar{W}_{20} / \partial p$$

and the derivatives of the image height

$$\frac{1}{\eta} \cdot \frac{\partial \eta}{\partial p} \ \text{(or } \ \frac{1}{\bar{\mu}'} , \ \frac{\partial \bar{\mu}'}{\partial p} \ \text{, see equation 6.32)}$$

3. Finite-ray trace subroutine - this computes and stores the ray data for the pupil ray, axial ray and the extra-axial ray subroutines.
4. Pupil-ray subroutine - here the effective pupil rays are traced for each of the different image heights. The parameters (a,b,c) of the elliptical pupil, which are obtained from the pupil exploration (see Chapter 4), are required. Various

quantities associated with the pupil ray (e.g. $\bar{H}_s$, $\bar{H}_o$, distortion etc.) are computed and stored.

5. T and $\bar{T}$ rays subroutine - here the T and $\bar{T}$ rays are traced using the pupil ray data from the above subroutine. The parameter derivatives $\partial W_{T1}/\partial p$, $\partial \bar{W}_{T1}/\partial p$ are calculated and stored.

6. Axial-ray subroutine (τ=o) - this deals with the starting data for the axial ray trace. Rays of the axial pencil are traced and the values of W, $\partial W/\partial p$ are found for each of the rays and for the desired parameters, p.

7. Extra-axial ray subroutine. ($\tau \neq$o) - this deals with the starting data for the extra-axial ray trace, making use of the data from the pupil exploration. It also computes W, $\partial W/\partial p$ associated with the rays of extra-axial pencils.

8. Output subroutine - this prints out the values of W, $\partial W/\partial p$ as well as the input data (intermediate data may also be printed out when desired).

The dimension statements in the programme impose the following limitations:

maximum number of surfaces = 30

" " " figured surfaces = 5

" " " fields for the extra-axial images = 5

" " " wavelengths = 5

" " " rays (2A+S) = 13 (see Chapter 5)

The programme has, however, been so designed that, if at any

time, when larger quantities than those specified above are to be accommodated, one simply modifies the dimension statements accordingly; the programme will remain valid.

The computation of W, $\partial W/\partial p$ and the associated quantities demands working and storage spaces which greatly exceed the available magnetic core units of the I.B.M. 7094 computer (32,768 word locations of 36 information bits each). Consequently eight magnetic tapes are called for; these are Tapes 1,2,3,4,19, 20, 21 and 22.

In what follows, the formulae involved in each subroutine are listed separately and precede the corresponding flow-diagram. In each case, explanatory notes concerning the subroutine are also presented.

(VII-i) Input stage (see also flow diagram VII-i)

The following code names are used in the input subroutine:-

NOWAVE = number of different wavelengths to be considered.

NOSURF = total number of surfaces, including the stop (which is counted as a surface).

NOSTOP = surface number (as counted from the left) corresponding to the stop position.

MQDRIC = maximum number of quadric surfaces (with $Q \neq o$).

MFIGRD = maximum number of figured surfaces.

NOFIGR = counter for the figured surfaces.

FIGRED(K) = o if the K^{th} surface is not figured.

= 1 if the K^{th} surface is figured.

FIGCOF = figuring coefficient.

NFIELD = number of extra-axial image points.

NCODE 2 = 0 if the angular size of the object and image are to be considered.

= 1 for a telescopic system.

= 2 if the linear size of the object and image are to be considered.

A = number of "A" rays (see ray pattern, Chapter 5, Section V-i) for the mean wavelength.

S = number of skew rays for the mean wavelength.

T = number of skew rays for the wavelengths other than the mean.

B = number of "A" rays for the wavelengths other than the mean.

E(K) = refractive index of the K^{th} medium.

IDATA = o if no intermediate data are required to be printed.

= 1 print intermediate data.

HTMAX(K)= value of the normalising radius at the figured surface K.

INDEX = counter for quantities depending on both the surface no. and the wavelength.

JB(i) = see below.

The code names B and T are introduced to give a more flexible ray pattern; thus when $\lambda \neq \bar{\lambda}$, the number of "A" rays to be traced (=B) may be less than or equal to that traced for $\bar{\lambda}$; similarly, $T \leqslant S$. In practice, each selected ray for $\lambda \neq \bar{\lambda}$

will be traced with the same entrance pupil co-ordinates as the corresponding ray for $\lambda = \bar{\lambda}$ - see flow diagrams (VII-vi) and (VII-vii) later. It is convenient here to describe this in detail :

The ray pattern to be used for tracing the $\bar{\lambda}$ rays is that described in Chapter 5, section (V-i). Referring to the particular ray pattern (for $\bar{\lambda}$) given in Fig. 5-1, with A=3 (j=1,2,3) and S=3 (the skew rays are numbered with j= 4,5 and 9); if, for $\lambda \neq \bar{\lambda}$, B=2 rays are to be traced and the ray co-ordinates selected are those corresponding to j=1 and j=3, then we employ subscripted variables JB(i), with i=1,2 such that

JB(1) = 1 (corresponding to j=1)

JB(2) = 3 (corresponding to j=3)

are stored at the input stage. The aperture rays for $\lambda \neq \bar{\lambda}$ will later be traced with the pupil co-ordinates given by

$$\left.\begin{aligned} y &= \pm\sqrt{\frac{(A+1)-j}{A}} \\ x &= o \end{aligned}\right\} \quad j = JB(1) \text{ and } JB(2)$$

(see equations 5-4 and 5-9)

In general, each of the B rays will have its selected ray number stored in JB(i), with i=1,2,...,B, at the input stage, and a total of 2B aperture rays will be traced for each $\lambda \neq \bar{\lambda}$ and field point.

As mentioned above, Fig. 5.1 shows the ray pattern for the skew rays corresponding to the case S=3, j=4,5,9 for $\lambda = \bar{\lambda}$. If,

for $\lambda \neq \bar{\lambda}$, T =1 ray is to be traced, then its pupil ray co-ordinates must be (x=1, y=o), corresponding to j=5. If T=2, then two skew rays are to be traced for $\lambda \neq \bar{\lambda}$, corresponding to j=4 and 9, with the pupil co-ordinates given by

$$\left.\begin{aligned} y &= \pm\sqrt{\frac{A-1}{2A}} \\ x &= +\sqrt{\frac{A-1}{2A}} \end{aligned}\right\} \text{ see equations (5.2),(5.5) and (5.6).}$$

If T=3, then three skew rays are to be traced for $\lambda \neq \bar{\lambda}$, corresponding to j=4,5,9, the same as that for $\bar{\lambda}$.

The notation used for the system data, such as C,D,E, δn, ϕ_{max}, and for the ray data such as $h_1, \bar{h}_1, u_1, \bar{u}_1$ etc. are as previously defined in Chapter 4.

In the subsequent flow-diagrams, the wavelength cycles will always have that for $\bar{\lambda}$ preceeding that for λ. Referring to the read-in cycle for E and δn in the flow-diagram (VII-i), the values of E(1) → E(NOSURF+1) and $\delta n(1) \rightarrow \delta n$ (NOSURF + 1) are reserved for $\bar{\lambda}$; E (NOSURF+2) → E(2xNOSURF+2) and δn(NOSURF+2) → δn(2xNOSURF + 2) are for the next wavelength, etc. The counter used for the subscripted variables E and δn is INDEX (see part B, flow-diagram VII-i).

The read-in cycle for the figuring coefficients has NOFIGR as the counter and its value is initially set to unity; when the first figured surface has been detected (i.e. when FIGRED(K)=1 for the K^{th} surface) its figuring coefficients

FIGCOF(M) must be read-in - see part A, flow-diagram (VII-i), where M has the values 4,3,2 and 1, corresponding to the figuring coefficients A_4, A_6, A_8 and A_{10} in that order. At the end of the cycle J, the counter NOFIGR is increased by unity so that the value of M for the next figured surface will be correct.

The values of the fractional image height (τ) are read in via the cycle L, part C in the flow-diagram (VII-i). For each τ value, the values of (a,b,c) for the elliptical pupil corresponding to each wavelength are also required; these are stored as $a_{pupil}(N)$, $b_{pupil}(N)$ and $c_{pupil}(N)$, where N is a counter which depends on both λ and τ.

The surface cycle K has K=1,2,3,...,NOSURF +1; the last quantity refers to the image space (i.e. the medium to the right of the last surface). An example of the print-out of the input data is presented at the end of this chapter.

FLOW-DIAGRAM (VII-i)

Input Subroutine.

Rewind all tapes (NOS. 1, 2, 3, 4, 19, 20, 21, 22)

Read : IDATA, NOWAVE, NOSURF, NOSTOP, MQDRIC, MFIGRD.

Cycle I (I = 1, 2, ..., NOWAVE)
Read λ(I) — i.e. the values of different λ's in A.
Repeat I

Set NOFIGR = 1

Cycle K (K = 1, 2, ..., NOSURF+1)

K = NOSURF+1 ?
Y → P.T.O.
N →

Read C(K), f_{max}(K), Q(K), FIGRED(K)

K = 1 ?
Y →
N → Read D(K) (= axial separation)

FIGRED(K) = 1 ?
N →
Y →

PART "A"
Read HTMAX(K)
Cycle J (J = 1, 2, 3, 4)
M = 4 × NOFIGR − J + 1
Read FIGCOF(M)
Repeat J
Set NOFIGR = NOFIGR + 1

P.T.O.

PART "B"

Cycle I (I = 1, 2, ···, NOWAVE)
Set INDEX = (I − 1) . (NOSURF + 1) + K
Read E(INDEX), δn(INDEX)
Repeat I

Repeat K (next surface)

Read NFIELD

PART "C"

Cycle L (L = 1, 2, ···, NFIELD)
Read τ(L)
Cycle M (M = 1, 2, ···· NOWAVE)
Set N = M + (L − 1) . NOWAVE
Read
$a_{pupil}(N)$; $b_{pupil}(N)$; $c_{pupil}(N)$
Repeat M
Repeat L

Read h_1, u_1, $\bar{h}_1$, $\bar{u}_1$, A, B, S, T, NCODE2

Cycle i (i = 1, 2, ···, B)
Read JB(i)
Repeat i

Call the following subroutines (in this order):

(i) Gaussian Optics and Distortion
(ii) Pupil Rays
(iii) T and $\bar{T}$ Rays
(iv) Axial Rays
(v) Extra-axial Rays
(vi) Out-put.

(VII-ii) Gaussian Optics, Distortion and Primary Aberrations
(see also flow-diagram VII-ii)

For each wavelength, a paraxial ray is traced from the axial object point with an aperture angle defined by $\mu_1 = \sin \alpha_1$, where α_1 is the angular semi-aperture; a paraxial pupil ray is also traced for the maximum object height (τ=1), using $\bar{u}_1 = \tan\bar{\alpha}_1$, where $\bar{\alpha}_1$ is the angular semi-field. The paraxial quantities $(h,u),(\bar{h},\bar{u}),i,i',\bar{i},\bar{i}'$ at each surface for the current wavelength are required for the calculation of the parameter derivatives at each surface (each K value). As mentioned in Chapter 6, the parameter derivatives of the Gaussian image height (i.e. $\frac{1}{H}\frac{\partial H}{\partial p}$) are only required for the mean wavelength. The formulae for $\frac{1}{\bar{H}}\frac{\partial H}{\partial p}$ corresponding to the linear-size case (see equation 6.31) require the values of $\bar{u}'$ and u' at the last surface; for this particular case, the cycle K1 is employed - see part A, flow-diagram (VII-ii). At the end of each wavelength cycle (i.e. for each value of I) the paraxial parameter derivatives corresponding to each surface are stored on the magnetic tapes (1),(2),(3) and (20). The different tapes are employed for the following purposes:
(a) Tape (1) - the paraxial data stored on this tape will be required for the axial-ray and extra-axial ray subroutines.
(b) Tape (2) - data stored on this tape will be augmented with additional quantities later in the T-ray subroutine; the purpose of this tape is for storing information for the print-out, to be called by the out-put subroutine.

(c) Tape (3) - this stores the $\bar{\lambda}$ quantities, and also additional data from the pupil ray subroutine later. This tape will be called by the T-ray subroutine.

(d) Tape (20) - the paraxial quantities stored on this tape, plus additional information to be recorded later, constitute the input data to the optimisation stage (see Chapter 8).

The use of the tapes in the manner indicated in the flow-diagram (VII-ii) result in economising the dimension of the subscripted variables; thus, in the dimension statement, the maximum size of the arrays need not exceed the maximum number of surfaces involved, regardless of the number of wavelengths to be considered.

For the mean wavelength, the following quantities are also calculated :

1. the total throw (i.e. the object to image distance).
2. the magnification (linear or angular, depending on the case considered).
3. focal lengths (f and f').
4. the positions of the principal planes P and P' as measured by the distances δ and δ' from the first and last surfaces respectively.
5. The primary aberrations (see formulae VII-ii).

FORMULAE (VII - ii)

The following formulae are required for the Flow Diagram (VII-ii). The subscripts 1 and k refer to the first and the last surfaces respectively

$$H = n_1 [\bar{h} u - \bar{u} h]_1 \quad ; \quad A = (n i) \quad ; \quad B = (n \bar{i})$$

Paraxial parameter derivatives :

$p = c$

$$\frac{\partial_0 W_{20}}{\partial c} = \frac{1}{2} \cdot h^2 \Delta(n)$$

$$\frac{\partial_1 W_{11}}{\partial c} = \bar{h} h \Delta(n)$$

$$\frac{\partial_0 \bar{W}_{20}}{\partial c} = \frac{1}{2} \cdot \bar{h}^2 \Delta(n)$$

$p = z$

$$\frac{\partial_0 W_{20}}{\partial z} = -\frac{1}{2} \Delta(n u^2)$$

$$\frac{\partial_1 W_{11}}{\partial z} = -\Delta(n \bar{u} u)$$

$$\frac{\partial_0 \bar{W}_{20}}{\partial z} = -\frac{1}{2} \Delta(n \bar{u}^2)$$

$p = \mu$

$$\frac{\partial_0 W_{20}}{\partial \mu} = \frac{1}{2} [(h i') - (h i)_{+1}]$$

$$\frac{\partial_1 W_{11}}{\partial \mu} = (h \bar{i}') - (h \bar{i})_{+1}$$

$$\frac{\partial_0 \bar{W}_{20}}{\partial \mu} = \frac{1}{2} [(\bar{h} \bar{i}') - (\bar{h} \bar{i})_{+1}]$$

Parameter derivative of the gaussian image height:

$$\frac{1}{\eta'} \cdot \frac{\partial \eta'}{\partial p} = \frac{1}{H} \cdot \frac{\partial H}{\partial p} = \frac{2}{H} \left(\frac{\bar{u}_k'}{u_k'} \right) \frac{\partial_0 W_{20}}{\partial p} - \frac{1}{H} \frac{\partial_1 W_{11}}{\partial p} \quad \text{(see 6-31)}$$

where, for small values of u' [NCODE2 ≤ 1], use

$$\frac{1}{\bar{u}'} \cdot \frac{\partial \bar{u}'}{\partial p} = \frac{1}{H} \cdot \frac{\partial H}{\partial p} = -\frac{1}{H} \cdot \frac{\partial_1 W_{11}}{\partial p} \quad \text{(see 6-32)}$$

$p = c, \mu, z$.

The primary aberration coefficients are calculated using the formulae as given by Hopkins[1], namely

Spherical aberration $$S_I = \sum_1^k A^2 h \Delta\left(\frac{u}{n}\right)$$

Coma $$S_{II} = \sum_1^k AB h \Delta\left(\frac{u}{n}\right)$$

Astigmatism $$S_{III} = \sum_1^k B^2 h \Delta\left(\frac{u}{n}\right)$$

Petzval Curvature $$S_{IV} = \sum_1^k H^2 P \quad ; \quad P = -c\Delta\left(\frac{1}{n}\right)$$

Distortion $$S_V = \sum_1^k \left\{-H^2 PB h E_1 (2 + A h E_1) - B^3 h \Delta\left(\frac{1}{n}\right)\right\}$$

where $E_1 = \frac{1}{H}\left(\frac{\bar{h}}{h}\right)$

Spherical Aberration of the pupil

$$\bar{S}_I = \sum_1^k B^2 \bar{h} \Delta\left(\frac{\bar{u}}{n}\right)$$

Longitudinal Chromatic variation of focus $$C_L = \sum_1^k A h \Delta\left(\frac{\delta n}{n}\right)$$

Transverse Chromatic variation of focus $$C_T = \sum_1^k B h \Delta\left(\frac{\delta n}{n}\right)$$

$\delta n = n_F - n_C$ when only $\bar{\lambda}$ is involved

$\left.\begin{aligned} \delta n_1 &= n_d - n_C \\ \delta n_2 &= n_F - n_d \end{aligned}\right\}$ i.e. the partial dispersions are used when $\lambda \neq \bar{\lambda}$.

FLOW DIAGRAM (VII - ii)

Paraxial Ray Subroutine (Gaussian Optics and Distortion)

Set $S_I = S_{II} = S_{III} = S_{IV} = S_V = \bar{S}_I = C_L(1) = C_L(2) = C_L(3) = C_L(4) = 0.0$

$C_T(1) = C_T(2) = C_T(3) = C_T(4) = 0.0$

Cycle I ($I = 1, 2, \cdots,$ NOWAVE)

Cycle K ($K = 1, 2, \cdots,$ NOSURF)

Set INDEX $= K + (I-1).(NOSURF+1)$

Calculate the paraxial quantities on the right hand side; primed and unprimed quantities correspond to the storage locations (K+1) and (K) respectively; n and n' are the refractive indices referred to the storage locations (INDEX) and (INDEX+1) respectively; starting values u_1, h_1, $\bar{u}_1$, $\bar{h}_1$ corresponding to K=1 have already been stored.

$$H = n_1(u_1 \bar{h}_1 - \bar{u}_1 h_1)$$

$$u' = \frac{n u + h(n'-n)c}{n'}$$

$$\bar{u}' = \frac{n \bar{u} + \bar{h}(n'-n)c}{n'}$$

$$h_{+1} = h - d\,u'$$

$$\bar{h}_{+1} = \bar{h} - d\,\bar{u}$$

$$i = ch - u$$

$$\bar{i} = c\bar{h} - \bar{u}$$

$$i' = \frac{(n\,i)}{n'}$$

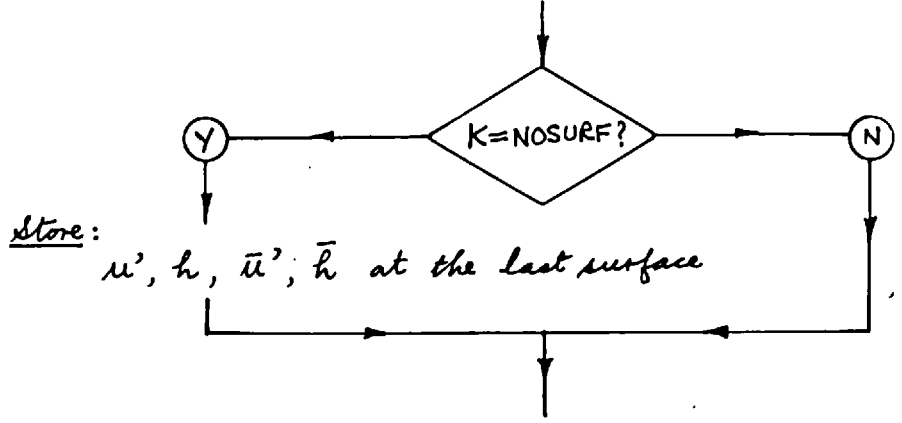
$$\bar{i}' = \frac{(n\,\bar{i})}{n'}$$

K=NOSTOP ?

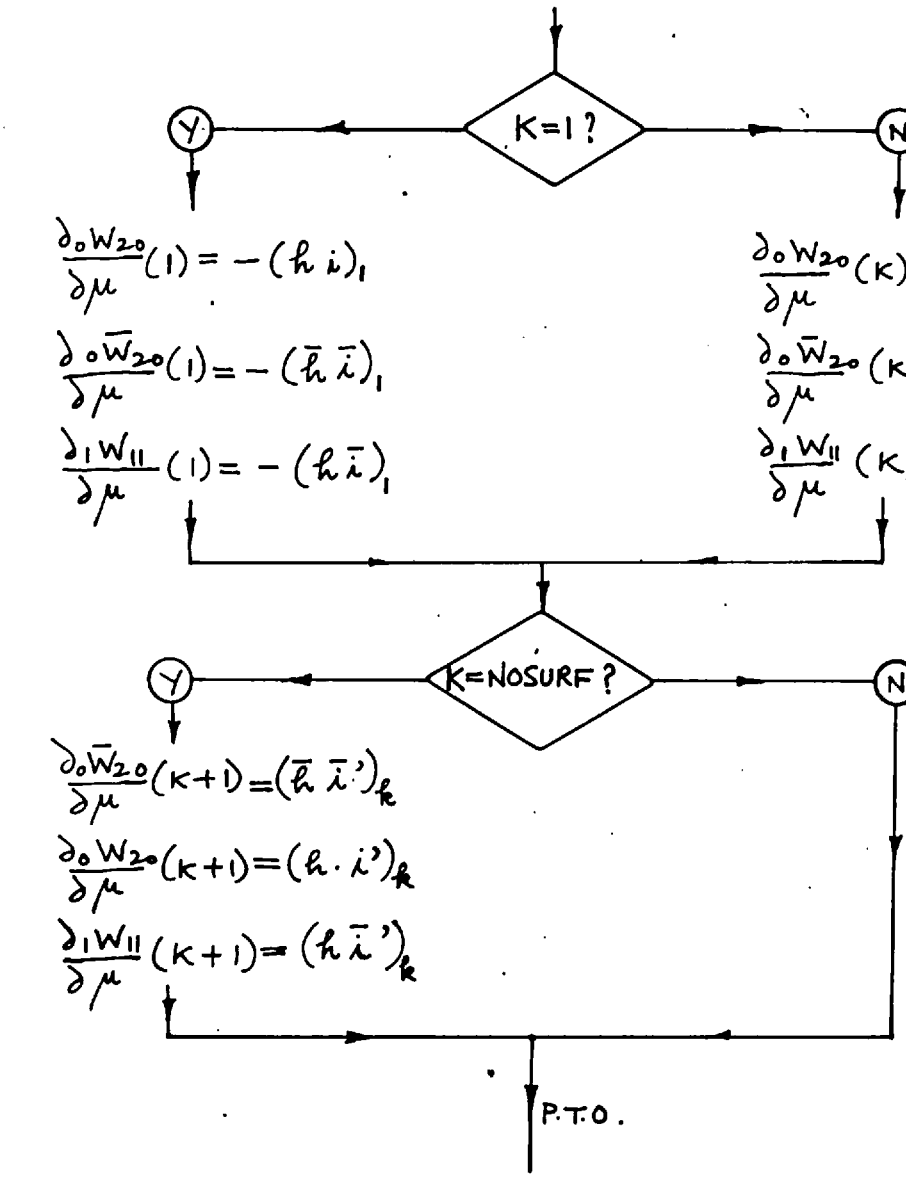
(Y) Store: (u, h) at "STOP"

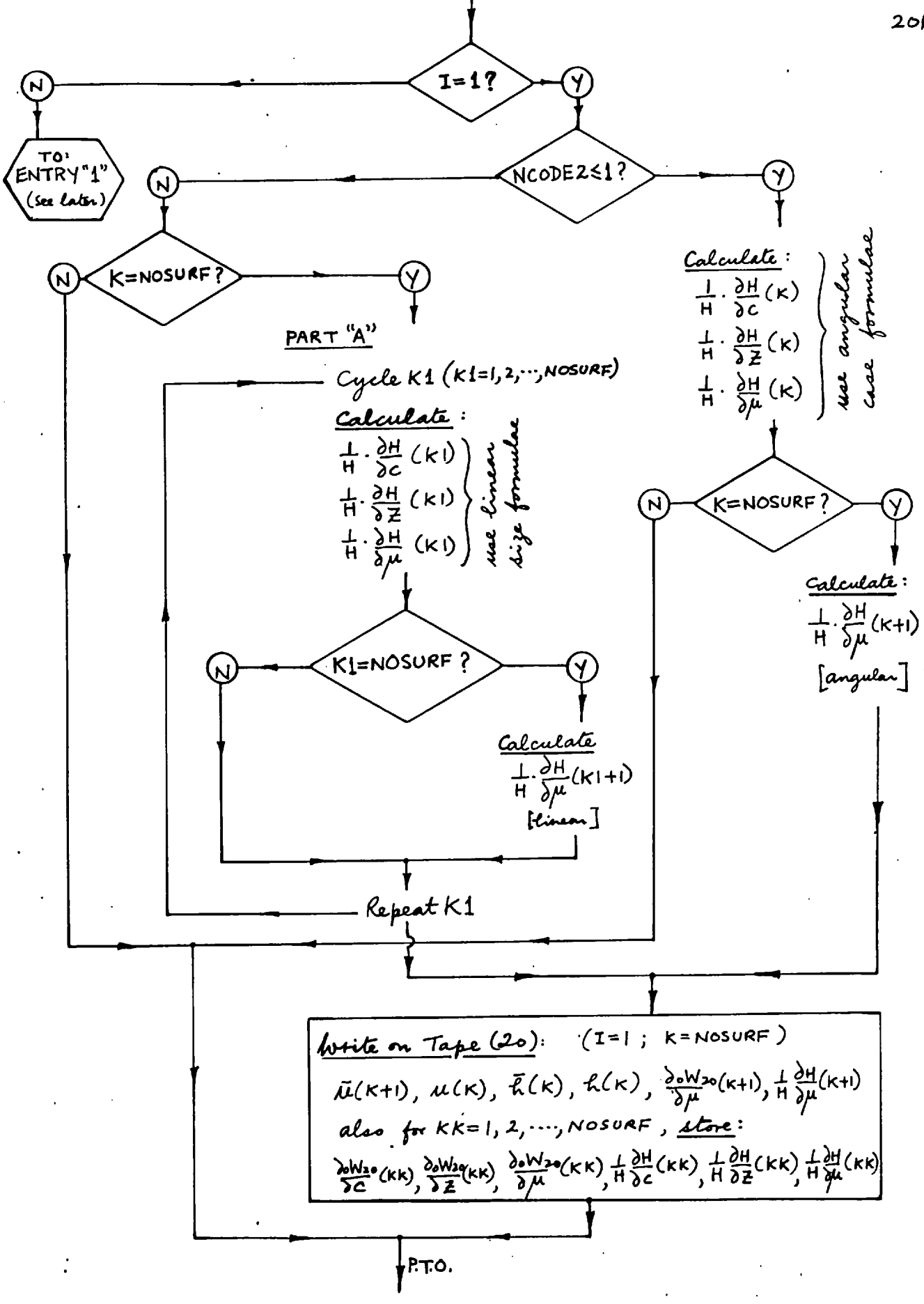
(N)

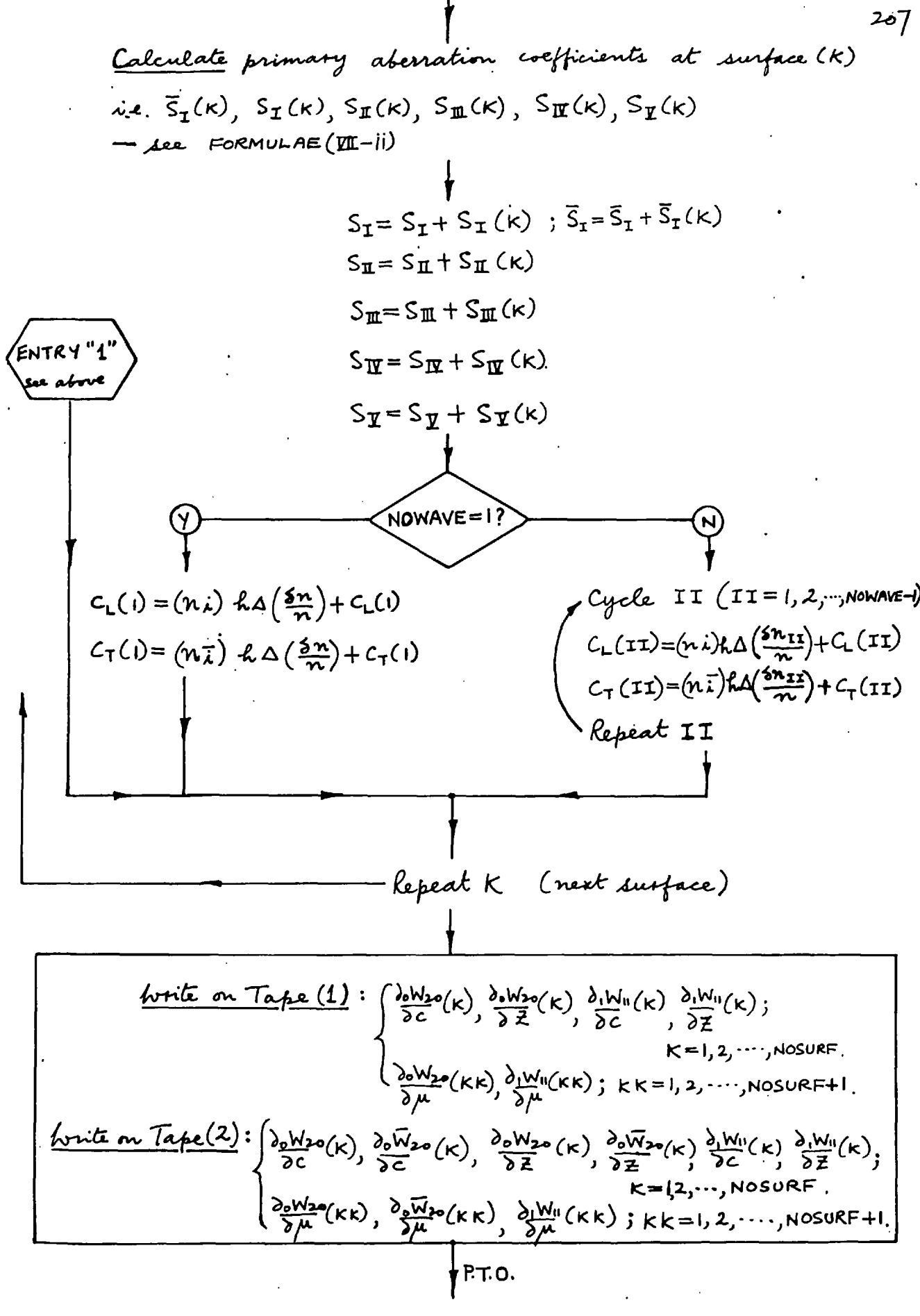
P.T.O.



Calculate $\left\{ \begin{array}{l} \frac{\partial_0 W_{20}}{\partial c}(K) \; ; \; \frac{\partial_0 W_{20}}{\partial z}(K) \; ; \; \frac{\partial_1 W_{11}}{\partial c}(K) \; ; \; \frac{\partial_1 W_{11}}{\partial z}(K) \\ \frac{\partial_0 \bar{W}_{20}}{\partial c}(K) \; ; \; \frac{\partial_0 \bar{W}_{20}}{\partial z}(K) \; ; \end{array} \right.$ see FORMULAE (VII-ii)




I=1?
N
Y
TO: ENTRY "1" (see later)
NCODE2≤1?
K=NOSURF?
PART "A"
Calculate:
$\frac{1}{H}\cdot\frac{\partial H}{\partial c}(K)$
$\frac{1}{H}\cdot\frac{\partial H}{\partial z}(K)$
$\frac{1}{H}\cdot\frac{\partial H}{\partial \mu}(K)$
use angular case formulae
Cycle K1 (K1=1,2,···,NOSURF)
Calculate:
$\frac{1}{H}\cdot\frac{\partial H}{\partial c}(K1)$
$\frac{1}{H}\cdot\frac{\partial H}{\partial z}(K1)$
$\frac{1}{H}\cdot\frac{\partial H}{\partial \mu}(K1)$
use linear case formulae
K1=NOSURF?
Calculate
$\frac{1}{H}\cdot\frac{\partial H}{\partial \mu}(K1+1)$
[linear]
Repeat K1
K=NOSURF?
Calculate:
$\frac{1}{H}\cdot\frac{\partial H}{\partial \mu}(K+1)$
[angular]
Write on Tape (20): (I=1 ; K=NOSURF)
$\bar{\mu}(K+1)$, $\mu(K)$, $\bar{h}(K)$, $h(K)$, $\frac{\partial_0 W_{20}}{\partial \mu}(K+1)$, $\frac{1}{H}\frac{\partial H}{\partial \mu}(K+1)$
also for KK=1, 2, ···, NOSURF, store:
$\frac{\partial_0 W_{20}}{\partial c}(KK)$, $\frac{\partial_0 W_{20}}{\partial z}(KK)$, $\frac{\partial_0 W_{20}}{\partial \mu}(KK)$, $\frac{1}{H}\frac{\partial H}{\partial c}(KK)$, $\frac{1}{H}\frac{\partial H}{\partial z}(KK)$, $\frac{1}{H}\frac{\partial H}{\partial \mu}(KK)$
P.T.O.


Calculate primary aberration coefficients at surface (K)
i.e. $\bar{S}_I(K)$, $S_I(K)$, $S_{II}(K)$, $S_{III}(K)$, $S_{IV}(K)$, $S_V(K)$
— see FORMULAE (VIII-ii)
$S_I = S_I + S_I(K)$; $\bar{S}_I = \bar{S}_I + \bar{S}_I(K)$
$S_{II} = S_{II} + S_{II}(K)$
$S_{III} = S_{III} + S_{III}(K)$
$S_{IV} = S_{IV} + S_{IV}(K)$
$S_V = S_V + S_V(K)$
ENTRY "1" see above
NOWAVE=1?
Y
N
$C_L(1) = (ni)\, h\Delta\left(\frac{\delta n}{n}\right) + C_L(1)$
$C_T(1) = (n\bar{i})\, h\Delta\left(\frac{\delta n}{n}\right) + C_T(1)$
Cycle II (II = 1, 2, ···, NOWAVE−1)
$C_L(II) = (ni)\, h\Delta\left(\frac{\delta n_{II}}{n}\right) + C_L(II)$
$C_T(II) = (n\bar{i})\, h\Delta\left(\frac{\delta n_{II}}{n}\right) + C_T(II)$
Repeat II
Repeat K (next surface)
Write on Tape (1): $\frac{\partial_0 W_{20}}{\partial c}(K)$, $\frac{\partial_0 W_{20}}{\partial \bar{z}}(K)$, $\frac{\partial_1 W_{11}}{\partial c}(K)$, $\frac{\partial_1 W_{11}}{\partial \bar{z}}(K)$; K = 1, 2, ···, NOSURF.
$\frac{\partial_0 W_{20}}{\partial \mu}(KK)$, $\frac{\partial_1 W_{11}}{\partial \mu}(KK)$; KK = 1, 2, ···, NOSURF+1.
Write on Tape (2): $\frac{\partial_0 W_{20}}{\partial c}(K)$, $\frac{\partial_0 \bar{W}_{20}}{\partial c}(K)$, $\frac{\partial_0 W_{20}}{\partial \bar{z}}(K)$, $\frac{\partial_0 \bar{W}_{20}}{\partial \bar{z}}(K)$, $\frac{\partial_1 W_{11}}{\partial c}(K)$, $\frac{\partial_1 W_{11}}{\partial \bar{z}}(K)$; K = 1, 2, ···, NOSURF.
$\frac{\partial_0 W_{20}}{\partial \mu}(KK)$, $\frac{\partial_0 \bar{W}_{20}}{\partial \mu}(KK)$, $\frac{\partial_1 W_{11}}{\partial \mu}(KK)$; KK = 1, 2, ···, NOSURF+1.
P.T.O.

I = 1 ?

(Y) $\lambda = \bar{\lambda}$

(N)

Write on Tape (3) $\left\{ \begin{array}{l} \frac{\partial_1 W_{11}}{\partial z}(K) \;,\; \frac{\partial_1 W_{11}}{\partial c}(K) \;\; ; K = 1, 2, \cdots, \text{NOSURF}. \\ \frac{\partial_1 W_{11}}{\partial \mu}(K1) \;; \;\; K = 1, 2, \cdots, (\text{NOSURF}+1) \end{array} \right.$

NCODE2 > 1 ?

(Y)

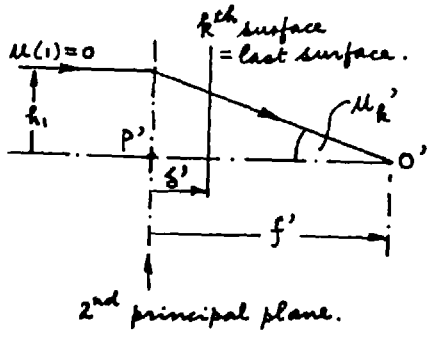
(N) angular magnification $= \frac{\bar{u}'_k}{\bar{u}_1}$

Calculate:

$$\text{THROW} = \sum_{i=2}^{k} d_i + \frac{h'_k}{u'_k} - \frac{h_1}{u_1}$$

Linear magnification $= \frac{\eta'}{\eta} = \frac{n_1 u_1}{n'_k u'_k}$

TO: ENTRY "2" see later



Set $H(1) = h_1$

$u(1) = 0$

Cycle K (K = 1, 2, ···, NOSURF)

$$u(K+1) = \frac{E(K) \cdot u(K) + H(K)[E(K+1) - E(K)] \cdot C(K)}{E(K+1)}$$

K = NOSURF ?

(N) $H(K+1) = H(K) - D(K+1) \cdot u(K+1)$

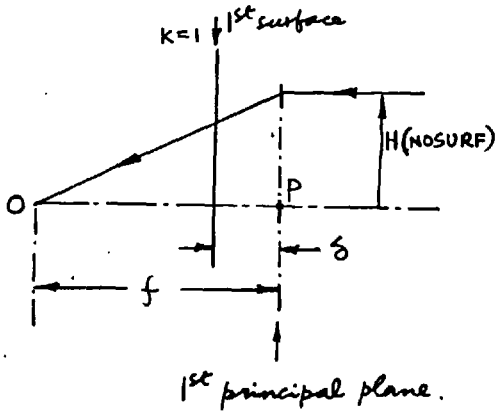
(Y)

Repeat K

Calculate: Second focal length: $f' = \frac{H(1)}{u'_k}$; $u'_k = u(\text{NOSURF}+1)$

$$\delta' = \frac{H(1) - H(\text{NOSURF})}{u'_k}$$

P.T.O.



Calculate: f and δ

Set $u(NOSURF+1) = 0$

$H(NOSURF) = 0.5 \times \rho_{max}(NOSURF)$

Cycle K ($K = NOSURF, NOSURF-1, \cdots, 2, 1$)

$$u(K) = \{E(K+1) \cdot u(K+1) - H(K)[E(K+1) - E(K)]c(K)\} / E(K)$$

K=1 ?

N

Y

$H(K-1) = H(K) + D(K) \cdot U(K)$

Repeat K

Calculate: 1st focal length $f = \dfrac{H(NOSURF)}{u(1)}$

$$\delta = \frac{H(1) - H(NOSURF)}{u(1)}$$

ENTRY "2" see above

Repeat I (next λ)

RETURN
END

(VII-iii) Finite-ray trace.

The finite ray trace subroutine to be described in the flow-diagram (VII-iii) calculates and stores the ray-trace data as well as the partial values of $\partial W/\partial p$. This subroutine is to be called by the axial ray, the pupil ray and the extra-axial ray subroutines; each of these will provide the appropriate starting data required for the finite ray trace, together with a set of code values for the purpose of identification.

For each wavelength, object point and ray, the starting data required are the direction cosines (L,M,N) of the ray in the object space, and the (X,Y) co-ordinates of the point $\underline{P}_1$ (see Fig. 4.8 and equation 4.26).

The code values required are :

NCODE4 - the number of image fields to be considered.
NCODE4=1 for the axial ray case.

NCODE6 - the number of rays involved.
NCODE6 = 2A+S for $\lambda = \bar{\lambda}$
= 2B+T for $\lambda \neq \bar{\lambda}$

NCODE7 - = 1 for the pupil ray case.
= 0 if not pupil ray case.

NCODE8 - = 1 for the extra-axial ray case.
= 0 for the axial ray case.

NCODE9 - = 2B+T, the number of rays to be traced for the extra-axial ray case, $\lambda \neq \bar{\lambda}$.

In this subroutine, the parameter derivatives $\partial W/\partial p$ are either calculated in parts or as a whole; for the former case, the notation $(\partial W/\partial p)_1$, $(\partial W/\partial p)_2$ is used to denote the partial values, and these are calculated using the data directly from the ray trace. In all cases, the parameter derivatives are calculated at each surface (each K value), and for each ray, wavelength and field. Preceeding the $\partial W/\partial p$ computations, the code values of NCODE7 and NCODE8 are tested to ensure that only the correct formulae are employed. At the end of the surface cycle, the values of $\partial W/\partial p$ calculated for p = c,z,u are stored on magnetic tapes, as indicated in the flow-diagram (VII-iii). In this way, the dimension of each of the subscripted variables $\partial W/\partial p$ need not exceed the maximum number of surfaces involved, independent of the number of rays, wavelengths and fields considered.

The subscripted variable $\partial W/\partial Q$ employs a special storage counter IQDRIC in the following way :

at the start of the subroutine, the counter LQDRIC is set to zero. Immediately outside the surface (K) cycle, the counter NQDRIC is set to unity. Inside the K cycle, the counter

$$IQDRIC = NQDRIC + LQDRIC$$

is introduced. When the surface (K) has Q(K) not equal to zero, then the quantity $\partial W/\partial Q$ calculated for this surface will be stored in $\partial W/\partial Q$ (IQDRIC); the same applies to the partial values $(\partial W/\partial Q)_1$ and $(\partial W/\partial Q)_2$; the counter NQDRIC is increased

by unity before the next surface is called. Immediately outside the surface cycle, the counter LQDRIC is set to a new value as follows :

a) LQDRIC = o if NCODE8 = 1, i.e. if it is the extra-axial ray case; and all $\partial W/\partial Q$'s are to be stored on to the tape (4).

b) LQDRIC=LQDRIC+MQDRIC if NCODE8=0 or

if NCODE7=1

MQDRIC being the total number of quadric surfaces ($Q \neq o$) in the system. If MQDRIC = 5, then the value of IQDRIC will not exceed 75.

The procedure employed for storing the derivatives $\partial W/\partial A_{2m}$ with m=2,3,4,5 is similar to the above case; the corresponding counters used for the quadric surface case and the figured surface case are listed below :

QUADRIC SURFACE	FIGURED SURFACE	
LQDRIC	LFIGRD	
NQDRIC	NOFIGR	see (VII-i)
MQDRIC	MFIGRD	see (VII-i)
IQDRIC	IFIGRD	

The subscripted variable used for the figured surface case is $\partial W/\partial A_{2m}$ (IFIGRD).

Referring to the flow-diagram (VII-iii), the starting data X(1), Y(1), L(1), M(1), N(1) for each field point, wavelength and ray are provided by the subroutine which precedes this ray trace subroutine. For the extra-axial case, the starting data are read from the Tape (22); for other cases, these data are obtained from the storage location (L20).

The ray-tracing part of the subroutine is essentially the same as that described in Chapter 4, section (IV-ii), for the pupil exploration; the notation and the formulae employed here are as described earlier. In the present subroutine, intermediate data such as (X,Y,Z), (L,M,N), corresponding to the first surface, the stop, the last surface and the image space (where appropriate, see part A, flow diagram VII-iii) are also stored; these will be required later in other subroutines.

In part B of the flow-diagram (VII-iii), two special counters, namely the KAPUPL and the KOPUPL, are employed; these are independent of the number of rays involved in the ray trace and are the correct counters to be used for the subscripted variables which are associated with the pupil rays.

Finally, the number of rays to be traced for the mean wavelength is 2A+S, and that for the other wavelengths is 2B+T ($\leqslant$2A+S). This is considered at the end of the subroutine.

FORMULAE (VII-iii)

Paraxial parameter derivatives. Extra-axial images:

(a) *Spherical surfaces:*

$p = c$ $\quad \dfrac{\partial W_{TI}}{\partial c} = \bar{y}\, h_T\, \Delta(n\cos\bar{I}) + \frac{1}{2}(\bar{y}^2+\bar{z}^2)(n\sin\bar{I})\,\Delta(\mu_T)$

store $\quad \left(\dfrac{\partial W_{TI}}{\partial c}\right)_1 = \bar{y}\,\Delta(n\cos\bar{I})$

$\left(\dfrac{\partial W_{TI}}{\partial c}\right)_2 = \frac{1}{2}(\bar{y}^2+\bar{z}^2)(n\sin\bar{I})$

$p = z$ $\quad \dfrac{\partial W_{TI}}{\partial z} = \Delta(n\bar{M}\mu_T)$

$p = \mu$ $\quad \dfrac{\partial W_{TI}}{\partial \mu} = (h_T \sin\bar{I}')_{-1} - (h_T \sin\bar{I})$

(b) *Aspheric surfaces:*

$p = c$ $\quad \dfrac{\partial W_{TI}}{\partial c} = \dfrac{\bar{y}\,h_T}{\bar{\gamma}(\bar{\beta}^2+\bar{\gamma}^2)}\cdot\Delta(n\cos\bar{I}) + \dfrac{\bar{y}^2}{(1+\bar{\gamma})\bar{\gamma}}\,\Delta(n\bar{M}\mu_T)$

store $\quad \left(\dfrac{\partial W_{TI}}{\partial c}\right)_1 = \dfrac{\bar{y}\,\Delta(n\cos\bar{I})}{\bar{\gamma}(\bar{\beta}^2+\bar{\gamma}^2)}$; $\left(\dfrac{\partial W_{TI}}{\partial c}\right)_2 = \dfrac{\bar{y}^2}{\bar{\gamma}(1+\bar{\gamma})}$

$p = Q$ $\quad \dfrac{\partial W_{TI}}{\partial Q} = \dfrac{(\bar{y}c)^3 h_T}{2\bar{\gamma}(\bar{\beta}^2+\bar{\gamma}^2)}\Delta(n\cos\bar{I}) + \dfrac{c\bar{z}_s^2}{2\bar{\gamma}}\Delta(n\bar{M}\mu_T)$

store $\quad \left(\dfrac{\partial W_{TI}}{\partial Q}\right)_1 = \dfrac{(\bar{y}c)^3\Delta(n\cos\bar{I})}{2\bar{\gamma}(\bar{\beta}^2+\bar{\gamma}^2)}$; $\left(\dfrac{\partial W_{TI}}{\partial Q}\right)_2 = \dfrac{c\bar{z}_s}{2\gamma}$

$p = A_{2m}$ $\quad \dfrac{\partial W_{TI}}{\partial A_{2m}} = \dfrac{2m\bar{y}\bar{k}^2}{\rho_{max}^2}\,p^{(m-1)}\,h_T\,\Delta(n\cos\bar{I}) + p^m\Delta(n\bar{M}\mu_T)$

$m = 2, 3, 4, 5$ $\quad$ *store* $\quad \left(\dfrac{\partial W_{TI}}{\partial A_{2m}}\right)_1 = \dfrac{2m\bar{y}(\bar{k}^2)p^{(m-1)}\Delta(n\cos\bar{I})}{\rho_{max}^2}$

$\left(\dfrac{\partial W_{TI}}{\partial A_{2m}}\right)_2 = p^m$

$p = \mu$ $\quad \dfrac{\partial W_{TI}}{\partial \mu} = (h_T\sin\bar{I}') - (h_T\sin\bar{I})_{+1}$

$p = z$ $\quad \dfrac{\partial W_{TI}}{\partial z} = \Delta(n\bar{M}\mu_T)$

Finite-ray parameter derivatives. Axial images:

(a) Spherical surfaces:

$p = c \qquad \frac{\partial W_i}{\partial c} = \frac{1}{2}(Y^2 + Z^2)\,\Delta(n\cos I)$

store $\left(\frac{\partial W}{\partial c}\right)_1 = \frac{1}{2}(Y^2+Z^2)\,\Delta(n\cos I)$

$p = z \qquad \frac{\partial W_i}{\partial z} = (1 - cz)\,\Delta(n\cos I) - \Delta(n)$

store $\left(\frac{\partial W}{\partial z}\right)_1 = (1-cz)\,\Delta(n\cos I) - \Delta(n)$

$p = \mu \qquad \frac{\partial W_i}{\partial \mu} = d - D$

$= \left[e'(\kappa - 1) - e(\kappa)\right]$

(b) Aspheric surfaces:

$p = c \qquad \frac{\partial W_i}{\partial c} = \frac{Y^2}{(1+\gamma)\sqrt{\alpha^2+\beta^2+\gamma^2}}\,\Delta(n\cos I)$

$p = Q \qquad \frac{\partial W_i}{\partial Q} = \frac{c\,Z_s^2}{2\sqrt{\alpha^2+\beta^2+\gamma^2}}\,\Delta(n\cos I)$

$p = A_{2m} \qquad m = 2, 3, 4, 5 \qquad \frac{\partial W_i}{\partial A_{2m}} = k\,p^m\,\Delta(n\cos I)$

$p = z \qquad \frac{\partial W_i}{\partial z} = k\,\Delta(n\cos I) - \Delta(n)$

store $\left(\frac{\partial W}{\partial z}\right)_2 = \Delta(n)$

$p = \mu \qquad \frac{\partial W_i}{\partial \mu} = e'(\kappa - 1) - e(\kappa)$

Finite-ray parameter derivatives. Extra-axial images:

(a) Spherical surfaces

$p = c$ $\quad \dfrac{\partial W_i}{\partial c} = \frac{1}{2}(x^2+y^2+z^2)\Delta(n\cos I) - \frac{1}{2}(\bar{y}^2+\bar{z}^2)\Delta(n\cos\bar{I})$

Store $\quad \left(\dfrac{\partial W}{\partial c}\right)_1 = \frac{1}{2}(x^2+y^2+z^2)\Delta(n\cos I)$

$\left(\dfrac{\partial W}{\partial c}\right)_2 = \frac{1}{2}(\bar{y}^2+\bar{z}^2)\Delta(n\cos\bar{I})$

$p = z$ $\quad \dfrac{\partial W_i}{\partial z} = (1-cz)\Delta(n\cos I) - (1-c\bar{z})\Delta(n\cos\bar{I})$

Store $\quad \left(\dfrac{\partial W}{\partial z}\right)_1 = (1-cz)\Delta(n\cos I)$

$\left(\dfrac{\partial W}{\partial z}\right)_2 = (1-c\bar{z})\Delta(n\cos\bar{I})$

$p = \mu$ $\quad \dfrac{\partial W_i}{\partial \mu} = \bar{D}-D = \left[e'(\kappa-1) - e(\kappa)\right] - \left[\bar{e}'(\kappa-1) - \bar{e}(\kappa)\right]$

Store $\quad \left(\dfrac{\partial W}{\partial \mu}\right)_2 = \left[\bar{e}'(\kappa-1) - \bar{e}(\kappa)\right]$

(b) Aspheric Surfaces

$p = c$ $\quad \dfrac{\partial W_i}{\partial c} = \dfrac{x^2+y^2}{(1+\gamma)\sqrt{\alpha^2+\beta^2+\gamma^2}}\Delta(n\cos I) - \dfrac{\bar{y}^2}{(1+\bar{\gamma})\sqrt{\bar{\beta}^2+\bar{\gamma}^2}}\Delta(n\cos\bar{I})$

Store $\quad \left(\dfrac{\partial W}{\partial c}\right)_2 = \dfrac{\bar{y}^2\,\Delta(n\cos\bar{I})}{(1+\bar{\gamma})\sqrt{\bar{\beta}^2+\bar{\gamma}^2}}$

$p = Q$ $\quad \dfrac{\partial W_i}{\partial Q} = \dfrac{c\,z_s^2}{2\cdot\sqrt{\alpha^2+\beta^2+\gamma^2}}\Delta(n\cos I) - \dfrac{c\,\bar{z}_s^2}{2\cdot\sqrt{\bar{\beta}^2+\bar{\gamma}^2}}\Delta(n\cos\bar{I})$

Store $\quad \left(\dfrac{\partial W}{\partial Q}\right)_2 = \dfrac{c\,z_s^2}{2\cdot\sqrt{\bar{\beta}^2+\bar{\gamma}^2}}$

$p = A_{2m}$, $m = 2,3,4,5$ $\quad \dfrac{\partial W_i}{\partial A_{2m}} = k\,p^m\Delta(n\cos I) - \bar{k}\,\bar{p}^m\Delta(n\cos\bar{I})$

Store $\left(\dfrac{\partial W}{\partial A_{2m}}\right)_2 = \bar{k}\,\bar{p}^m\Delta(n\cos\bar{I})$

$p = z$ $\quad \dfrac{\partial W_i}{\partial z} = k\Delta(n\cos I) - \bar{k}\Delta(n\cos\bar{I})$

Store $\left(\dfrac{\partial W}{\partial z}\right)_1 = k\Delta(n\cos I)$; $\left(\dfrac{\partial W}{\partial z}\right)_2 = \bar{k}\Delta(n\cos\bar{I})$

$p = \mu$ $\quad \dfrac{\partial W_i}{\partial \mu} = \bar{D} - D = [e'(k-1) - e(k)] - [\bar{e}'(k-1) - \bar{e}(k)]$

Store $\left(\dfrac{\partial W}{\partial \mu}\right)_2 = [\bar{e}'(k-1) - \bar{e}(k)]$

Tangential curvature C_T for aspheric surfaces

$$C_T = \frac{c}{(\beta^2+\gamma^2)^{3/2}} + k^3\left\{\left[\frac{2Y}{\rho^2_{max}}\right]^2 \cdot \frac{\partial^2 A}{\partial \rho^2} + \frac{2}{\rho^2_{max}} \cdot \frac{\partial A}{\partial \rho}\right\}$$

which reduces to c for $Q = 0$, $A_4 = A_6 = A_8 = A_{10} = 0$.
The values of β, γ, k and the derivatives

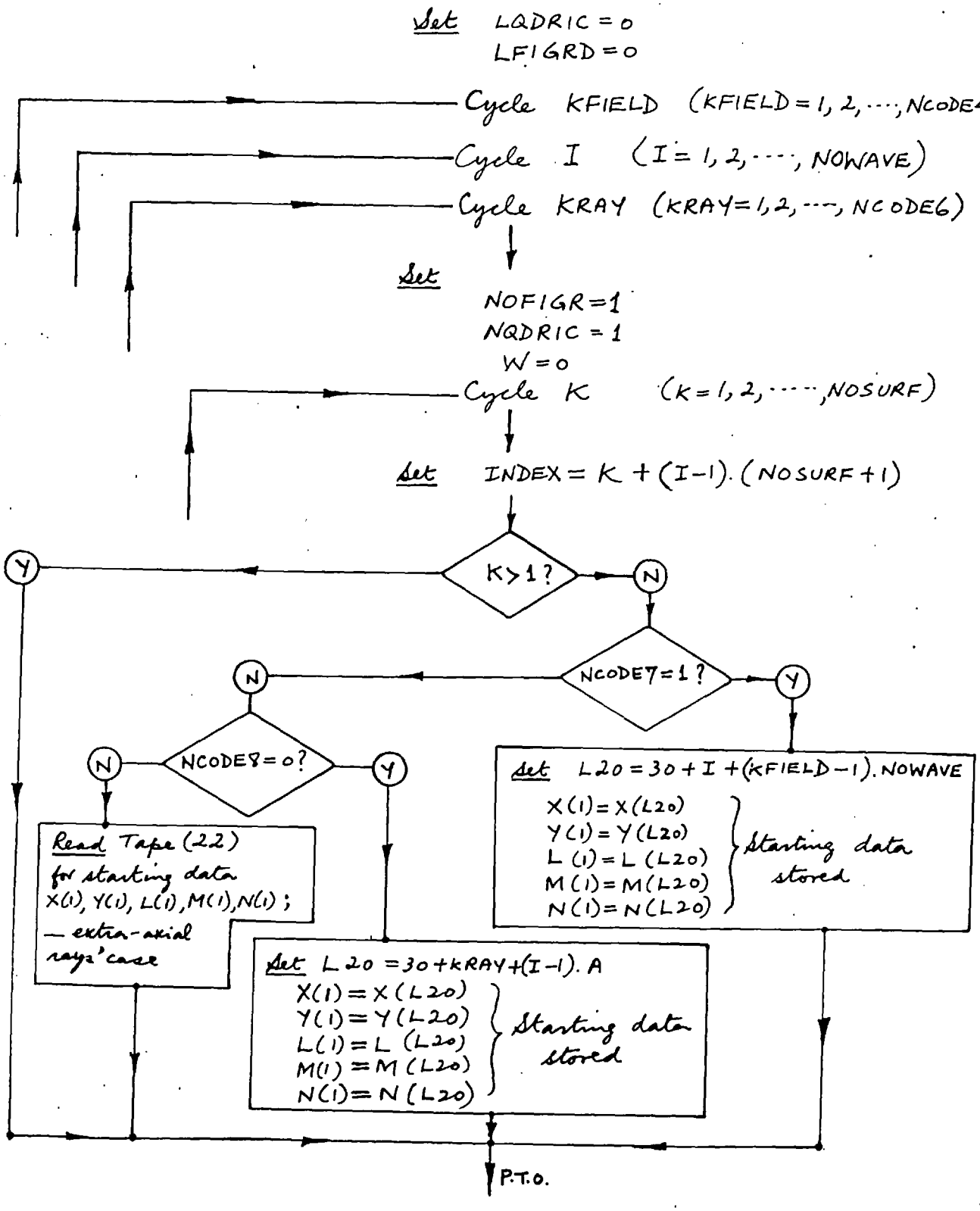
$$\frac{\partial A}{\partial \rho} = 2A_4\rho + 3A_6\rho^2 + 4A_8\rho^3 + 5A_{10}\rho^4$$

$$\frac{\partial^2 A}{\partial \rho^2} = 2A_4 + 6A_6\rho + 12A_8\rho^2 + 20A_{10}\rho^3$$

are calculated for the pupil ray (see Flow Diagram III-iii) at whose point of incidence with the surface the curvature C_T is required to be known. The value of C_T at each aspheric surface, and for each pupil ray and each λ is required to be calculated and stored for the T-ray subroutine.

FLOW DIAGRAM (VII - iii)

Finite Ray Trace Subroutine



Set $IQDRIC = NQDRIC + LQDRIC$

$IFIGRD = NOFIGR + LFIGRD$

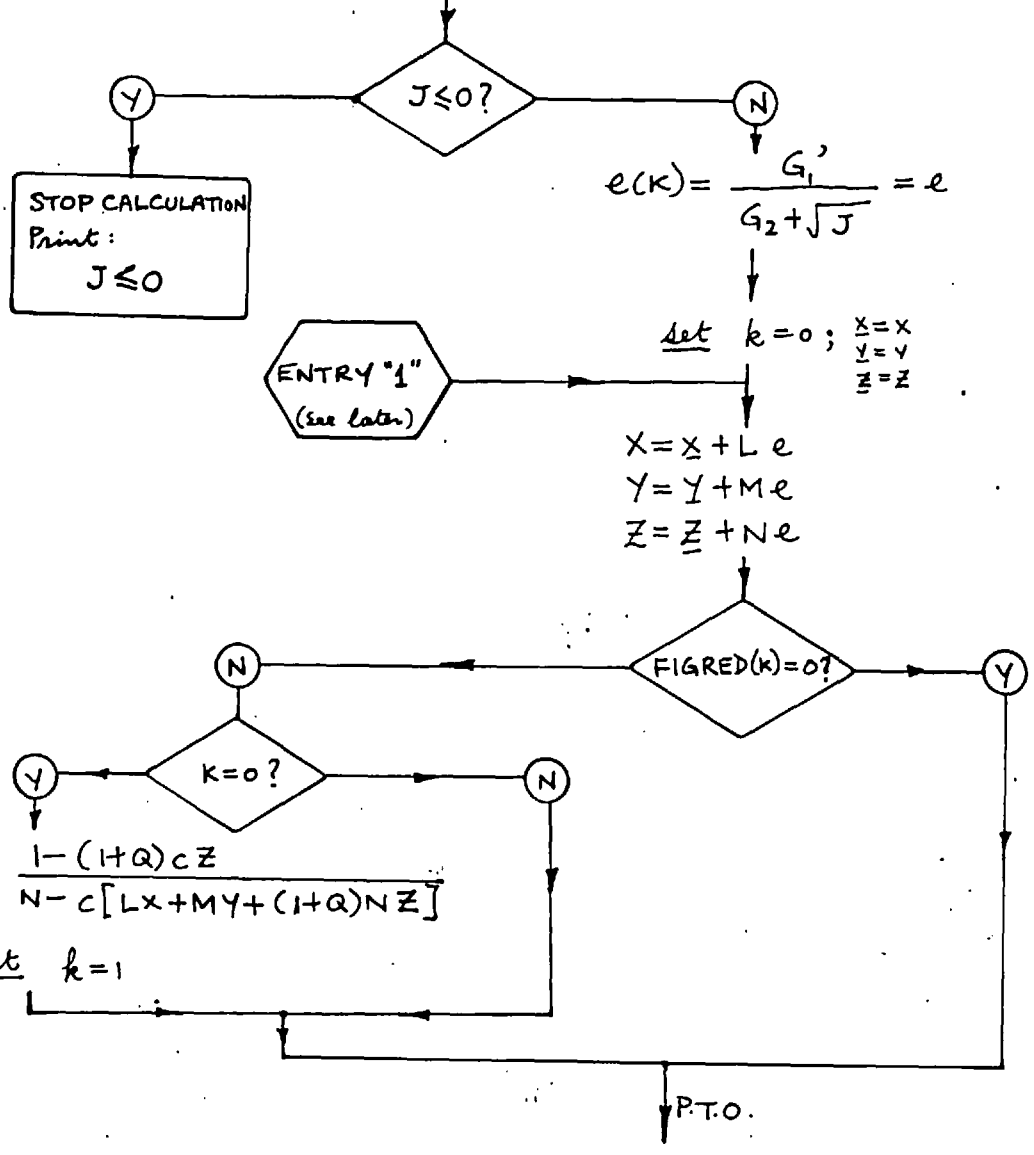
$$Z(K) = -\left\{\frac{L(K)\,X(K) + M(K)\,Y(K)}{1 + N(K)}\right\}$$

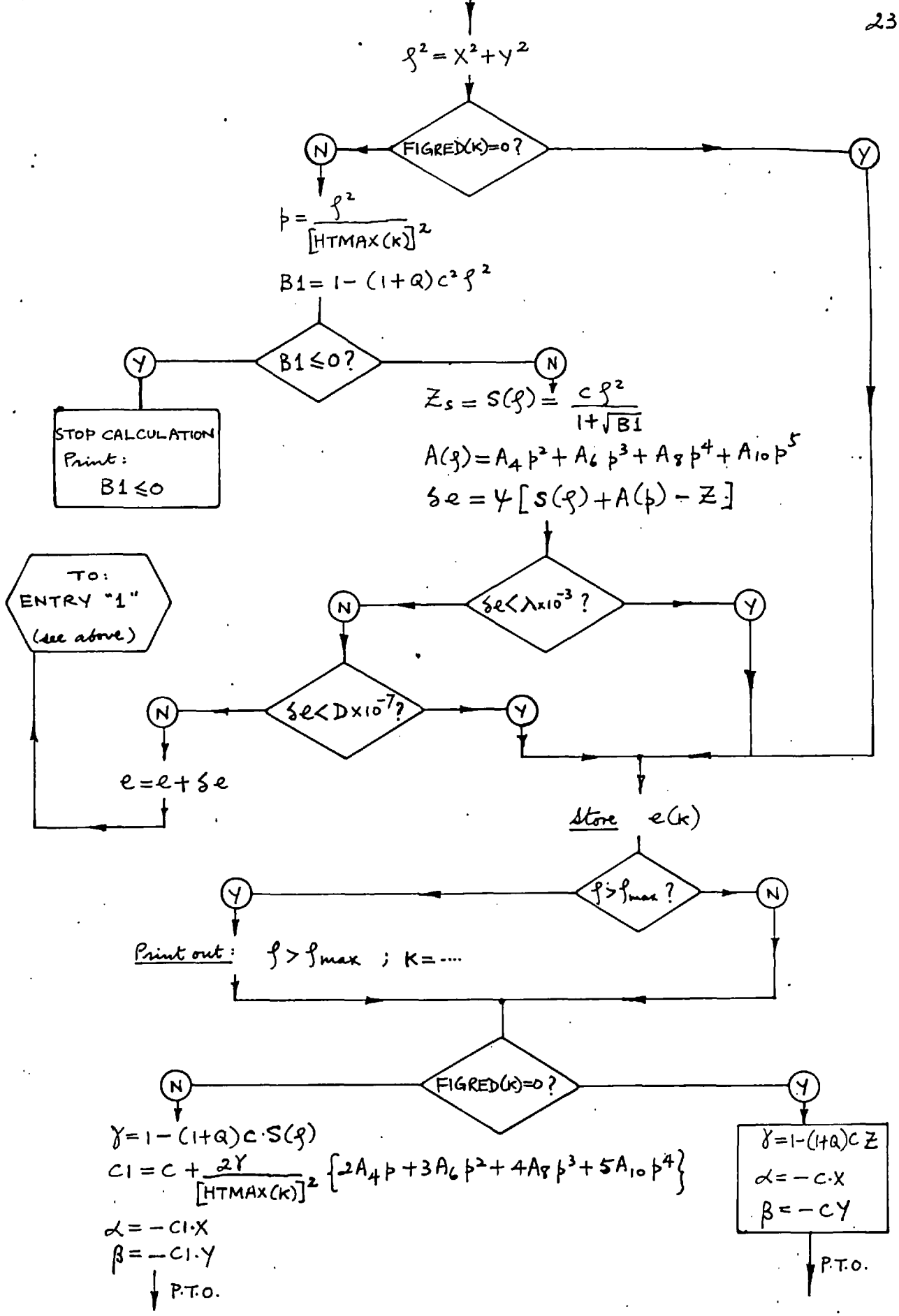
all the following quantities are associated with the current surface "K"

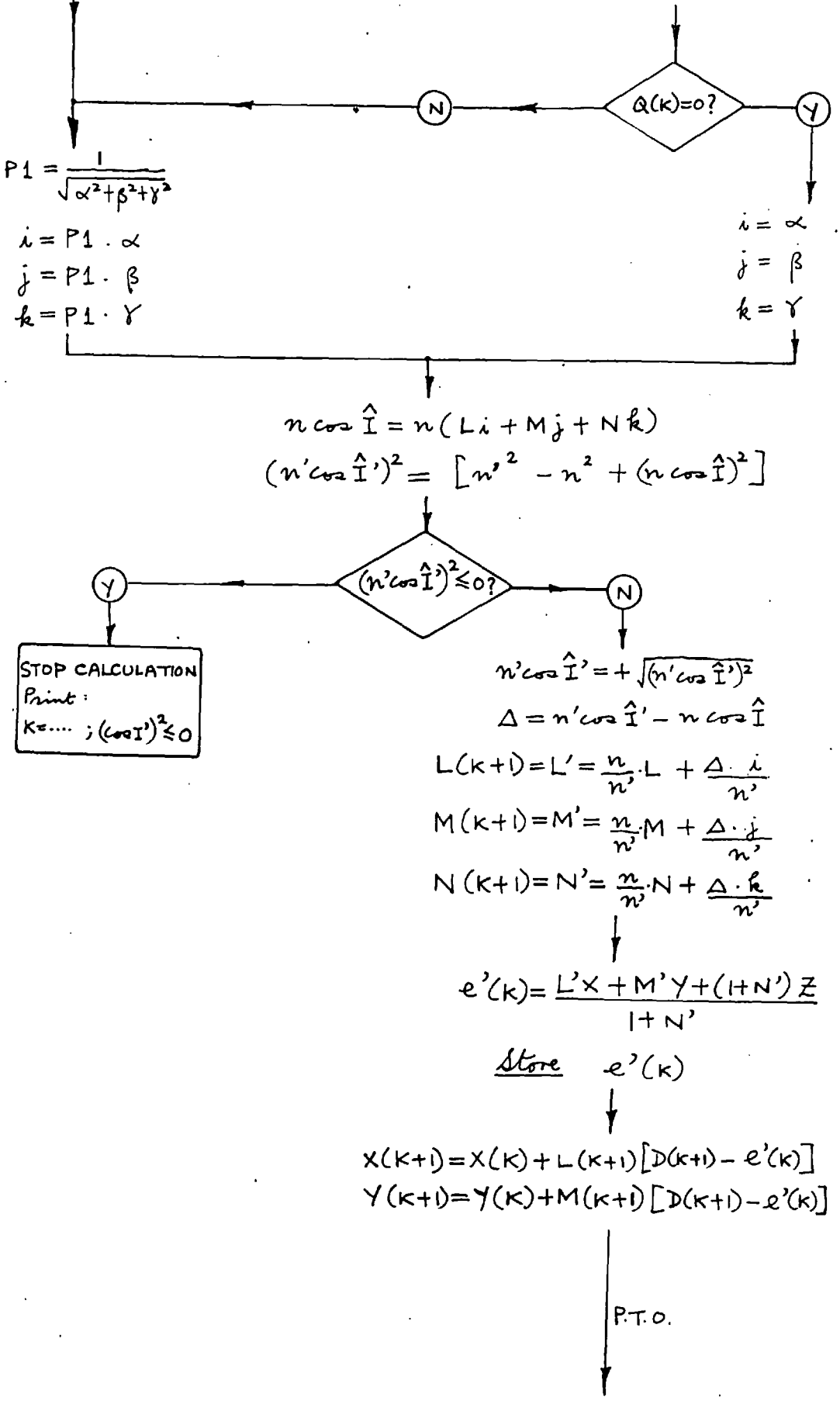
$$G_1' = c\left[X^2 + Y^2 + (1+Q)Z^2\right] - 2Z$$

$$G_2 = N - c\left[LX + MY + (1+Q)NZ\right]$$

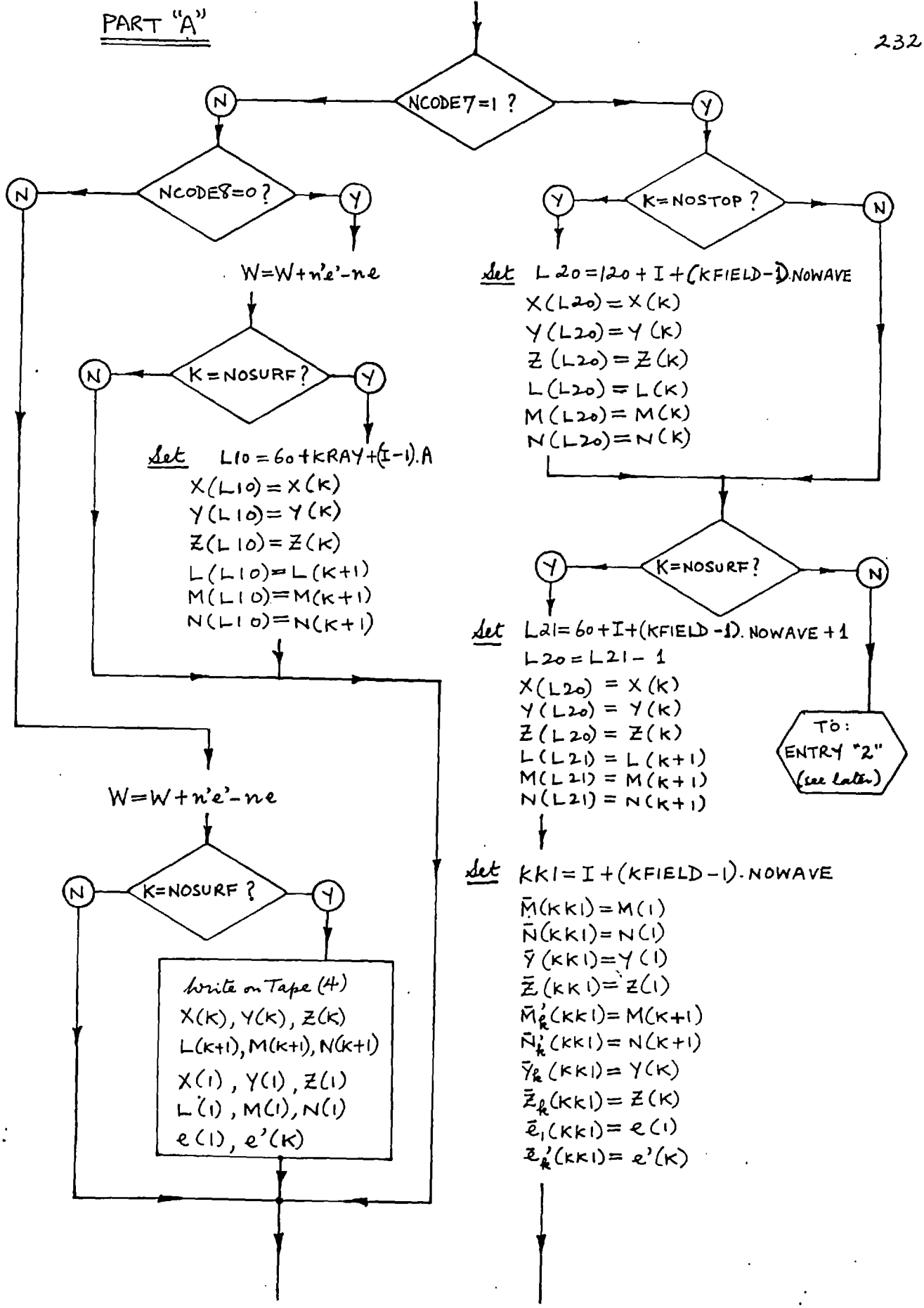
$$J = G_2^2 - (1+QN^2)\,c\,G_1'$$

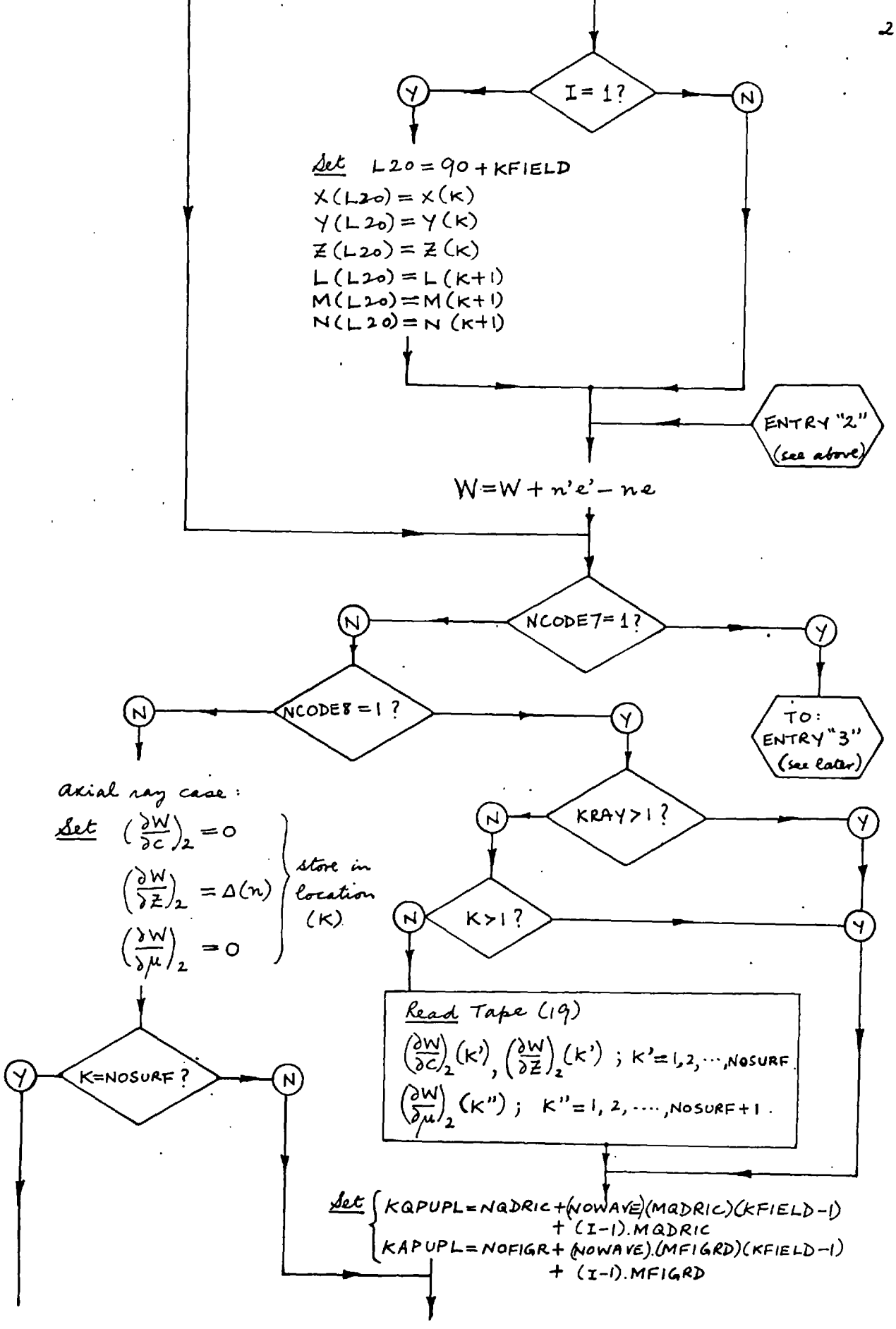


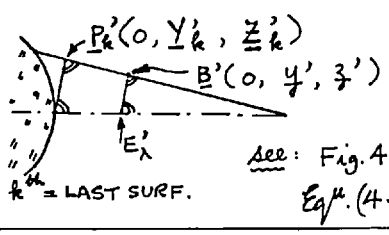

$\rho^2 = X^2 + Y^2$
FIGRED(K)=0 ?
N
Y
$p = \frac{\rho^2}{[HTMAX(K)]^2}$
$B1 = 1 - (1+Q)c^2\rho^2$
B1 ≤ 0 ?
Y
N
STOP CALCULATION
Print:
B1 ≤ 0
$Z_s = S(\rho) = \frac{c\rho^2}{1+\sqrt{B1}}$
$A(p) = A_4 p^2 + A_6 p^3 + A_8 p^4 + A_{10} p^5$
$\delta e = \psi [S(\rho) + A(p) - Z]$
TO:
ENTRY "1"
(see above)
$\delta e < \lambda \times 10^{-3}$?
N
Y
$\delta e < D \times 10^{-7}$?
N
Y
$e = e + \delta e$
Store e(K)
$\rho > \rho_{max}$?
Y
N
Print out: $\rho > \rho_{max}$; K=....
FIGRED(K)=0 ?
N
Y
$\gamma = 1 - (1+Q)c \cdot S(\rho)$
$C1 = C + \frac{2\gamma}{[HTMAX(K)]^2} \{2A_4 p + 3A_6 p^2 + 4A_8 p^3 + 5A_{10} p^4\}$
$\alpha = -C1 \cdot X$
$\beta = -C1 \cdot Y$
P.T.O.
$\gamma = 1 - (1+Q)cZ$
$\alpha = -c \cdot X$
$\beta = -cY$
P.T.O.


Q(K)=0?
N
Y
$P1 = \frac{1}{\sqrt{\alpha^2+\beta^2+\gamma^2}}$
$i = P1 \cdot \alpha$
$j = P1 \cdot \beta$
$k = P1 \cdot \gamma$
$i = \alpha$
$j = \beta$
$k = \gamma$
$n\cos\hat{I} = n(Li + Mj + Nk)$
$(n'\cos\hat{I}')^2 = [n'^2 - n^2 + (n\cos\hat{I})^2]$
$(n'\cos\hat{I}')^2 \leq 0$?
Y
N
STOP CALCULATION
Print:
K=.... ; $(\cos I')^2 \leq 0$
$n'\cos\hat{I}' = +\sqrt{(n'\cos\hat{I}')^2}$
$\Delta = n'\cos\hat{I}' - n\cos\hat{I}$
$L(K+1) = L' = \frac{n}{n'} \cdot L + \frac{\Delta \cdot i}{n'}$
$M(K+1) = M' = \frac{n}{n'} \cdot M + \frac{\Delta \cdot j}{n'}$
$N(K+1) = N' = \frac{n}{n'} \cdot N + \frac{\Delta \cdot k}{n'}$
$e'(K) = \frac{L'X + M'Y + (1+N')Z}{1+N'}$
Store $e'(K)$
$X(K+1) = X(K) + L(K+1)[D(K+1) - e'(K)]$
$Y(K+1) = Y(K) + M(K+1)[D(K+1) - e'(K)]$
P.T.O.

PART "A"


NCODE7=1 ?
N
Y
NCODE8=0 ?
N
Y
W=W+n'e'-ne
K=NOSURF ?
N
Y
Set L10=60+KRAY+(I-1).A
X(L10)=X(K)
Y(L10)=Y(K)
Z(L10)=Z(K)
L(L10)=L(K+1)
M(L10)=M(K+1)
N(L10)=N(K+1)
W=W+n'e'-ne
K=NOSURF ?
N
Y
Write on Tape (4)
X(K), Y(K), Z(K)
L(K+1), M(K+1), N(K+1)
X(1), Y(1), Z(1)
L(1), M(1), N(1)
e(1), e'(K)
K=NOSTOP ?
Y
N
Set L20=120+I+(KFIELD-1).NOWAVE
X(L20)=X(K)
Y(L20)=Y(K)
Z(L20)=Z(K)
L(L20)=L(K)
M(L20)=M(K)
N(L20)=N(K)
K=NOSURF ?
Y
N
Set L21=60+I+(KFIELD-1).NOWAVE+1
L20=L21-1
X(L20)=X(K)
Y(L20)=Y(K)
Z(L20)=Z(K)
L(L21)=L(K+1)
M(L21)=M(K+1)
N(L21)=N(K+1)
TO: ENTRY "2" (see later)
Set KK1=I+(KFIELD-1).NOWAVE
M̄(KK1)=M(1)
N̄(KK1)=N(1)
Ȳ(KK1)=Y(1)
Z̄(KK1)=Z(1)
M̄'k(KK1)=M(K+1)
N̄'k(KK1)=N(K+1)
Ȳk(KK1)=Y(K)
Z̄k(KK1)=Z(K)
ē1(KK1)=e(1)
ē'k(KK1)=e'(K)


I = 1?
Y
N
Set L20 = 90 + KFIELD
X(L20) = X(K)
Y(L20) = Y(K)
Z(L20) = Z(K)
L(L20) = L(K+1)
M(L20) = M(K+1)
N(L20) = N(K+1)
ENTRY "2" (see above)
W = W + n'e' − ne
NCODE7 = 1?
NCODE8 = 1?
TO: ENTRY "3" (see later)
axial ray case:
Set (∂W/∂C)2 = 0
(∂W/∂Z)2 = Δ(n)
(∂W/∂μ)2 = 0
store in location (K)
KRAY > 1?
K > 1?
Read Tape (19)
(∂W/∂C)2 (K'), (∂W/∂Z)2 (K') ; K' = 1, 2, ···, NOSURF
(∂W/∂μ)2 (K") ; K" = 1, 2, ···, NOSURF + 1.
K = NOSURF?
Set KQPUPL = NQDRIC + (NOWAVE)(MQDRIC)(KFIELD − 1) + (I − 1).MQDRIC
KAPUPL = NOFIGR + (NOWAVE).(MFIGRD)(KFIELD − 1) + (I − 1).MFIGRD



see: Fig. 4-3
Eqμ. (4-12)

Locate P_k' : PART "C"

$$Y_k' = Y(K) - M'(K+1)\cdot e'(K)$$

$$B' \begin{cases} y' = \dfrac{n_k'}{H_\lambda}\left(-\bar{u}_k' Y_k' - M_k' \bar{h}_k\right) \\ z' = -\left(\dfrac{M_k' y'}{1+N_k'}\right) \end{cases}$$

L 20 = KRAY + (I−1).A

Store y' and z' in the location (L 20).

Q(K)=0?

Calculate:

$$\left(\frac{\partial W_i}{\partial C}\right)(K) = \left(\frac{\partial W}{\partial C}\right)_1 - \left(\frac{\partial W}{\partial C}\right)_2$$

$$\left(\frac{\partial W_i}{\partial Z}\right)(K) = \left(\frac{\partial W}{\partial Z}\right)_1 - \left(\frac{\partial W}{\partial Z}\right)_2$$

K=1 ?

$$\left(\frac{\partial W_i}{\partial \mu}\right)(K) = e'(K-1) - e(K) - \left(\frac{\partial W}{\partial \mu}\right)_2$$

$$\frac{\partial W_i}{\partial \mu} = -e(1) - \left(\frac{\partial W}{\partial \mu}\right)_2$$

K=NOSURF?

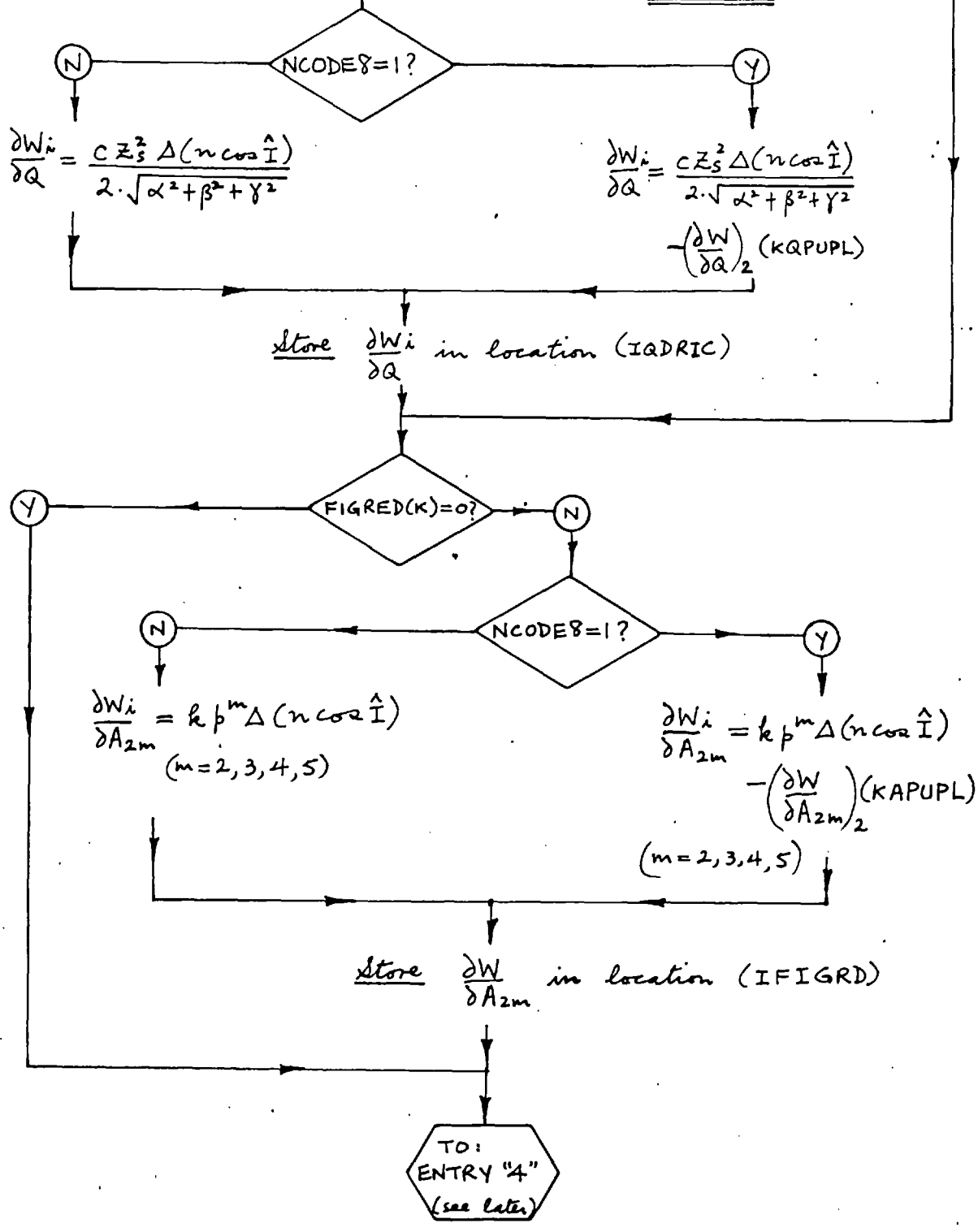
$$\frac{\partial W_i}{\partial \mu}(K+1) = e'(K) - \left(\frac{\partial W}{\partial \mu}\right)_2(K+1)$$

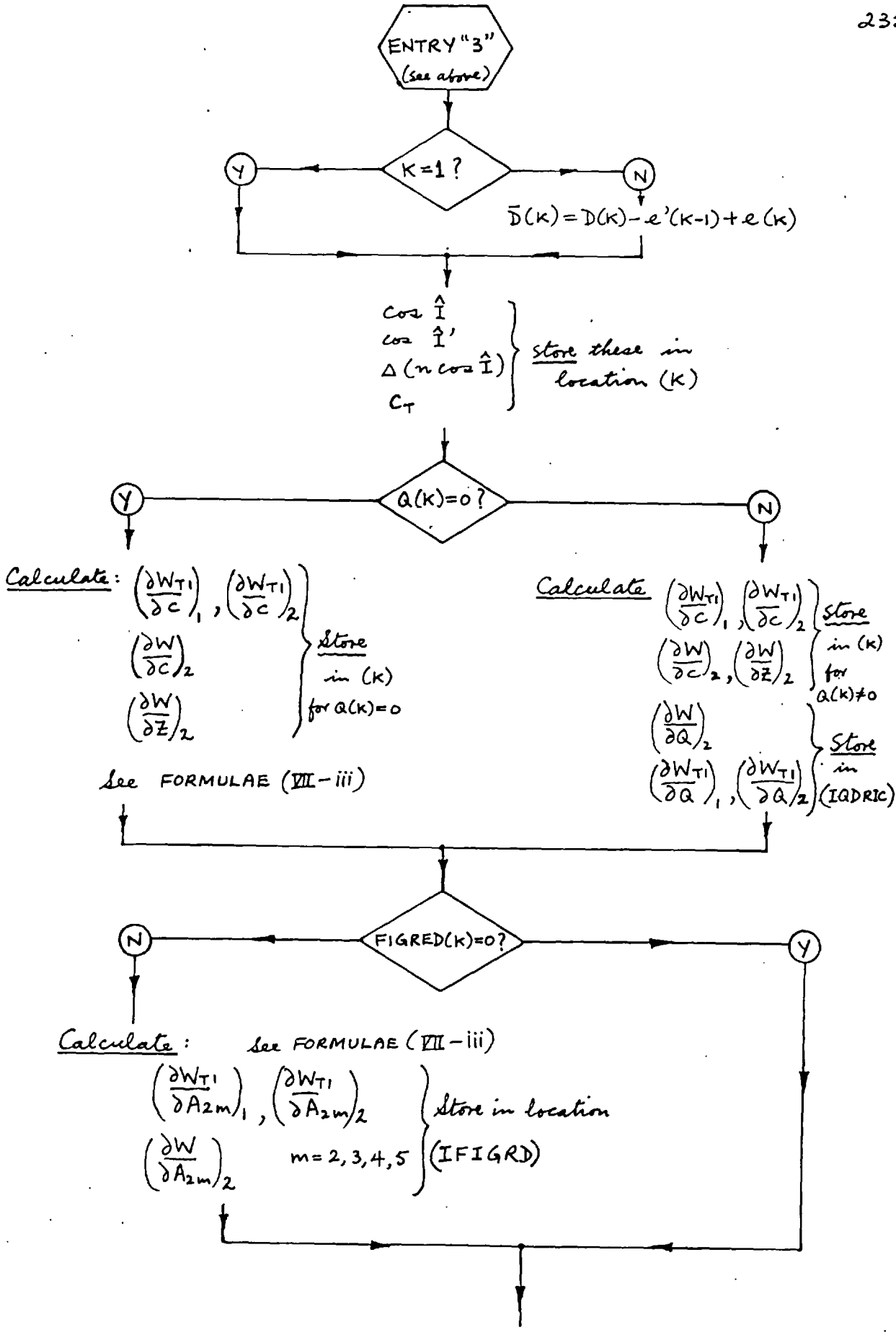
Q(K)=0?

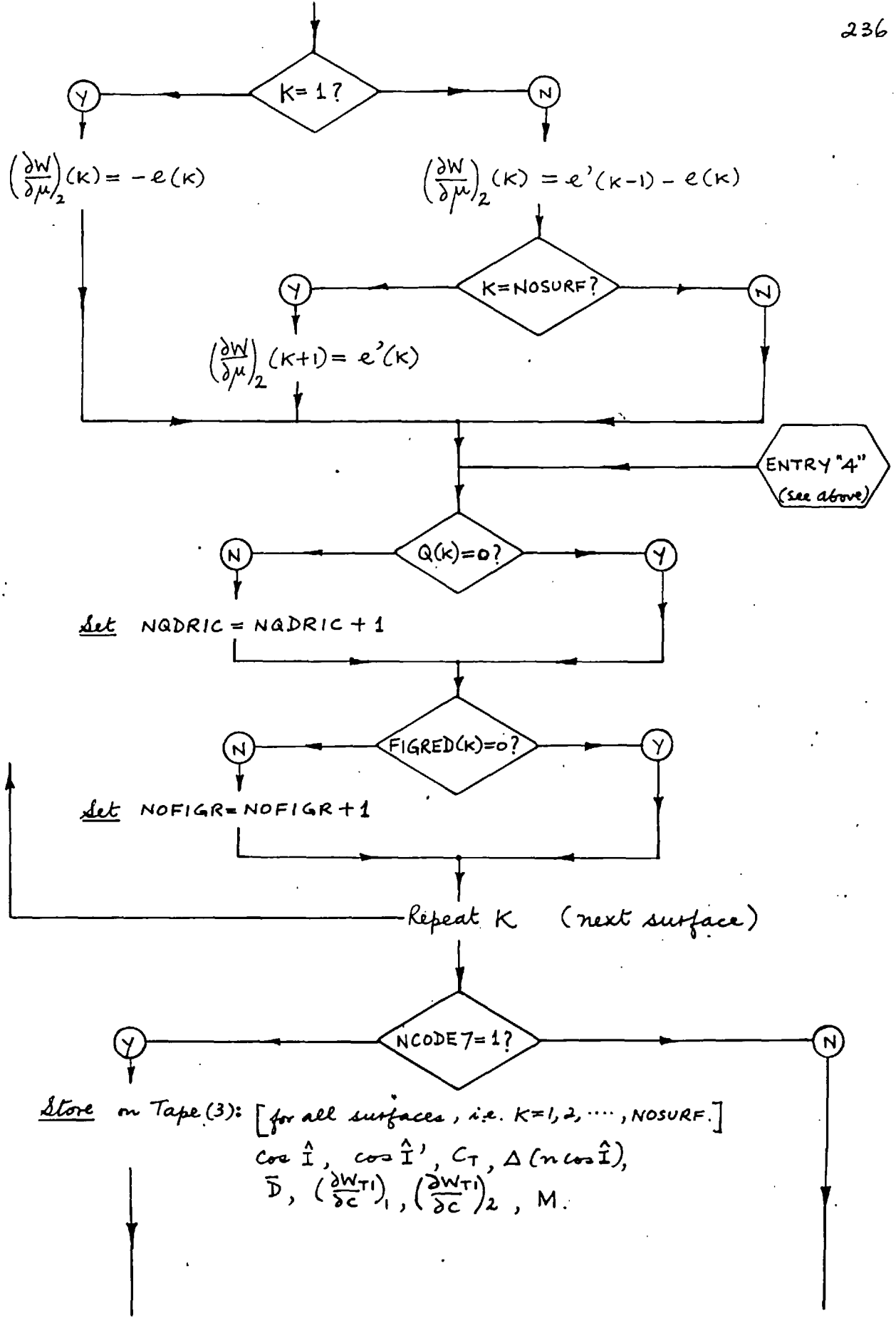
$$\left(\frac{\partial W_i}{\partial C}\right)(K) = \frac{(X^2+Y^2)\,\Delta(n\cos\hat{I})}{(1+\gamma)\sqrt{\alpha^2+\beta^2+\gamma^2}} - \left(\frac{\partial W}{\partial C}\right)_2$$

$$\left(\frac{\partial W_i}{\partial Z}\right)(K) = k\,\Delta(n\cos\hat{I}) - \left(\frac{\partial W}{\partial Z}\right)_2$$

PART "B"


NCODE8=1?
N
Y
$\frac{\partial W_i}{\partial Q} = \frac{c Z_s^2 \Delta (n \cos \hat{I})}{2 \cdot \sqrt{\alpha^2+\beta^2+\gamma^2}}$
$\frac{\partial W_i}{\partial Q} = \frac{c Z_s^2 \Delta (n \cos \hat{I})}{2 \cdot \sqrt{\alpha^2+\beta^2+\gamma^2}} - \left(\frac{\partial W}{\partial Q}\right)_2$ (KQPUPL)
Store $\frac{\partial W_i}{\partial Q}$ in location (IQDRIC)
FIGRED(K)=0?
Y
N
NCODE8=1?
N
Y
$\frac{\partial W_i}{\partial A_{2m}} = k \rho^m \Delta (n \cos \hat{I})$
$(m=2,3,4,5)$
$\frac{\partial W_i}{\partial A_{2m}} = k \rho^m \Delta (n \cos \hat{I}) - \left(\frac{\partial W}{\partial A_{2m}}\right)_2$ (KAPUPL)
$(m=2,3,4,5)$
Store $\frac{\partial W}{\partial A_{2m}}$ in location (IFIGRD)
TO: ENTRY "4" (see later)


ENTRY "3"
(see above)
K=1 ?
Y
N
$\bar{D}(K) = D(K) - e'(K-1) + e(K)$
$\cos \hat{I}$
$\cos \hat{I}'$
$\Delta(n \cos \hat{I})$
C_T
store these in location (K)
Q(K)=0 ?
Y
N
Calculate: $\left(\frac{\partial W_{TI}}{\partial C}\right)_1, \left(\frac{\partial W_{TI}}{\partial C}\right)_2$
$\left(\frac{\partial W}{\partial C}\right)_2$
$\left(\frac{\partial W}{\partial \bar{Z}}\right)_2$
Store in (K) for Q(K)=0
See FORMULAE (VII-iii)
Calculate $\left(\frac{\partial W_{TI}}{\partial C}\right)_1, \left(\frac{\partial W_{TI}}{\partial C}\right)_2$
$\left(\frac{\partial W}{\partial C}\right)_2, \left(\frac{\partial W}{\partial \bar{Z}}\right)_2$
Store in (K) for Q(K)≠0
$\left(\frac{\partial W}{\partial Q}\right)_2$
$\left(\frac{\partial W_{TI}}{\partial Q}\right)_1, \left(\frac{\partial W_{TI}}{\partial Q}\right)_2$
Store in (IQDRIC)
FIGRED(K)=0?
N
Y
Calculate: See FORMULAE (VII-iii)
$\left(\frac{\partial W_{TI}}{\partial A_{2m}}\right)_1, \left(\frac{\partial W_{TI}}{\partial A_{2m}}\right)_2$
$\left(\frac{\partial W}{\partial A_{2m}}\right)_2$
m=2,3,4,5
Store in location (IFIGRD)


K= 1?
Y
N
$\left(\frac{\partial W}{\partial \mu}\right)_2(K) = -e(K)$
$\left(\frac{\partial W}{\partial \mu}\right)_2(K) = e'(K-1) - e(K)$
K=NOSURF?
Y
N
$\left(\frac{\partial W}{\partial \mu}\right)_2(K+1) = e'(K)$
ENTRY "4" (see above)
Q(K)=0?
N
Y
Set NQDRIC = NQDRIC + 1
FIGRED(K)=0?
N
Y
Set NOFIGR=NOFIGR+1
Repeat K (next surface)
NCODE7=1?
Y
N
Store on Tape(3): [for all surfaces, i.e. K=1, 2, ···, NOSURF.]
$\cos \hat{I}$, $\cos \hat{I}'$, C_T, $\Delta(n \cos \hat{I})$,
$\bar{D}$, $\left(\frac{\partial W_{TI}}{\partial c}\right)_1$, $\left(\frac{\partial W_{TI}}{\partial c}\right)_2$, M.

Store on Tape (19): [for all surfaces, i.e. K=1, 2, ···, NOSURF]

$\left(\frac{\partial W}{\partial c}\right)_2, \left(\frac{\partial W}{\partial z}\right)_2, \left(\frac{\partial W}{\partial \mu}\right)_2$; also $\left(\frac{\partial W}{\partial \mu}\right)_2$(NOSURF+1)

Set L20 = I + (KFIELD−1). NOWAVE.

Store W_{pupil}(L20) = W

NCODE8=0?

(Y) → Store on Tape (4): [for all surfaces, i.e. K=62, ···, NOSURF]

$\frac{\partial W_i}{\partial c}, \frac{\partial W_i}{\partial z}, \frac{\partial W_i}{\partial \mu}$; also $\frac{\partial W_i}{\partial \mu}$(NOSURF+1)

Set L20 = KRAY + (I−1). A

W_{axial}(L20) = W

(N) → Store WAXIAS(KRAY) = W

Store the following on Tape (4): [for all surfaces, i.e. K=1, 2, ···, NOSURF]

$\frac{\partial W_i}{\partial c}, \frac{\partial W_i}{\partial z}, \frac{\partial W_i}{\partial \mu}$; also $\frac{\partial W_i}{\partial \mu}$(NOSURF+1)

WAXIAS(KRAY)

MQDRIC=0?

(Y) →

Set

LQDRIC = LQDRIC + MQDRIC

LFIGRD = LFIGRD + MFIGRD

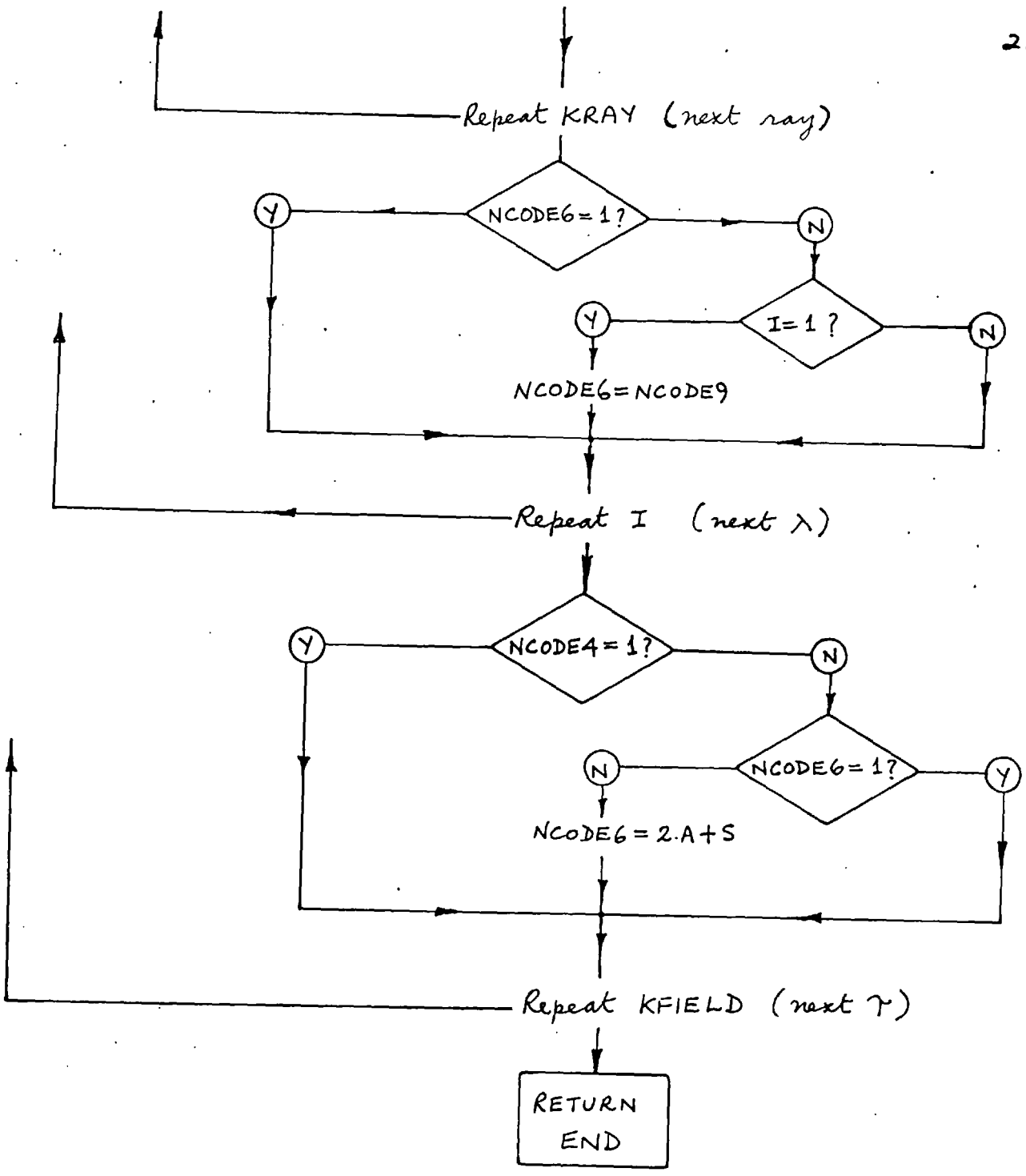
(N) → Store on Tape (4):

$\frac{\partial W_i}{\partial Q}$(L20) ; L20 = 1, 2, ···, MQDRIC

$\frac{\partial W_i}{\partial A_{2m}}$(L20) ; m = 2, 3, 4, 5.

Set LQDRIC = 0

LFIGRD = 0


Repeat KRAY (next ray)
NCODE6=1?
Y
N
I=1?
Y
N
NCODE6=NCODE9
Repeat I (next λ)
NCODE4=1?
Y
N
NCODE6=1?
N
Y
NCODE6=2.A+S
Repeat KFIELD (next γ)
RETURN
END

(VII-iv) Pupil ray subroutine

Referring to part A of the flow-diagram (VII-iv), the cycles for the field and wavelength calculate the starting data (L,M,N), (X,Y) for the pupil ray trace; the latter is carried out in the finite ray trace subroutine. The values of NCODE4,6,7 serve to identify that the finite rays to be traced are the pupil rays (see section VII-iii).

The pupil ray for each field passes through the centre of the corresponding elliptical pupil and this is described later in the Formulae (VII-iv). The starting data are stored in the location (L20) as required by the input of the finite ray trace subroutine.

The pupil ray data required by the part B of the pupil ray subroutine are stored in the locations (KK), (KKK1), (LL2), (LL3) and (LL5). Thus, for example, the location

$$KK = 120 + I + (KFIELD-1)\ (NOWAVE)$$

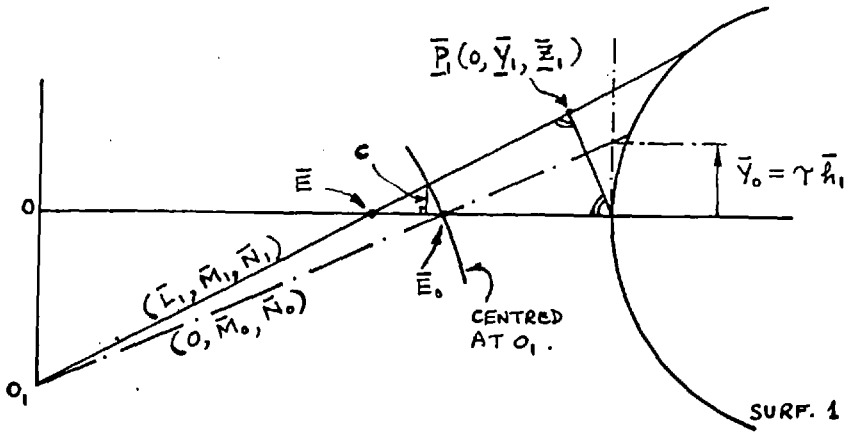
depends on the field and wavelength; the value of X(KK) is the value of the co-ordinate X of the pupil ray intersection point at the stop surface, corresponding to the particular values of the field and wavelength; this value of X(KK) has been stored in the finite ray trace specially for the present subroutine.

Part B of the flow-diagram (VII-iv) are presented in a detailed form; this is necessary since the various quantities involved in the calculations are obtained directly from the

finite ray trace subroutine; their dependence on the surface number, the wavelength and the field are such that their storage locations are required to be specifically defined.

A summary of the formulae employed in the computation are given in Formulae (VII-iv). The quantities such as $\bar{H}_s$, $\bar{H}_o$, etc. depend only on the field and the wavelength, and they may be conveniently stored in the location (KK1). The value of the distortion for each field is calculated only for the mean wavelength.

FORMULAE (VII - iv)



When the pupil has been fitted to the ellipse

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y-c}{b}\right)^2 = 1$$

the effective pupil ray passes through the points $\bar{E}$, $\bar{P}_1$. The pupil ray is traced using :

$$\begin{cases} \bar{L}_1 = 0 \\ \bar{M}_1 = \bar{M}_0 - (\mu_s)_0 \cdot c \\ \bar{N}_1 = +\sqrt{1 - \bar{M}_1^2} \\ \bar{X}_1 = 0 \\ \bar{Y}_1 = \bar{N}_1 \bar{Y}_0 + \left(\dfrac{1 + \bar{M}_1 \bar{M}_0 + \bar{N}_1 \bar{N}_0}{\bar{N}_1 + \bar{N}_0}\right) h_1 c \\ \bar{Z}_1 = -\dfrac{\bar{M}_1}{1 + \bar{N}_1} \cdot \bar{Y}_1 \end{cases}$$

where

$$\bar{N}_0 = \frac{1}{\sqrt{1 + (\bar{\mu}_1 \Upsilon)^2}}$$

$$\bar{M}_0 = -\bar{N}_0 (\bar{\mu}_1 \Upsilon)$$

$$\bar{Y}_0 = \bar{h}_1 \Upsilon$$

$$(\mu_s)_0 = \bar{N}_0 \mu_1$$

$$\bar{H}_S = \left[n(\bar{M} h + \bar{N} \bar{Y} u) \right]_{\text{at the stop surface.}}$$

$$(\bar{H}_0)_\lambda = \frac{n_\lambda}{n_{\bar{\lambda}}} \cdot H_{\bar{\lambda}} \cdot \tau$$

$$\bar{H}_0' = \frac{n'_{\bar{\lambda}}}{\bar{N}'_k} \left\{ (\bar{N}' \bar{Y} - \bar{M}' \bar{Z}) u_{\bar{\lambda}}' + \bar{M}' h_{\bar{\lambda}} \right\}_k \qquad (k = \text{last surface})$$

$$\text{Distortion} = \frac{\bar{H}_0' - \bar{H}_0}{\bar{H}_0} \qquad (\lambda = \bar{\lambda} \text{ only})$$

$$\delta \bar{H}_0' = (\bar{H}_0')_\lambda - (\bar{H}_0')_{\bar{\lambda}}$$

$$\frac{\tilde{N}'}{N'} = 1 + (\bar{M}')^2 \left(\frac{\delta \bar{H}_0'}{\bar{H}_0'} \right)$$

$$\frac{\tilde{M}'}{M'} = 1 - (\bar{N}')^2 \left(\frac{\delta \bar{H}_0'}{\bar{H}_0'} \right)$$

$$\therefore \delta \bar{H}_S' = \frac{\tilde{N}'}{N'} \cdot \left(\frac{\bar{H}_S}{\bar{H}_0'} \right) \cdot \delta \bar{H}_0'$$

$$\pi' = \frac{\tilde{N}'}{N'} \left(\frac{\bar{H}_S}{\bar{H}_0'} \right)$$

$$\delta \bar{y}_S' = \frac{n'_\lambda}{\bar{N}'_{\bar{\lambda}} (\bar{H}_S)_\lambda} \cdot \left[\bar{M}'_\lambda (\bar{N}'_{\bar{\lambda}} \bar{Y}_{\bar{\lambda}} - \bar{M}'_{\bar{\lambda}} \bar{Z}_{\bar{\lambda}}) - \bar{M}'_{\bar{\lambda}} (\bar{N}'_\lambda \bar{Y}_\lambda - \bar{M}'_\lambda \bar{Z}_\lambda) \right]_k$$

FLOW DIAGRAM (VII-iv)

Pupil Ray Subroutine

PART "A"

Set NCODE7 = 1
NCODE6 = 1
NCODE4 = NFIELD

Cycle KFIELD (KFIELD = 1, 2, ···, NFIELD)

Cycle I (I = 1, 2, ···, NOWAVE)

$$\bar{N}_0 = \frac{1}{\sqrt{1+(\bar{u}_1\tau)^2}}$$

$$\bar{M}_0 = -\bar{N}_0\,(\bar{u}_1\tau)$$

$$\bar{Y}_0 = \bar{h}_1\tau$$

$$(u_s)_0 = \bar{N}_0\,u_1$$

Set L20 = 30 + I + (KFIELD − 1). NOWAVE
KKI = I + (KFIELD − 1). NOWAVE

Store
Starting data for the pupil ray trace

X(L20) = 0
Y(L20) = $\bar{Y}_1$ see FORMULAE (VII-iv)
L(L20) = 0
M(L20) = $\bar{M}_0 - (u_s)_0 \cdot C_{pupil}(KKI)$
N(L20) = $+\sqrt{1-[M(L20)]^2}$

Repeat I (next λ)

Repeat KFIELD (next τ)

CALL FINITE RAY TRACE (VII-iii) SUBROUTINE →

PART "B" (P.T.O.)

PART "B"

Cycle KFIELD (KFIELD = 1, 2, ···· , NFIELD)

Cycle I (I = 1, 2, ···· , NOWAVE)

Set
$$\begin{cases} INDEX = (I-1).(NOSURF+1) \\ KK1 = I + .(KFIELD-1).NOWAVE \\ LL1 = NOSTOP + INDEX \\ KK = 120 + I + (KFIELD-1).NOWAVE \end{cases}$$

$$\bar{H}_s(KK1) = E(LL1).\left[M(KK).(h_{stop})_{\bar{\lambda}} + N(KK).Y(KK)(u_{stop})_{\bar{\lambda}}\right]$$

$$\bar{H}_o(KK1) = \frac{E(INDEX+1)}{E(1)} \cdot H_{\bar{\lambda}} \cdot \Upsilon$$

Set $KKK1 = 60 + I + (KFIELD-1).NOWAVE + 1$

$LL5 = KKK1 - 1$

Calculate:

$$\bar{H}_o'(KK1) = \frac{E(NOSURF+1)}{N(KKK1)}\left\{\left[N(KKK1).Y(LL5) - M(KKK1).Z(LL5)\right](u_{image})_{\bar{\lambda}} + M(KKK1).(h_{image})_{\bar{\lambda}}\right\}$$

I = 1 ?

(Y) → Calculate:

DISTORTION (KFIELD) $= \left(\frac{\bar{H}_o' - \bar{H}_o}{\bar{H}_o}\right)_{(KK1)}$

Store this on Tape (20)

(N) →

Set $JJ = 1 + (KFIELD-1).NOWAVE$

Calculate:
$$\delta\bar{H}_o'(KK1) = \bar{H}_o'(KK1) - \bar{H}_o'(JJ)$$
$$\frac{\tilde{N}'}{\bar{N}'}(KK1) = 1 + [M(KKK1)]^2 \cdot \frac{\delta\bar{H}_o'(KK1)}{\bar{H}_o'(KK1)}$$
$$\frac{\tilde{M}'}{\bar{M}'}(KK1) = 1 - [N(KKK1)]^2 \cdot \frac{\delta\bar{H}_o'(KK1)}{\bar{H}_o'(KK1)}$$

Calculate:

$$\delta \bar{H}_s'(KK1) = \frac{\tilde{N}'}{\bar{N}'}(KK1) \cdot \frac{\bar{H}_s(KK1)}{\bar{H}_o'(KK1)} \cdot \delta \bar{H}_o'(KK1)$$

$$\pi'(KK1) = \frac{\tilde{N}'}{\bar{N}'}(KK1) \cdot \frac{\bar{H}_s(KK1)}{\bar{H}_o'(KK1)}$$

Set: $LL2 = INDEX + NOSURF + 1$

$LL3 = 90 + I$

Calculate:

$$\delta \bar{y}_s'(KK1) = \frac{E(LL2)}{N(LL3) \cdot \bar{H}_s(KK1)} \cdot \Big\{ M(KKK1)\big[N(LL3) \cdot Y(LL3) - M(LL3) \cdot Z(LL3)\big] - M(LL3)\big[N(KKK1) \cdot Y(KKK1-1) - M(KKK1) \cdot Z(KKK1-1)\big] \Big\}$$

Repeat I (next λ)

Repeat KFIELD (next γ)

RETURN

END.

(VII-v) T and $\bar{T}$-rays subroutine.

In this subroutine, two close (paraxial) tangential rays are traced relative to the pupil ray for each field and wavelength. The tangential ray traced from the object space is referred to as the T-ray; the other tangential ray is the $\bar{T}$ ray, which is traced in the reverse direction (right to left) starting from the centre of the effective exit pupil. The starting data required, namely the values of (u_T, h_T) at the first surface and $(\bar{u}'_T, \bar{h}_T)$ at the last surface, are given later in Formulae (VII-v).

Referring to the flow-diagram (VII-v), the pupil ray data required for the T and $\bar{T}$ - ray trace are read-in via Tape (3). The part A is concerned with the $\bar{T}$-ray trace, and the values of $(\bar{u}_T, \bar{h}_T)$ at each surface are stored for use in the part B; the latter deals first with the calculations of (u_T, h_T) of the T-ray, and then follow the computations of $\partial W_{T1}/\partial p$ and $\partial \bar{W}_{T1}/\partial p$, which require the data from the T and $\bar{T}$-ray trace.

The counters employed for the quadric and the figured surfaces are the same as described in the section (VII-iii) for the finite ray trace. At the end of the surface cycle of part B, the values of $\partial W_{T1}/\partial p$ and $\partial \bar{W}_{T1}/\partial p$ are stored on Tapes (2) and (19); Tape (19) will be called by the extra-axial ray subroutine, and Tape (2) is specially adapted for the print-out subroutine. Part C of the flow-diagram (VII-V) deals with the distortion derivatives, $\partial D/\partial p$'s, which are computed for different fields,

but only for $\lambda=\bar{\lambda}$.

The quantities calculated in the T and $\bar{T}$-rays sub-routine are summarised in Formulae (VII-V) which follows.

FORMULAE (VII-V)

Opening formulae.

T-ray:
$$(u_T)_1 = \bar{N}_1 \left(\frac{\bar{H}_s}{\bar{H}_0}\right) u_1$$
$$(h_T)_1 = \left(\frac{\bar{H}_s}{\bar{H}_0}\right) \cdot \left(\frac{h_1 - u_1 \bar{Z}_1}{\cos \bar{I}_1}\right)$$

$\bar{T}$-ray:
$$\left.\begin{aligned}(\bar{u}_T')_k &= -\bar{M}_k' \\ (\bar{h}_T)_k &= \left(\frac{\bar{Y}_k}{\cos \bar{I}_k'}\right)\end{aligned}\right\} \quad k = \text{last surface.}$$

Tangential Ray-tracing formulae.

T-ray:
$$(u_T') = \left\{n(\cos \bar{I})\, u_T + h_T . c_T . \Delta(n \cos \bar{I})\right\} / n' \cos \bar{I}'$$
$$(h_T)_{+1} = \frac{h_T \cos \bar{I} - \bar{D}' u_T'}{(\cos \bar{I})_{+1}}$$

$\bar{T}$-ray:
$$(u_T) = \left\{n' (\cos \bar{I}')\, u_T' - h_T . c_T . \Delta(n \cos \bar{I})\right\} / n \cos \bar{I}$$
$$(h_T)_{-1} = \frac{h_T \cos \bar{I} + \bar{D}\, u_T}{(\cos \bar{I})_{-1}}$$

Paraxial parameter derivatives.

(a) Spherical surfaces.

$p = c$
$$\frac{\partial W_{T1}}{\partial c} = \bar{Y} h_T \Delta(n \cos \bar{I}) + \tfrac{1}{2}(\bar{Y}^2 + \bar{Z}^2)(n \cos \bar{I}) \Delta(u_T)$$

$p = z$
$$\frac{\partial W_{T1}}{\partial z} = \Delta(n \bar{M} u_T)$$

$p = \mu$
$$\frac{\partial W_{T1}}{\partial \mu} = (h_T \cdot \sin \bar{I}')_{-1} - (h_T \sin \bar{I})$$

(b) Aspheric surfaces.

$p=c$ $\quad \dfrac{\partial W_{T1}}{\partial c} = \dfrac{\bar{Y}\,h_T}{\bar{\gamma}(\bar{\beta}^2+\bar{\gamma}^2)}\,\Delta(n\cos\bar{I}) + \dfrac{\bar{Y}^2}{(1+\bar{\gamma})\bar{\gamma}}\,\Delta(n\bar{M}\,u_T)$

$p=Q$ $\quad \dfrac{\partial W_{T1}}{\partial Q} = \dfrac{(\bar{Y}c)^3}{2\bar{\gamma}(\bar{\beta}^2+\bar{\gamma}^2)}\,\Delta(n\cos\bar{I}) + \dfrac{c\bar{Z}_S^2}{2\bar{\gamma}}\,\Delta(n\bar{M}\,u_T)$

$p=A_{2m}$ $\quad \dfrac{\partial W_{T1}}{\partial A_{2m}} = \dfrac{2m\,\bar{Y}\bar{h}^2}{\rho_{max}^2}\cdot \rho^{(m-1)}\,h_T\,\Delta(n\cos\bar{I}) + \rho^m\,\Delta(n\bar{M}\,u_T)$

$(m=2,3,4,5)$

$p=Z$ $\quad \dfrac{\partial W_{T1}}{\partial Z} = \Delta(n\bar{M}\,u_T)$

$p=\mu$ $\quad \dfrac{\partial W_{T1}}{\partial \mu} = (h_T \sin\bar{I}') - (h_T\sin\bar{I})_{+1}$

The values of $\dfrac{\partial \bar{W}_{T1}}{\partial p}$ are obtained by replacing (h_T, u_T) by $(\bar{h}_T, \bar{u}_T)$ in the above formulae (for both (a) and (b)).

$$\bar{H}_T' = \bar{N}_k'\,n_k'\,(\bar{M}'h_T'\cos\bar{I}' + \bar{\gamma}\,u_T')_k$$

$$\rho' = \bar{H}_S / \bar{H}_T' \qquad (k=\text{last surf.})$$

Distortion derivatives.

$$\frac{\partial D}{\partial p} = \frac{\partial}{\partial p}\left(\frac{\delta\bar{H}_o'}{\bar{H}_o}\right) = \left(\frac{\bar{H}_o'}{\bar{H}_o}\right)\left\{-\frac{1}{\bar{H}_T'}\cdot\frac{\partial W_{T1}}{\partial p} + \frac{1}{H}\cdot\frac{\partial_1 W_{11}}{\partial p}\right\}$$

with $p = c, z, \mu, A_{2m}, Q.$

FLOW DIAGRAM (VI - V)

T and $\bar{T}$ Rays Subroutine

Read Tape (3): $\begin{cases} \frac{\partial_1 W_{11}}{\partial Z}(K),\ \frac{\partial_1 W_{11}}{\partial c}(K) \quad ; \ K = 1, 2, \cdots, NOSURF. \\ \frac{\partial_1 W_{11}}{\partial \mu}(K1) \ ; \quad K1 = 1, 2, \cdots, (NOSURF+1). \end{cases}$

↓

Set: $\begin{cases} LQDRIC = 0 \\ LFIGRD = 0 \\ LQ = 0 \\ LF = 0 \end{cases}$

↓

Cycle KFIELD (KFIELD = 1, 2, ···, NFIELD)

Cycle I (I = 1, 2, ···, NOWAVE)

↓

Set KK1 = I + (KFIELD − 1). NOWAVE.

↓

Read Tape (3): $\begin{cases} \cos \hat{I}(K),\ \cos \hat{I}'(K),\ C_T(K),\ \Delta(n \cos I)_K, \\ \bar{D}_T(K),\ \left(\frac{\partial W_{T1}}{\partial c}\right)_1(K),\ \left(\frac{\partial W_{T1}}{\partial c}\right)_2(K) \quad ; \ K = 1, 2, \cdots, NOSURF. \\ M(K1) \ ; \ K1 = 1, 2, \cdots, (NOSURF+1). \end{cases}$

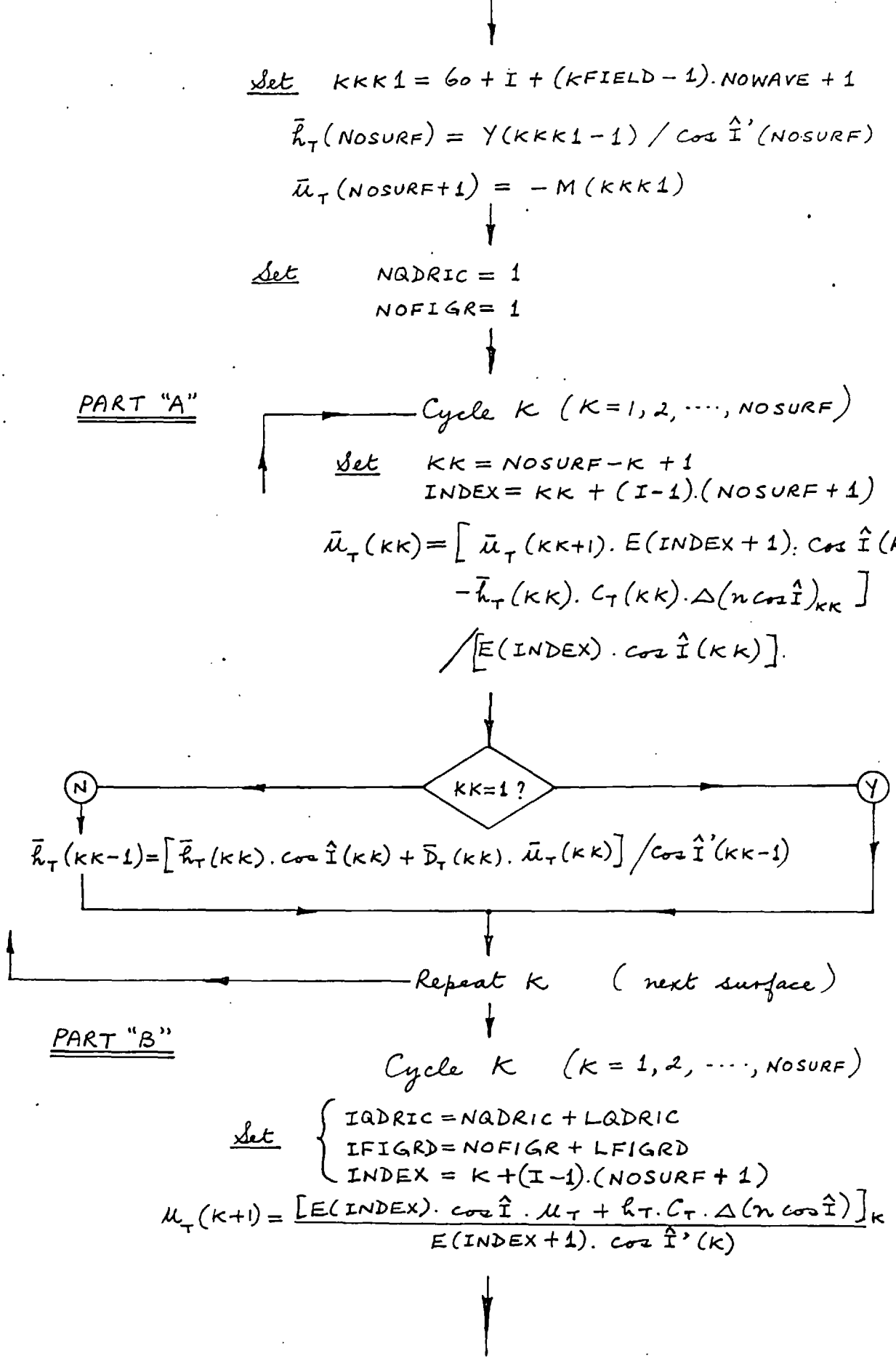
↓

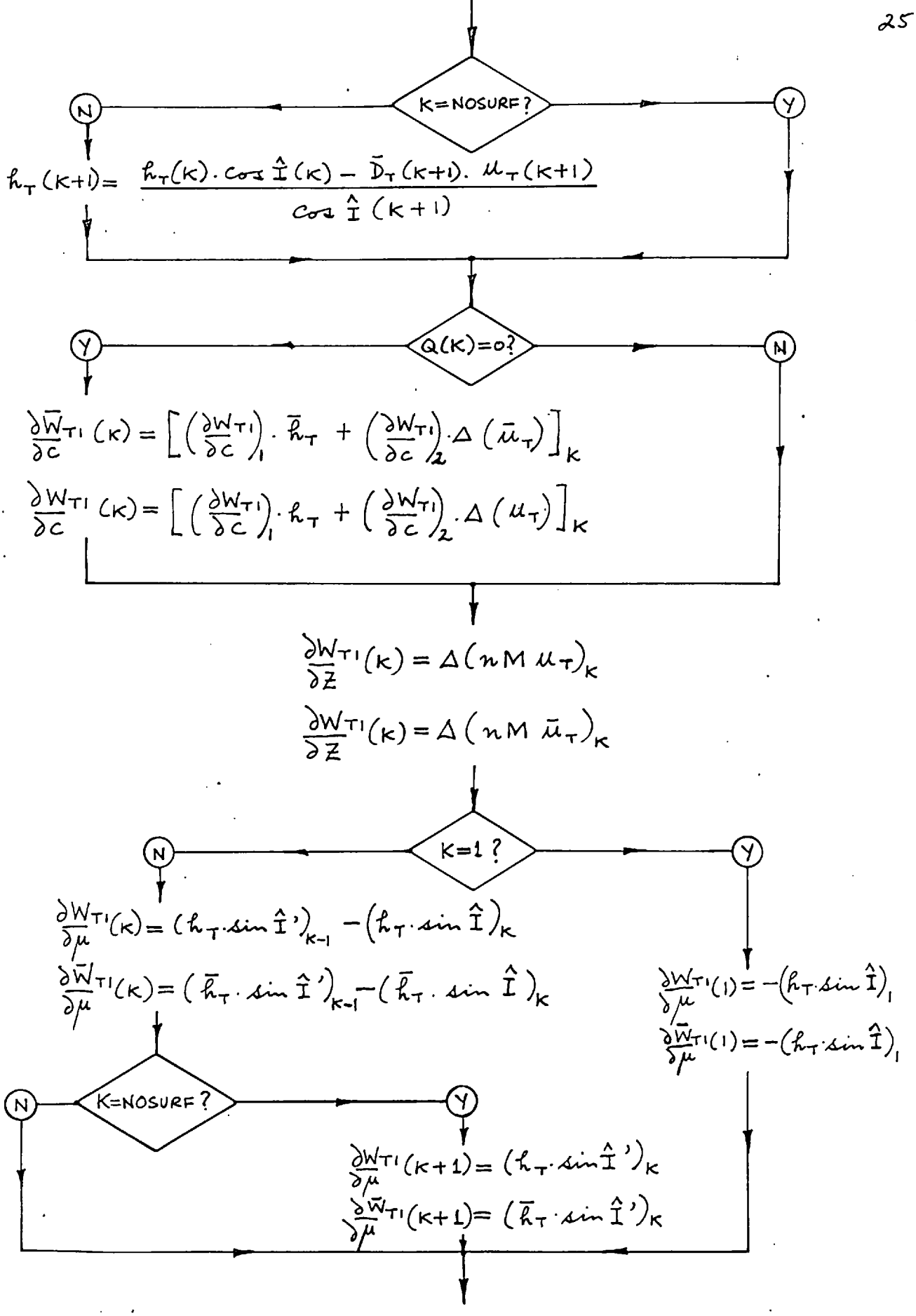
Set L20 = 30 + I + (KFIELD − 1). NOWAVE

Calculate: $\mu_T(1) = N(L20) \cdot \mu_1 \cdot \bar{H}_S(KK1) \,/\, \bar{H}_0(KK1)$

$$h_T(1) = \frac{[h_1 - \mu_1 \cdot Z(L20)] \cdot \bar{H}_S(KK1)}{\bar{H}_0(KK1) \cdot \cos \hat{I}(1)}$$

↓


Set KKK1 = 60 + I + (KFIELD − 1).NOWAVE + 1
$\bar{h}_T(NOSURF) = Y(KKK1-1) / \cos\hat{I}'(NOSURF)$
$\bar{u}_T(NOSURF+1) = -M(KKK1)$
Set NQDRIC = 1
NOFIGR = 1
PART "A"
Cycle K (K = 1, 2, ···, NOSURF)
Set KK = NOSURF − K + 1
INDEX = KK + (I − 1).(NOSURF + 1)
$\bar{u}_T(KK) = [\bar{u}_T(KK+1).E(INDEX+1).\cos\hat{I}(KK) - \bar{h}_T(KK).C_T(KK).\Delta(n\cos\hat{I})_{KK}] / [E(INDEX).\cos\hat{I}(KK)]$
KK=1 ?
N
Y
$\bar{h}_T(KK-1) = [\bar{h}_T(KK).\cos\hat{I}(KK) + \bar{D}_T(KK).\bar{u}_T(KK)] / \cos\hat{I}'(KK-1)$
Repeat K (next surface)
PART "B"
Cycle K (K = 1, 2, ···, NOSURF)
Set IQDRIC = NQDRIC + LQDRIC
IFIGRD = NOFIGR + LFIGRD
INDEX = K + (I − 1).(NOSURF + 1)
$u_T(K+1) = \frac{[E(INDEX).\cos\hat{I}.u_T + h_T.C_T.\Delta(n\cos\hat{I})]_K}{E(INDEX+1).\cos\hat{I}'(K)}$


K=NOSURF?
N
Y
$h_T(K+1) = \frac{h_T(K)\cdot\cos\hat{I}(K) - \bar{D}_T(K+1)\cdot\mu_T(K+1)}{\cos\hat{I}(K+1)}$
Q(K)=0?
Y
N
$\frac{\partial\bar{W}_{TI}}{\partial c}(K) = \left[\left(\frac{\partial W_{TI}}{\partial c}\right)_1\cdot\bar{h}_T + \left(\frac{\partial W_{TI}}{\partial c}\right)_2\cdot\Delta(\bar{\mu}_T)\right]_K$
$\frac{\partial W_{TI}}{\partial c}(K) = \left[\left(\frac{\partial W_{TI}}{\partial c}\right)_1\cdot h_T + \left(\frac{\partial W_{TI}}{\partial c}\right)_2\cdot\Delta(\mu_T)\right]_K$
$\frac{\partial W_{TI}}{\partial Z}(K) = \Delta(nM\,\mu_T)_K$
$\frac{\partial W_{TI}}{\partial Z}(K) = \Delta(nM\,\bar{\mu}_T)_K$
K=1 ?
N
Y
$\frac{\partial W_{TI}}{\partial\mu}(K) = (h_T\cdot\sin\hat{I}')_{K-1} - (h_T\cdot\sin\hat{I})_K$
$\frac{\partial\bar{W}_{TI}}{\partial\mu}(K) = (\bar{h}_T\cdot\sin\hat{I}')_{K-1} - (\bar{h}_T\cdot\sin\hat{I})_K$
$\frac{\partial W_{TI}}{\partial\mu}(1) = -(h_T\cdot\sin\hat{I})_1$
$\frac{\partial\bar{W}_{TI}}{\partial\mu}(1) = -(h_T\cdot\sin\hat{I})_1$
K=NOSURF?
N
Y
$\frac{\partial W_{TI}}{\partial\mu}(K+1) = (h_T\cdot\sin\hat{I}')_K$
$\frac{\partial\bar{W}_{TI}}{\partial\mu}(K+1) = (\bar{h}_T\cdot\sin\hat{I}')_K$

Q(K)=0?

(N)

$$\frac{\partial \bar{W}_{TI}}{\partial c}(K) = \left[\left(\frac{\partial W_{TI}}{\partial c} \right)_1 \cdot \bar{h}_T + \left(\frac{\partial W_{TI}}{\partial c} \right)_2 \cdot \left(\frac{\partial \bar{W}_{TI}}{\partial \bar{z}} \right) \right]_K$$

$$\frac{\partial W_{TI}}{\partial c}(K) = \left[\left(\frac{\partial W_{TI}}{\partial c} \right)_1 \cdot h_T + \left(\frac{\partial W_{TI}}{\partial c} \right)_2 \cdot \left(\frac{\partial W_{TI}}{\partial z} \right) \right]_K$$

$$\frac{\partial \bar{W}_{TI}}{\partial Q}(IQDRIC) = \left[\left(\frac{\partial W_{TI}}{\partial Q} \right)_1 (IQDRIC) \cdot \bar{h}_T(K) + \left(\frac{\partial W_{TI}}{\partial Q} \right)_2 (IQDRIC) \cdot \frac{\partial \bar{W}_{TI}}{\partial \bar{z}}(K) \right]$$

$$\frac{\partial W_{TI}}{\partial Q}(IQDRIC) = \left[\left(\frac{\partial W_{TI}}{\partial Q} \right)_1 (IQDRIC) \cdot h_T(K) + \left(\frac{\partial W_{TI}}{\partial Q} \right)_2 (IQDRIC) \cdot \frac{\partial W_{TI}}{\partial z}(K) \right]$$

Set NQDRIC = NQDRIC + 1

(Y)

FIGRED(K)=0?

(N)

$$\frac{\partial \bar{W}_{TI}}{\partial A_{2m}}(IFIGRD) = \left(\frac{\partial W_{TI}}{\partial A_{2m}} \right)_1 (IFIGRD) \cdot \bar{h}_T(K) + \left(\frac{\partial W_{TI}}{\partial A_{2m}} \right)_2 (IFIGRD) \cdot \frac{\partial \bar{W}_{TI}}{\partial \bar{z}}(K)$$

$$\frac{\partial W_{TI}}{\partial A_{2m}}(IFIGRD) = \left(\frac{\partial W_{TI}}{\partial A_{2m}} \right)_1 (IFIGRD) \cdot h_T(K) + \left(\frac{\partial W_{TI}}{\partial A_{2m}} \right)_2 (IFIGRD) \cdot \frac{\partial W_{TI}}{\partial z}(K)$$

$(m = 2, 3, 4, 5)$

Set NOFIGR = NOFIGR + 1

(Y)

Repeat K (next surface)

Store on Tape (19) and on Tape (2):

$\frac{\partial W_{TI}}{\partial c}(K)$, $\frac{\partial W_{TI}}{\partial z}(K)$, $\frac{\partial \bar{W}_{TI}}{\partial c}(K)$, $\frac{\partial \bar{W}_{TI}}{\partial \bar{z}}(K)$; $K = 1, 2, \ldots, NOSURF.$

$\frac{\partial W_{TI}}{\partial \mu}(K1)$, $\frac{\partial \bar{W}_{TI}}{\partial \mu}(K1)$; $K1 = 1, 2, \ldots, (NOSURF+1).$

Calculate:

$$\bar{H}_T'(KK1) = N(KKK1) \cdot E(INDEX+1)\Big[M(KKK1)\cdot h_T(NOSURF)\cdot Cos\,\hat{I}'(NOSURF) + Y(KKK1-1)\cdot u_T(NOSURF+1)\Big]$$

$$\rho'(KK1) = \bar{H}_S'(KK1) \,/\, \bar{H}_T'(KK1)$$

$$RR1 = \bar{H}_o'(KK1) \,/\, \bar{H}_o(KK1)$$

I = 1 ? (Y / N)

N → TO: ENTRY "1" (see later)

Y → Set { NQDRIC = 1, NOFIGR = 1 }

Part "C"

Cycle K (K = 1, 2, ···, NOSURF)

Set { IQDRIC = NQDRIC + LQ, IFIGRD = NOFIGR + LF }

$$\frac{\partial D}{\partial Z}(K) = RR1\left[-\frac{\partial W_{TI}}{\partial Z}(K) \,/\, \bar{H}_T'(KK1) + \frac{\partial_1 W_{II}}{\partial Z}(K) \,/\, H\right]$$

$$\frac{\partial D}{\partial \mu}(K) = RR1\left[-\frac{\partial W_{TI}}{\partial \mu}(K) \,/\, \bar{H}_T'(KK1) + \frac{\partial_1 W_{II}}{\partial \mu}(K) \,/\, H\right]$$

$$\frac{\partial D}{\partial c}(K) = RR1\left[-\frac{\partial W_{TI}}{\partial c}(K) \,/\, \bar{H}_T'(KK1) + \frac{\partial_1 W_{II}}{\partial c}(K) \,/\, H\right]$$

Q(K) = 0 ? (N / Y)

N → $\frac{\partial D}{\partial Q}(IQDRIC) = RR1\cdot\left[-\frac{\partial W_{TI}}{\partial Q}(IQDRIC) \,/\, \bar{H}_T'(KK1)\right]$

Set NQDRIC = NQDRIC + 1

FIGRED(K) = 0 ? (N / Y)

N → $\frac{\partial D}{\partial A_{2m}}(IFIGRD) = RR1\cdot\left[-\frac{\partial W_{TI}}{\partial A_{2m}}(IFIGRD) \,/\, \bar{H}_T'(KK1)\right]$ (m = 2, 3, 4, 5)

Set NOFIGR = NOFIGR + 1

Repeat K (next surface)

P.T.O.

Calculate:

$$\frac{\partial D}{\partial \mu}(NOSURF+1) = RR1 \cdot \left[-\frac{\partial W_{TI}}{\partial \mu}(NOSURF+1) \Big/ \bar{H}_T'(KK1) + \frac{\partial_1 W_{II}}{\partial \mu}(NOSURF+1) \Big/ H \right]$$

Store on Tape (21) and Tape (20):

$\frac{\partial D}{\partial z}(K)$, $\frac{\partial D}{\partial c}(K)$; $K = 1, 2, \cdots, NOSURF$.

$\frac{\partial D}{\partial \mu}(K1)$; $K1 = 1, 2, \cdots, (NOSURF+1)$

MQDRIC = 0?

(N)

Store on Tape (20):

$\frac{\partial D}{\partial Q}(K')$; $\begin{cases} K' = K + LQ \\ K = 1, 2, \cdots, MQDRIC \end{cases}$

Set $LQ = LQ + MQDRIC$

(Y)

MFIGRD = 0?

(N)

Store on Tape (20):

$\frac{\partial D}{\partial A_{2m}}(K')$; $\begin{cases} K' = K + LF \\ K = 1, 2, \cdots, MFIGRD. \\ \cdot (m = 2, 3, 4, 5) \end{cases}$

Set $LF = LF + MFIGRD$

(Y)

ENTRY "1" (see above)

Set $\begin{cases} LQDRIC = LQDRIC + MQDRIC \\ LFIGRD = LFIGRD + MFIGRD \end{cases}$

Repeat I (next λ)

Repeat KFIELD (next τ)

Rewind Tapes (19) and (2)

RETURN

END.

(VII-vi) Axial ray subroutine.

Referring to the flow-diagram (VII-vi), part A, the starting data required for an axial ray trace are of two types :

(i) the co-ordinates (X,Y) of the point $\underline{P}_1$ and the direction cosines (L,M,N) are the opening values for the axial ray trace (see equations 4-3, 4-4 and Fig.4-2), and are calculated and stored for each ray and wavelength. The ray pattern employed is that which has been described in Chapter 5; and for axial object points, the value of S is set to zero. The value of NCODE6 specifies the number of aperture rays to be traced in the azimuth $\phi=0$; for flexibility, NCODE6 is set to the value A for $\lambda = \bar{\lambda}$ and the value B for $\lambda \neq \bar{\lambda}$ $(B \leqslant A)$.

(ii) the code values, namely of the NCODE4, 7,8,9 are set in accordance with that defined in section (VII-iii) for the finite ray trace.

The values of $\partial W/\partial p$ for the axial ray are calculated in the finite ray trace subroutine, and are stored on Tape (4). The paraxial derivatives $\partial_o W_{20}/\partial p$ and $\partial_1 W_{11}/\partial p$ for each wavelength are read in via Tape (1). Part B of the flow-diagram (VII-vi) calculates the image space quantities which are described in detail in Chapter 4, part (a). The values of (L,M,N), (X,Y,Z)required are those in the image space and at the last surface respectively; these are stored in the location (LL_{10}) by the finite ray trace subroutine.

Part C of the flow-diagram (VII-vi) deals with the storage and computation of $\frac{\partial W_o}{\partial p}$, $\left(\frac{\partial W}{\partial p}\right)_\lambda$ and $\frac{\partial W_i}{\partial p}$, and also the values of the axial ray wave-aberration. The counters employed for the quadric and figured surfaces are those which have been described in the section (VII-iii). The values of $\frac{\partial W_o}{\partial p}$ are stored on to Tape (20) for the optimisation stage (see Chapter 8) and the values of $\left(\frac{\partial W}{\partial p}\right)_\lambda$, $\left(\frac{\partial W_i}{\partial p}\right)$ are stored on Tape (21) for the print-out subroutine.

The quantities calculated in the present subroutine are summarised in Formulae (VII-vi) which follows.

FORMULAE (VII - vi)

Image space quantities.

$$H_o' = \frac{n_k'}{N_k'} \left\{ (N'Y - M'Z)u' + M'h \right\}_k \qquad k = \text{last surface.}$$

$$\tilde{H}_o' = (n_k')_{\bar{\lambda}} \left\{ (\bar{h}_k)_\lambda \cdot (u_k')_{\bar{\lambda}} - (h_k)_{\bar{\lambda}} \cdot (\bar{u}_k')_\lambda \right\}$$

$$H_\lambda = \frac{n_\lambda}{n_{\bar{\lambda}}} H_{\bar{\lambda}}$$

$$\delta H_s' = \frac{H_\lambda \cdot H_o'}{\tilde{H}_o'}$$

$$\delta\Omega' = -\left(\frac{1}{1+N_k'}\right) \left\{ y' - M'(M'y + N'z') \right\}_k \cdot \delta H_s'$$

$$y'_{axial} = y' + M'\left(\frac{\bar{u}' \delta\Omega'}{H_\lambda}\right) \; ; \; z'_{axial} = z' + N'\left(\frac{\bar{u}' \cdot \delta\Omega'}{H_\lambda}\right)$$

Note the values of y' and z' are calculated in PART "C", FLOW DIAG. (VII - iii).

$$\frac{\partial W}{\partial (_oW_{20})_{\bar{\lambda}}} = -\frac{n_\lambda'}{n_{\bar{\lambda}}'} \left\{ y'^2_{axial} + z'^2_{axial} \right\}$$

$$\frac{\partial W}{\partial a_o} = -y \cdot \delta H_s' \quad \text{(derivative w.r.t. the axial aperture)}$$

Finite ray parameter derivatives : Axial images.

The formulae for $\frac{\partial W_i}{\partial p}$ are those which have been given in Formulae (VII - iii).

$$\left.\begin{aligned} \frac{\partial W_o}{\partial p} &= \frac{\partial W_i}{\partial p} - \frac{\delta H_s'}{H_\lambda} \cdot \left(\frac{\partial_1 W_{11}}{\partial p}\right) \cdot y'_{axial} \\ \left(\frac{\partial W}{\partial p}\right)_\lambda &= \frac{\partial W_o}{\partial p} + \frac{\partial W}{\partial (_oW_{20})_{\bar{\lambda}}} \cdot \left(\frac{\partial_o W_{20}}{\partial p}\right)_{\bar{\lambda}} \end{aligned}\right\} \; p = c, z, \mu.$$

$$\frac{\partial W_o}{\partial p} = \frac{\partial W_i}{\partial p} = \left(\frac{\partial W}{\partial p}\right)_\lambda \; ; \; p = \alpha, A_{2m} \quad (m = 2, 3, 4, 5)$$

– see also equations (6-15) and (6-19).

FLOW DIAGRAM (VII - vi)

Axial Ray Subroutine

PART "A"

Set { NCODE4 = 1
NCODE6 = A

Cycle I (I = 1, 2, ···, NOWAVE)

Cycle KRAY (KRAY = 1, 2, ···, NCODE6)

Set L20 = 30 + KRAY + (I - 1) . NCODE6

I = 1 ?

(Y) $y_j = \sqrt{\frac{(A+1) - KRAY}{A}}$

(N) $y_j = \sqrt{\frac{(A+1) - JB(KRAY)}{A}}$ (see Section VII - i)

X(L20) = 0

Y(L20) = $h_1 \cdot y_j$

L(L20) = 0

M(L20) = $-u_1 \cdot y_j$

N(L20) = $+\sqrt{1 - [M(L20)]^2}$

Repeat KRAY (next ray).

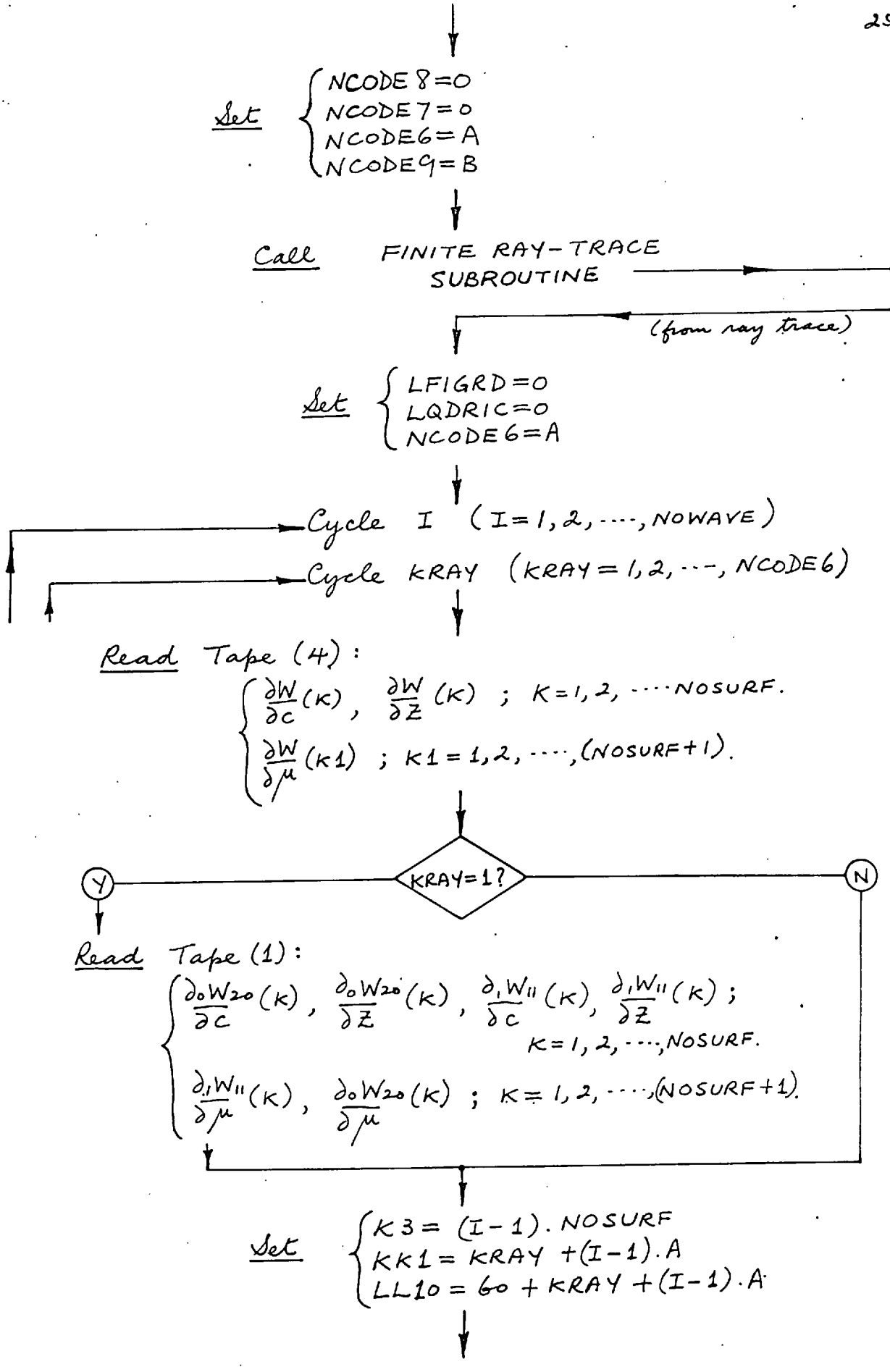
I = 1 ?

(Y) Set NCODE6 = B

(N)

Repeat I (next λ)

P.T.O.


Set
NCODE 8 = 0
NCODE 7 = 0
NCODE 6 = A
NCODE 9 = B
Call FINITE RAY-TRACE SUBROUTINE
(from ray trace)
Set
LFIGRD = 0
LQDRIC = 0
NCODE 6 = A
Cycle I (I = 1, 2, ···, NOWAVE)
Cycle KRAY (KRAY = 1, 2, ···, NCODE6)
Read Tape (4):
∂W/∂c (K), ∂W/∂Z (K) ; K = 1, 2, ···· NOSURF.
∂W/∂μ (K1) ; K1 = 1, 2, ····, (NOSURF+1).
KRAY=1?
Y
N
Read Tape (1):
∂₀W₂₀ (K)/∂c, ∂₀W₂₀ (K)/∂Z, ∂₁W₁₁ (K)/∂c, ∂₁W₁₁ (K)/∂Z ;
K = 1, 2, ···, NOSURF.
∂₁W₁₁ (K)/∂μ, ∂₀W₂₀ (K)/∂μ ; K = 1, 2, ····, (NOSURF+1).
Set
K3 = (I-1). NOSURF
KK1 = KRAY + (I-1). A
LL10 = 60 + KRAY + (I-1). A

PART "B"

Calculate:

$$H_o'(KK1) = \frac{E(NOSURF+1)}{N(LL10)} \left\{ \left[N(LL10)\cdot Y(LL10) - M(LL10)\cdot Z(LL10) \right] (\mu_{\bar{\lambda}}')_k + M(LL10)\cdot (h_{\bar{\lambda}})_k \right\}$$

$$\tilde{H}_o'(KK1) = E(NOSURF+1) \left[\bar{h}\,\mu_{\bar{\lambda}}' - h_{\bar{\lambda}}\,\bar{\mu}' \right]_k$$

$$H_\lambda = \frac{E(K3+1)}{E(1)} \cdot H_{\bar{\lambda}}$$

$$\delta H_s'(KK1) = \frac{H_\lambda \cdot H_o'(KK1)}{\tilde{H}_o'(KK1)}$$

$$\delta\Omega'(KK1) = -\frac{\delta H_s'(KK1)}{1+N(LL10)} \left\{ y'(KK1) - M(LL10)\left[M(LL10)\cdot y'(KK1) + N(LL10)\cdot z'(KK1) \right] \right\}$$

$$y'_{axial}(KK1) = y'(KK1) + \frac{M(LL10)\cdot \bar{\mu}_\lambda' \cdot \delta\Omega'(KK1)}{H_\lambda}$$

$$z'_{axial}(KK1) = z'(KK1) + \frac{N(LL10)\cdot \bar{\mu}_\lambda' \cdot \delta\Omega'(KK1)}{H_\lambda}$$

Set $LL12 = K3 + NOSURF + 1$

$$\frac{\partial W}{\partial (\delta W_{20})_{\bar{\lambda}}}(KK1) = -\frac{E(NOSURF+1)}{E(LL12)} \left\{ \left[y'_{axial}(KK1) \right]^2 + \left[z'_{axial}(KK1) \right]^2 \right\}$$

$$\frac{\partial W}{\partial a_o}(KK1) = -y'_{axial}(KK1)\cdot \delta H_s'(KK1)$$

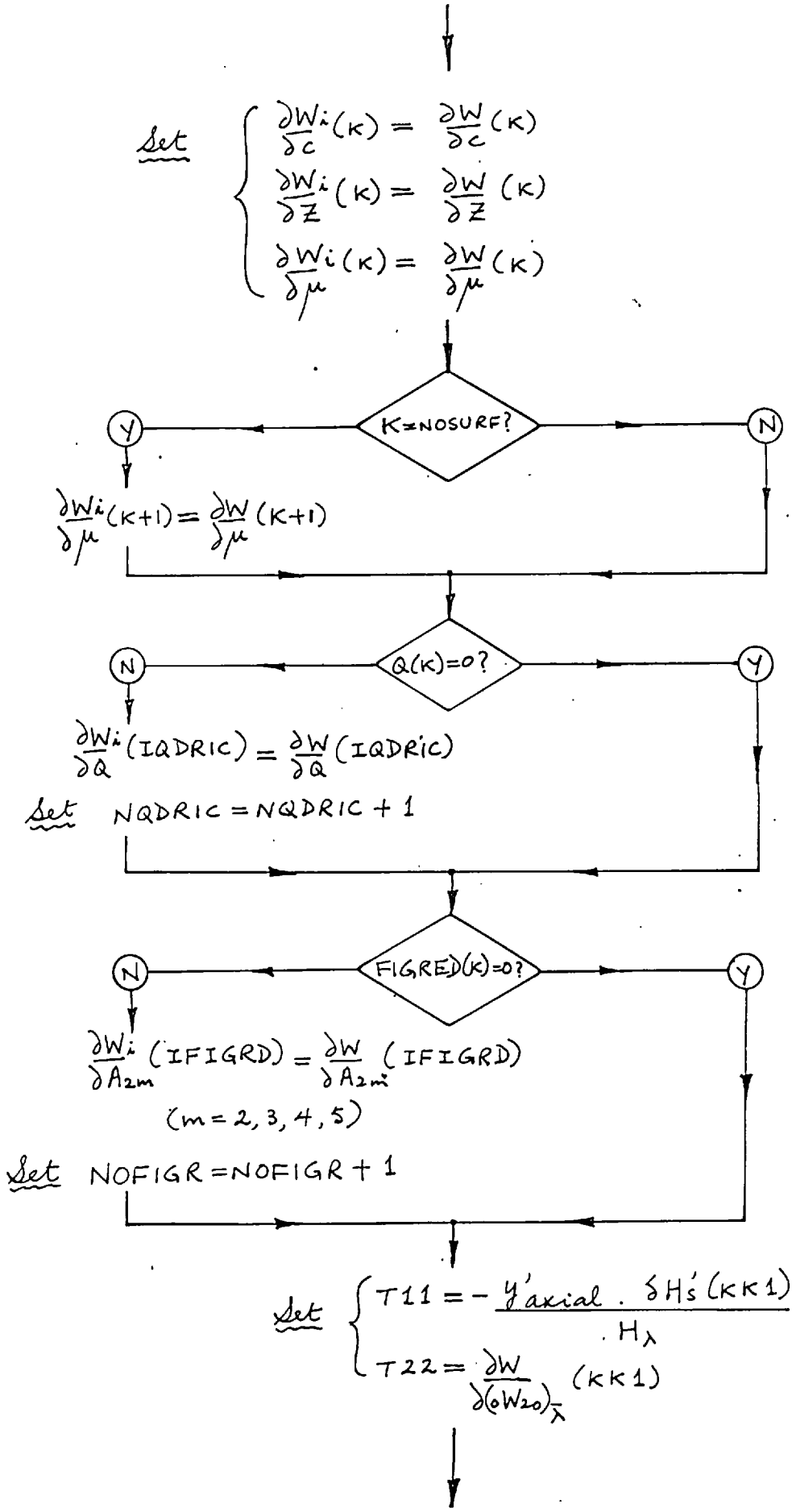
Set $\begin{cases} W = 0 \\ NOFIGR = 1 \\ NQDRIC = 1 \end{cases}$

PART "C"

Cycle K $(K = 1, 2, \cdots, NOSURF)$

Set $\begin{cases} LFIGRD = NOFIGR + LFIGRD \\ IQDRIC = NQDRIC + LQDRIC \end{cases}$

P.T.O.


Set
$\frac{\partial W_i}{\partial c}(K) = \frac{\partial W}{\partial c}(K)$
$\frac{\partial W_i}{\partial Z}(K) = \frac{\partial W}{\partial Z}(K)$
$\frac{\partial W_i}{\partial \mu}(K) = \frac{\partial W}{\partial \mu}(K)$
K=NOSURF?
Y
N
$\frac{\partial W_i}{\partial \mu}(K+1) = \frac{\partial W}{\partial \mu}(K+1)$
Q(K)=0?
N
Y
$\frac{\partial W_i}{\partial Q}(IQDRIC) = \frac{\partial W}{\partial Q}(IQDRIC)$
Set NQDRIC = NQDRIC + 1
FIGRED(K)=0?
N
Y
$\frac{\partial W_i}{\partial A_{2m}}(IFIGRD) = \frac{\partial W}{\partial A_{2m}}(IFIGRD)$
(m = 2, 3, 4, 5)
Set NOFIGR = NOFIGR + 1
Set
$T11 = -\frac{y'_{axial} \cdot \delta H's(KK1)}{H_\lambda}$
$T22 = \frac{\partial W}{\partial (\delta W_{20})_{\bar{\lambda}}}(KK1)$

$$\frac{\partial W_0}{\partial c}(K) = \frac{\partial W_i}{\partial c}(K) + T11 \cdot \frac{\partial_1 W_{11}}{\partial c}(K)$$

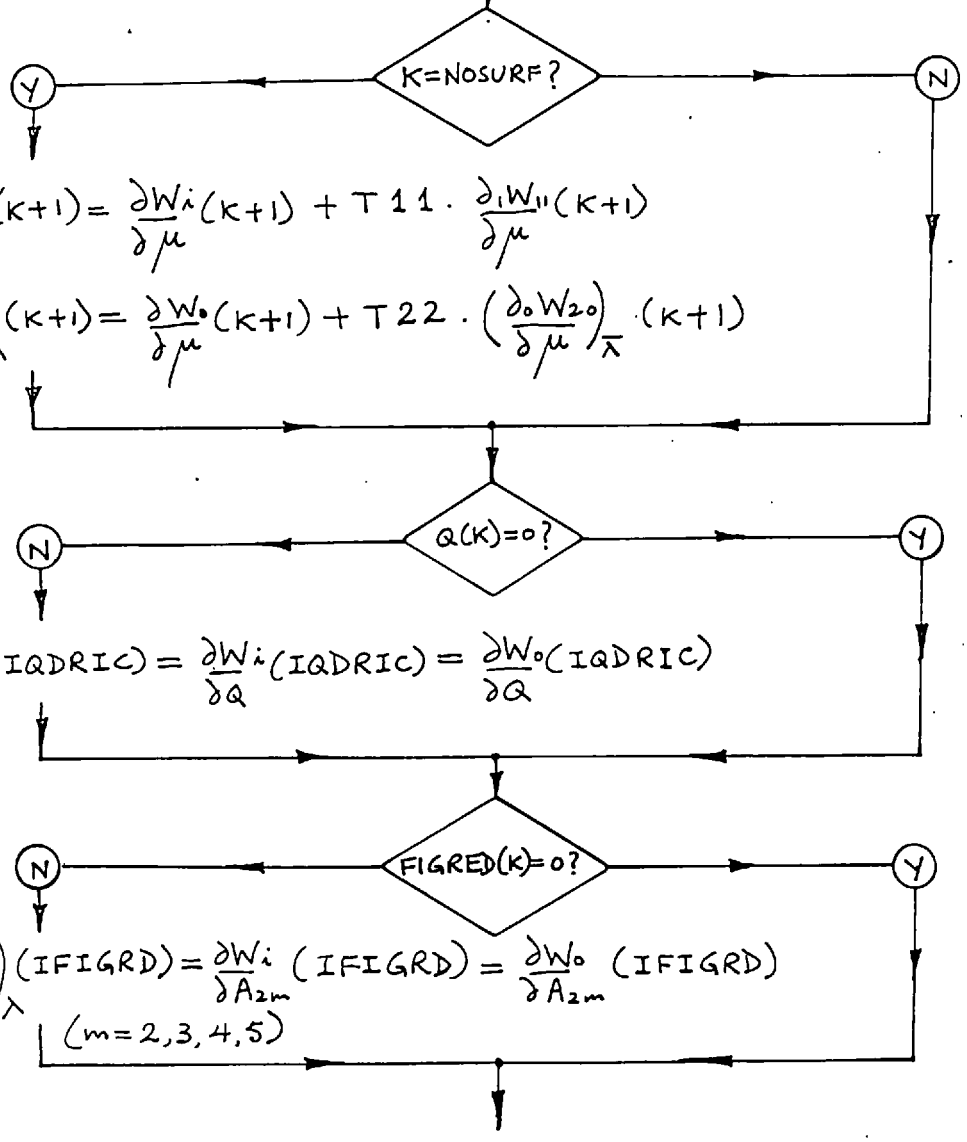
$$\left(\frac{\partial W}{\partial c}\right)_\lambda(K) = \frac{\partial W_0}{\partial c}(K) + T22 \cdot \left(\frac{\partial_0 W_{20}}{\partial c}\right)_{\bar{\lambda}}(K)$$

$$\frac{\partial W_0}{\partial z}(K) = \frac{\partial W_i}{\partial z}(K) + T11 \cdot \frac{\partial_1 W_{11}}{\partial z}(K)$$

$$\left(\frac{\partial W}{\partial z}\right)_\lambda(K) = \frac{\partial W_0}{\partial z}(K) + T22 \cdot \left(\frac{\partial_0 W_{20}}{\partial z}\right)_{\bar{\lambda}}(K)$$

$$\frac{\partial W_0}{\partial \mu}(K) = \frac{\partial W_i}{\partial \mu}(K) + T11 \cdot \frac{\partial_1 W_{11}}{\partial \mu}(K)$$

$$\left(\frac{\partial W}{\partial \mu}\right)(K) = \frac{\partial W_0}{\partial \mu}(K) + T22 \cdot \left(\frac{\partial_0 W_{20}}{\partial \mu}\right)_{\bar{\lambda}}(K)$$



Repeat K (next surface).

Set $\begin{cases} LFIGRD = LFIGRD + MFIGRD \\ LQDRIC = LQDRIC + MQDRIC \end{cases}$

$$W_{axial}(KK1) = \delta\Omega'(KK1) + W_{axial}(KK1)$$

Store on Tape (20):

$$\begin{cases} \left(\frac{\partial W_0}{\partial c}\right)(K), \left(\frac{\partial W_0}{\partial \bar{z}}\right)(K) \; ; \; K = 1, 2, \cdots, NOSURF. \\ \left(\frac{\partial W_0}{\partial \mu}\right)(K) \; ; \; K = 1, 2, \cdots, (NOSURF+1). \end{cases}$$

Store on Tape (21)

$$\begin{cases} \frac{\partial W_i}{\partial c}(K), \left(\frac{\partial W}{\partial c}\right)_\lambda(K), \frac{\partial W_i}{\partial \bar{z}}(K), \left(\frac{\partial W}{\partial \bar{z}}\right)_\lambda(K); \\ \qquad K = 1, 2, \cdots, NOSURF. \\ \frac{\partial W_i}{\partial \mu}(K), \left(\frac{\partial W}{\partial \mu}\right)_\lambda(K) \; ; \; K = 1, 2, \cdots, (NOSURF+1). \end{cases}$$

Repeat KRAY (next ray)

I=1?

Y

N

Set NCODE6 = B

Repeat I (next λ)

Rewind Tape (21)

RETURN.
END.

(VII-vii) Extra-axial ray subroutine.

As mentioned in Chapter 4, section (b), the starting data required for the extra axial ray trace are $(u_s,h_s)_1$ - see equation (4.18) and Fig. 4.6; these are calculated and stored for each field and wavelength in part A of flow-diagram (VII-vii). These are followed by part B, in which a_s and b_s, the aperture normalised major and minor semi-axes, are calculated. Additional details concerning (a_s,b_s) are given in Formulae (VII-vii). The starting values (X_1,Y_1) and (L_1,M_1,N_1) required for the ray trace are as those which have been defined by the equations (4.25), (4.24) and Fig. 4.8 in Chapter 4; these starting values are calculated and stored onto Tape (22) for each ray, wavelength and field in part C of the flow-diagram (VII-vii). The ray patterns adopted are those which have been described in Chapter 5, section (V-i).

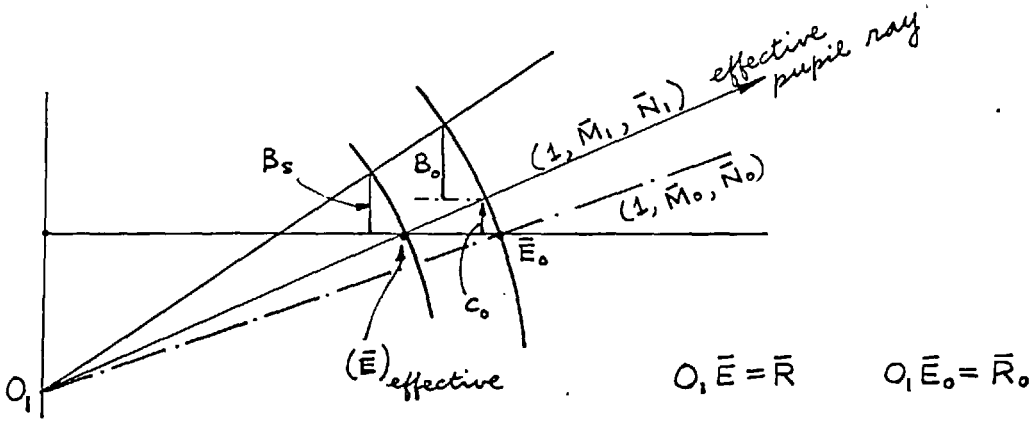
Before calling the finite ray trace subroutine, the coded values of NCODE4,6,7,8,9 are set in accordance with those defined in section (VII-iii).

In part D of the flow-diagram, the values of parameter derivatives calculated earlier are read from Tapes (1), (4) and (19). Tape (4) also provides the intermediate ray data which are required in the present subroutine. The image space quantities, such as $\underline{x}'$, $\underline{y}'$, $\underline{z}'$ etc., are computed in part E of the flow-diagram. An account of the image space variables, the wave-aberration calculations etc. can be found in Chapter 4.

In part F of the flow-diagram, the values of $\frac{\partial W_i}{\partial p}$, $\frac{\partial W_o}{\partial p}$ and $\left(\frac{\partial W}{\partial p}\right)_\lambda$ as defined in the equations (6.14), (6.15) and (6.19) are calculated and stored. In part G of the flow-diagram (VII-vii), the parameter derivatives are stored on to Tapes (20) and (21); the former tape is to be used for the optimisation stage and the latter is for the output subroutine and it stores also the image space data. Finally, in part H of the flow-diagram, the number of rays to be traced for $\lambda \neq \bar{\lambda}$ are set equal to 23+T (=3DB9), and this has been described in sections (VII-i) and (VII-iii).

FORMULAE (VII - vii)

Formulae for the aperture-normalised major and minor axes (a_s and b_s).



The effective pupil ray passes through the centre (0, c) of the (possibly vignetted) pupil, where c is defined as

$$c = \frac{c_0}{h}$$

From the diagram, $\frac{B_s}{\bar{R}} = \frac{B_0}{\bar{R}_0}$, and this may be expressed as

$$\frac{B_s}{h_s} \cdot \frac{h_s}{\bar{R}} = \frac{B_0}{h} \cdot \frac{h}{\bar{R}_0}$$

with the relations

$$(u_s)_1 = \frac{h_s}{\bar{R}} \quad ; \quad u_1 = \frac{h}{(\bar{R}_0/\bar{N}_0)} \quad ; \quad b = \frac{B_0}{h}$$

(b = semi-minor axis of the elliptical pupil)

and defining

$$b_s = \frac{B_s}{h_s}$$

the above relation gives

$$b_s = b \cdot \frac{u_1 \bar{N}_0}{(u_s)_1}$$

or, since $\frac{u_1}{(u_s)_1} = \frac{\bar{H}_0}{\bar{H}_s}$

$$\boxed{b_s = \bar{N}_0 \cdot \left(\frac{\bar{H}_0}{\bar{H}_s}\right) \cdot b}$$

The formula for a_s is found in a similar manner:

$$\boxed{a_s = \bar{N}_o \cdot \left(\frac{\bar{H}_o}{\bar{H}_s}\right) \cdot a}$$

where a is the semi-major axis of the elliptical pupil.

The formulae for the image space quantities are given in Chapter 4 (equations 4-29 to 4-39).

Focal shift derivatives:

$$\frac{\partial W}{\partial (\sigma W_{20})_{\bar{\lambda}}} = -\frac{n'_{\bar{\lambda}}}{n'_{\lambda}} \cdot \left[1 + \bar{M}'^{2} \cdot \frac{\partial \bar{H}'_o}{\bar{H}'_o}\right]_{\lambda} \left\{\frac{1}{\bar{N}'} \left(\frac{\bar{H}_s}{\bar{H}'_o}\right)^2 \left[(x'^2_s + y'^2_s + z'^2_s) + 2\,\partial \bar{y}'_s \cdot y'_s\right]\right\}_{\lambda}$$

where

$$\partial \bar{y}'_s = \frac{n'_{\lambda}}{\bar{N}'_{\bar{\lambda}} (\bar{H}_s)_{\lambda}} \left\{\bar{M}'_{\lambda}\left[\bar{N}'_{\bar{\lambda}}\, \bar{Y}_{\bar{\lambda}} - \bar{M}'_{\bar{\lambda}}\, \bar{Z}_{\bar{\lambda}}\right] - \bar{M}'_{\bar{\lambda}}\left[\bar{N}'_{\lambda}\, \bar{Y}_{\lambda} - \bar{M}'_{\lambda}\, \bar{Z}_{\lambda}\right]\right\}$$

$$\frac{\partial W}{\partial (W_{11})_{\bar{\lambda}}} = -\left(\frac{1}{\bar{H}_s}\right)_{\bar{\lambda}} \left\{\left[1 - \bar{N}'^{2} \cdot \frac{\partial \bar{H}'_o}{\bar{H}'_o}\right] \cdot \bar{H}_s \cdot y'_s\right\}_{\lambda}$$

$$\left(\frac{\partial W_{11}}{\partial P}\right)_{\bar{\lambda}} = \left[\frac{\bar{H}_s}{\bar{H}'_T} \cdot \frac{\partial W_{T1}}{\partial P}\right]_{\bar{\lambda}}$$

Derivatives of aberration with respect to aperture:

Since $x_s = a_s \cdot x$; $y_s = b_s \cdot y$; $\delta G'_s = -\frac{\partial W}{\partial x'_s}$; $\delta H'_s = -\frac{\partial W}{\partial y'_s}$, the derivatives of aberration w.r.t. a_s and b_s are

$$\frac{\partial W}{\partial a_s} = -x \cdot \delta G'_s \quad ; \quad \frac{\partial W}{\partial b_s} = -y \cdot \delta H'_s \,.$$

The parameter derivative $\left(\frac{\partial W}{\partial P}\right)_{\lambda}$ is calculated using formulae (6-19), (6-16) and (6-15).

FLOW - DIAGRAM (VII - vii)

Extra - axial Ray (ASRAYS) Subroutine

PART "A"

Rewind Tape (4)

→ Cycle KFIELD (KFIELD = 1, 2, ···, NFIELD)

→ Cycle I (I = 1, 2, ···, NOWAVE)

Set: $KK1 = I + (KFIELD - 1) \cdot NOWAVE$

$$(u_s)_1\,(KK1) = \frac{\bar{H}_s(KK1)}{\bar{H}_o(KK1)} \cdot u_1$$

$$(h_s)_1\,(KK1) = \frac{\bar{H}_s(KK1)}{\bar{H}_o(KK1)} \cdot \frac{(h_1 - u_1 \bar{z}_1)}{\bar{N}_1}$$

$\bar{z}_1$ and $\bar{N}_1$ for the pupil ray have been stored in the pupil ray subroutine

↓

← Repeat I (next λ)

← Repeat KFIELD (next τ)

↓

PART "B"

Set: NCODE9 = 2B + T

→ Cycle KFIELD (KFIELD = 1, 2, ···, NFIELD)

Set: NCODE6 = 2A + S

→ Cycle I (I = 1, 2, ···, NOWAVE)

Set: $KK1 = I + (KFIELD - 1) \cdot NOWAVE$

$$\bar{N}_o = \frac{1}{\sqrt{1 + (\bar{u}_1 \tau)^2}}$$

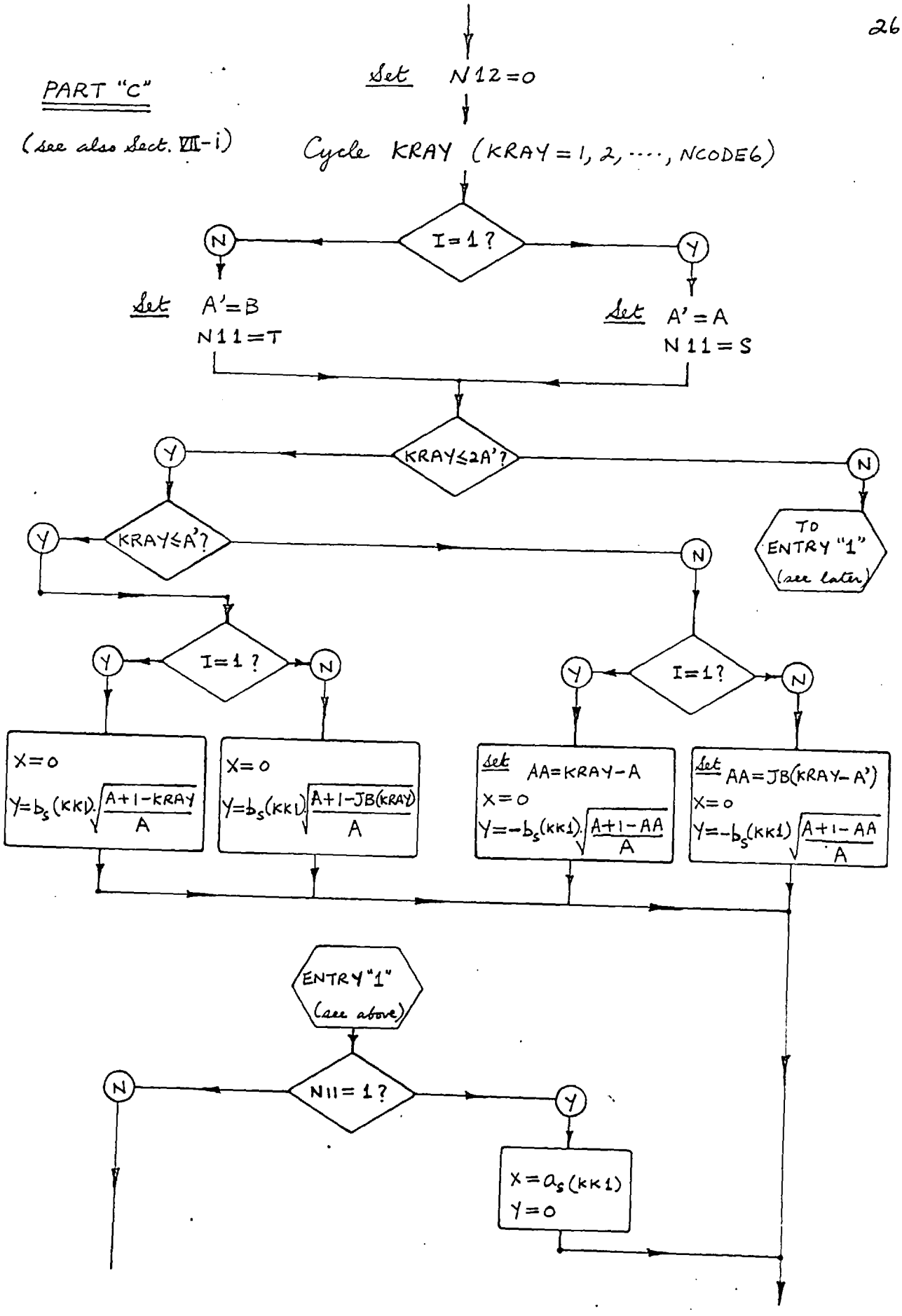
$$a_s(KK1) = \bar{N}_o \cdot \left(\frac{\bar{H}_o(KK1)}{\bar{H}_s(KK1)}\right) \cdot a_{pupil}\,(KK1)$$

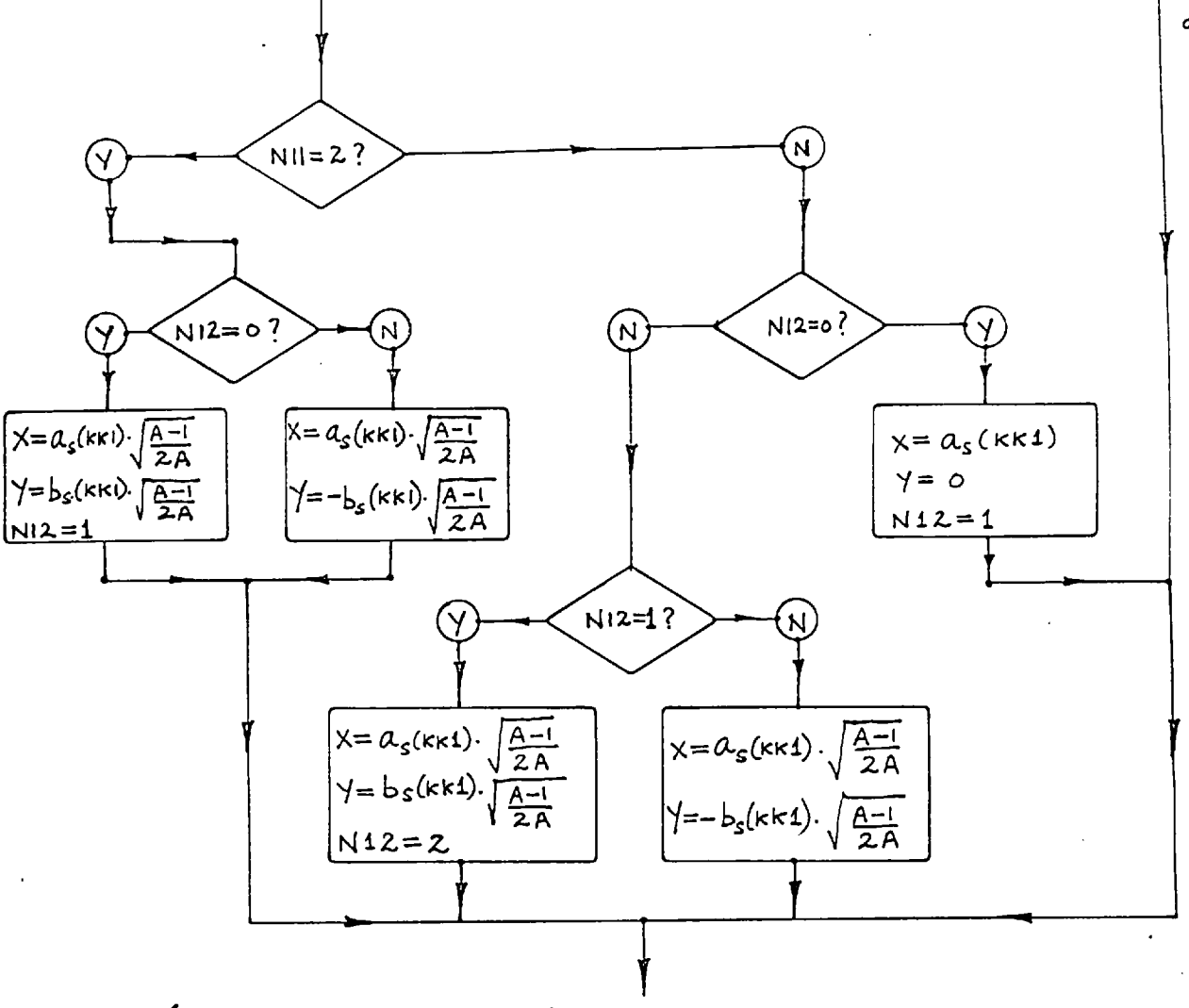
$$b_s(KK1) = \bar{N}_o \cdot \left(\frac{\bar{H}_o(KK1)}{\bar{H}_s(KK1)}\right) \cdot b_{pupil}\,(KK1)$$

↓

PART "C"

(see also Sect. VII-i)





Set

$$N15 = KRAY + NOWAVE.(KFIELD-1)(2A+S) + (I-1)(2A+S)$$

$$X_{EO}(N15) = X / a_s(KK1)$$

$$Y_{EO}(N15) = Y / b_s(KK1)$$

$$L_1 = -X \cdot (\mu_s)_1(KK1)$$

$$M_1 = -Y \cdot (\mu_s)_1(KK1) + \bar{M}_1$$

$$N_1 = +\sqrt{1-(L_1^2 + M_1^2)}$$

$$X = (h_s)_1(KK1) \cdot X$$

$$Y = \bar{Y}_1 + Y \cdot (h_s)_1(KK1)$$

$$Z = \bar{Z}_1 - \left(\frac{L_1 X + (M_1 + \bar{M}_1)(Y - \bar{Y}_1)}{N_1 + \bar{N}_1}\right)$$

$$e_1 = \frac{L_1 X + M_1 Y + (1+N_1) Z}{1+N_1}$$

$$X(1) = Y - L_1 \cdot e_1$$
$$Y(1) = Y - M_1 \cdot e_1$$

Store on Tape (22):

$X(1),\ Y(1),\ L_1,\ M_1,\ N_1$

Repeat KRAY (next ray)

I=1?

Y

N

Set NCODE6 = NCODE9

Repeat I (next λ)

Repeat KFIELD (next T)

Rewind Tapes (22) and (21)

Set:

NCODE4 = NFIELD
NCODE6 = 2A+S
NCODE7 = 0
NCODE8 = 1

CALL FINITE RAY TRACE SUBROUTINE

from ray trace

Set: LFIGRD = 0
LQDRIC = 0
NCODE9 = 2B+T

PART "D"

Cycle KFIELD (KFIELD = 1, 2, ···, NFIELD)

NCODE6 = 2A + S

Cycle I (I = 1, 2, ···, NOWAVE)

Read Tape (1):

$$\left\{ \frac{\partial_0 W_{20}}{\partial c}(K),\ \frac{\partial_0 W_{20}}{\partial z}(K),\ \frac{\partial_1 W_{11}}{\partial c}(K),\ \frac{\partial_1 W_{11}}{\partial z}(K);\quad K=1,2,\cdots,NOSURF. \right.$$

$$\left. \frac{\partial W_{11}}{\partial \mu}(K1),\ \frac{\partial_0 W_{20}}{\partial \mu}(K1);\quad K1=1,2,\cdots,(NOSURF+1). \right\}$$

Read Tape (19):

$$\left\{ \frac{\partial W_{TI}}{\partial c}(K),\ \frac{\partial W_{TI}}{\partial z}(K),\ \frac{\partial \bar{W}_{TI}}{\partial c}(K),\ \frac{\partial \bar{W}_{TI}}{\partial z}(K);\quad K=1,2,\cdots,NOSURF. \right.$$

$$\left. \frac{\partial W_{TI}}{\partial \mu}(K1),\ \frac{\partial \bar{W}_{TI}}{\partial \mu}(K1);\quad K1=1,2,\cdots,(NOSURF+1). \right\}$$

Cycle KRAY (KRAY = 1, 2, ···, NCODE6)

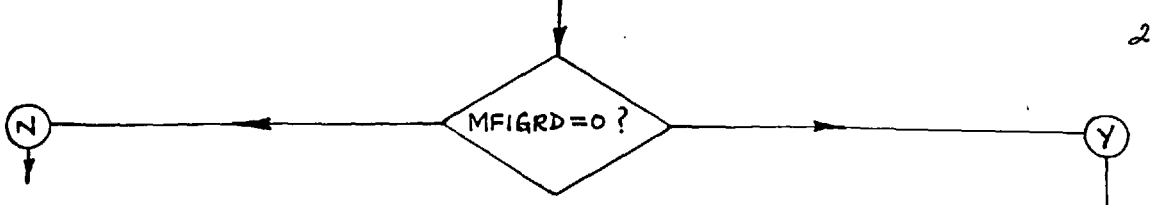
Read Tape (4):

$$\left\{ \begin{array}{l} X_k, Y_k, Z_k, L'_k, M'_k, N'_k, X_1, Y_1, Z_1, \\ L_1, M_1, N_1, e_1, e_k'. \\ W_{extra\text{-}axial}(KRAY);\ \frac{\partial W}{\partial c}(K),\ \frac{\partial W}{\partial z}(K);\quad K=1,2,\cdots,NOSURF. \\ \frac{\partial W}{\partial \mu}(K1);\quad K1=1,2,\cdots,(NOSURF+1). \end{array} \right.$$

MQDRIC = 0 ?

(N) Read Tape (4): $\frac{\partial W}{\partial Q}(K)$; K = 1, 2, ···, MQDRIC.

(Y)



Read Tape (4) : $\frac{\partial W}{\partial A_{2m}}(k)$; $k = 1, 2, \cdots, MFIGRD$. $m = 2, 3, 4, 5$

Set: $KK1 = KRAY$

PART "E"

$$H_o'(KK1) = \frac{E(NOSURF+1)}{N_k'}\left\{(N_k' Y_k - M_k' Z_k)\mu_\lambda' + M_k'(h_1)_{\bar\lambda}\right\}$$

$$G_o'(KK1) = \frac{E(NOSURF+1)}{N_k'}\left\{(N_k' X_k - L_k' Z_k)\mu_\lambda' + L_k'(h_1)_{\bar\lambda}\right\}$$

Set: $I_{12} = (KFIELD-1).NOWAVE + 1$

$I_{11} = I + (KFIELD-1).NOWAVE$

$$\delta H_s'(KK1) = \pi'(I11)\cdot[H_o'(KK1) - \bar H_o'(I12)]$$

$$\delta G_s'(KK1) = \pi'(I11).\ G_o'(KK1)$$

$$\tilde e_k'(KK1) = \frac{L_k' X_k + (M_k' + \bar M_k')(Y_k - \bar Y_k) + (N_k' + \bar N_k')(Z_k - \bar Z_k)}{1 + (L_k' \bar L_k' + M_k' \bar M_k' + N_k' \bar N_k')}$$

$$X = X_k - \tilde e_k'(KK1)\ .\ L_k'$$

$$Y = Y_k - \tilde e_k'(KK1)\ .\ M_k'$$

Set: $LL23 = (NOSURF+1).I$

$$\underline{x}'(KK1) = \frac{E(LL23)}{\bar H_s(I11)}.\left\{\bar M_k'.X - L_k' \bar Y_k\right\}$$

$$\underline{y}'(KK1) = \frac{E(LL23)}{\bar H_s(I11)}.\left\{\bar M_k'.Y - M_k' \bar Y_k\right\}$$

$$z'(kk1) = -\left\{\frac{L'_k \cdot \underline{x}'(kk1) + (M'_k + \bar{M}'_k)\, y'(kk1)}{N'_k + \bar{N}'_k}\right\}$$

$$TT = L'_k \cdot \underline{x}(kk1) + M'_k \cdot y(kk1) + N'_k \cdot z(kk1)$$

$$\delta\Omega'(kk1) = -\frac{1}{1 + M'_k \bar{M}'_k + N'_k \bar{N}'_k}\left\{\left[\underline{x}'(kk1) - L'_k \cdot TT\right]\cdot \delta G'_s(kk1) + \left[y'(kk1) - M'_k \cdot TT\right]\cdot \delta H'_s(kk1) + \delta \bar{H}'_s(111)\left[y'(kk1) - \bar{M}'_k\left(\bar{M}'_k \cdot y(kk1) + \bar{N}'_k \cdot z(kk1)\right)\right]\right\}$$

$$TT1 = \frac{\bar{M}'_k \cdot \delta\Omega'(kk1)}{\bar{H}_s(111)}$$

$$x'_s(kk1) = \underline{x}'(kk1) - L'_k \cdot TT1$$

$$y'_s(kk1) = y'(kk1) - M'_k \cdot TT1$$

$$z'_s(kk1) = z'(kk1) - N'_k \cdot TT1$$

$$\frac{\partial W}{\partial (\delta W_{20})_{\bar{\lambda}}}(kk1) = \frac{-E(NOSURF+1)}{\bar{N}'_k \cdot E(LL23)} \cdot \left[\left(\frac{\tilde{N}'}{\bar{N}'}\right)(I11)\right]\cdot\left\{\frac{\bar{H}_s(I11)}{\bar{H}'_0(I11)}\right\}^2 \times \left\{[x'_s(kk1)]^2 + [y'_s(kk1)]^2 + [z'_s(kk1)]^2 + 2\,\partial\bar{y}'_s(I11)\cdot y'_s(kk1)\right\}$$

$$\frac{\partial W}{\partial (W_{11})_{\bar{\lambda}}}(kk1) = -\frac{\bar{H}_s(I11)}{\bar{H}_s(I12)} \cdot \left[\left(\frac{\tilde{M}'}{\bar{M}'}\right)(I11)\right]\cdot y'_s(kk1)$$

Set $N15 = KRAY + NOWAVE.(KFIELD-1)(2A+S) + (I-1)(2A+S)$

$$\frac{\partial W}{\partial a_s}(kk1) = -x_{E0}(N15)\cdot \delta G'_s(kk1)$$

$$\frac{\partial W}{\partial b_s}(kk1) = -y_{E0}(N15)\cdot \delta H'_s(kk1)$$

$$\frac{\partial W}{\partial \bar{y}'_s}(kk1) = -f'(I11)\cdot \delta H'_s(kk1)$$

Set: NOFIGR=0
NQDRIC=0

PART "F"

Cycle K (K=1, 2, ···, NOSURF)

Set { IFIGRD=NOFIGR + LFIGRD
IQDRIC=NQDRIC + LQDRIC

For $p = c, z, \mu$:

Store: $\frac{\partial W_i}{\partial p}(K) = \frac{\partial W}{\partial p}(K)$

Calculate:

$$\frac{\partial W_0}{\partial p}(K) = \frac{\partial W}{\partial p}(K) - \delta G_s'(KK1) \cdot \frac{\bar{N}_k'^2 \cdot x_s'(KK1)}{\bar{H}_T'(I11)} \cdot \frac{\partial W_{TI}}{\partial p}(K)$$

$$- \delta H_s'(KK1) \cdot \left\{ \frac{\bar{N}_k'}{\bar{H}_s(I11)} \cdot \frac{\partial \bar{W}_{TI}}{\partial p}(K) + \frac{\bar{N}_k'^2}{\bar{H}_T'(I11)} \cdot \frac{\partial W_{TI}}{\partial p}(K) \cdot y_s'(KK1) \right\}$$

$$- \delta \bar{H}_s'(I11) \cdot \frac{\bar{N}_k'}{\bar{H}_s(I11)} \cdot \frac{\partial \bar{W}_{TI}}{\partial p}(K)$$

$$\left(\frac{\partial W}{\partial p}\right)_\lambda (K) = \frac{\partial W_0}{\partial p}(K) + \frac{\partial W}{\partial ({}_0W_{20})_{\bar{\lambda}}}(KK1) \cdot \left(\frac{\partial {}_0W_{20}}{\partial p}\right)_{\bar{\lambda}}(K) + \frac{\partial W}{\partial (W_{11})_{\bar{\lambda}}}(KK1) \left(\frac{\partial W}{\partial p}\right)_{\bar{\lambda}}(K)$$

K=NOSURF? — Y / N

Store: $\frac{\partial W_i}{\partial \mu}(K+1) = \frac{\partial W}{\partial \mu}(K+1)$

Calculate: { $\frac{\partial W_0}{\partial p}(K+1)$
$\left(\frac{\partial W}{\partial p}\right)_\lambda (K+1)$ } for $p = \mu$.

Q(K)=0? (N / Y)

$\frac{\partial W_i}{\partial Q}(IQDRIC) = \frac{\partial W}{\partial Q}(IQDRIC)$

Calculate: $\frac{\partial W_0}{\partial p}(IQDRIC)$; $p = Q$

Set: NQDRIC = NQDRIC + 1

FIGRED(K)=0? (N / Y)

$\frac{\partial W_i}{A_{2m}}(IFIGRD) = \frac{\partial W}{\partial A_{2m}}(IFIGRD)$

Calculate: $\frac{\partial W_0}{\partial p}(IFIGRD)$; $p = A_{2m}$ } $m = 2, 3, 4, 5$

Set: NOFIGR = NOFIGR + 1

Repeat K (next surface)

$$\tilde{e}_1 = \frac{L_1 X_1 + (M_1 + \bar{M}_1)(Y_1 - \bar{Y}_1) + (N_1 + \bar{N}_1)(Z_1 - \bar{Z}_1)}{1 + M_1 \bar{M}_1 + N_1 \bar{N}_1}$$

Set: JJ3 = (I − 1).(NOSURF + 1) + 1

Calculate:

$$W_{extra\text{-}axial}(KK1) = W_{extra\text{-}axial}(KK1) - W_{pupil}(I11) + E(JJ3).e_1 - E(JJ3+NOSURF).e'_k - E(JJ3).\bar{e}_1 + E(JJ3+NOSURF).\bar{e}'_k + E(JJ3+NOSURF).\tilde{e}'_k - E(JJ3).\tilde{e}_1 + \delta\Omega'(KK1)$$

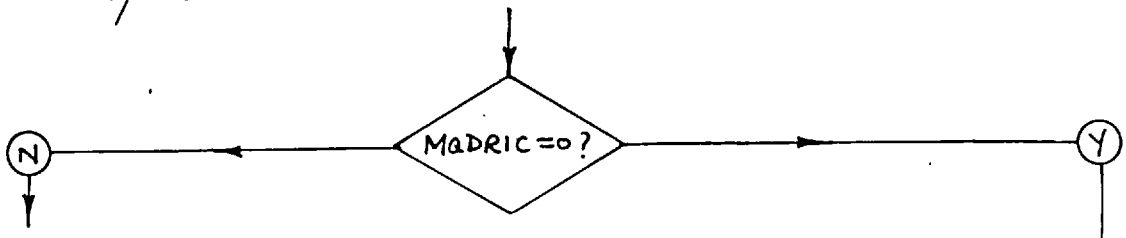
PART "G"

write on Tape (20):

$$\begin{cases} W_{\text{extra-axial}}(KK1); \quad \left(\frac{\partial W_0}{\partial c}\right)(K), \left(\frac{\partial W_0}{\partial z}\right)(K); \\ \qquad\qquad K = 1, 2, \cdots, NOSURF. \\ \left(\frac{\partial W_0}{\partial \mu}\right)(K1); \quad K1 = 1, 2, \cdots, (NOSURF+1) \end{cases}$$

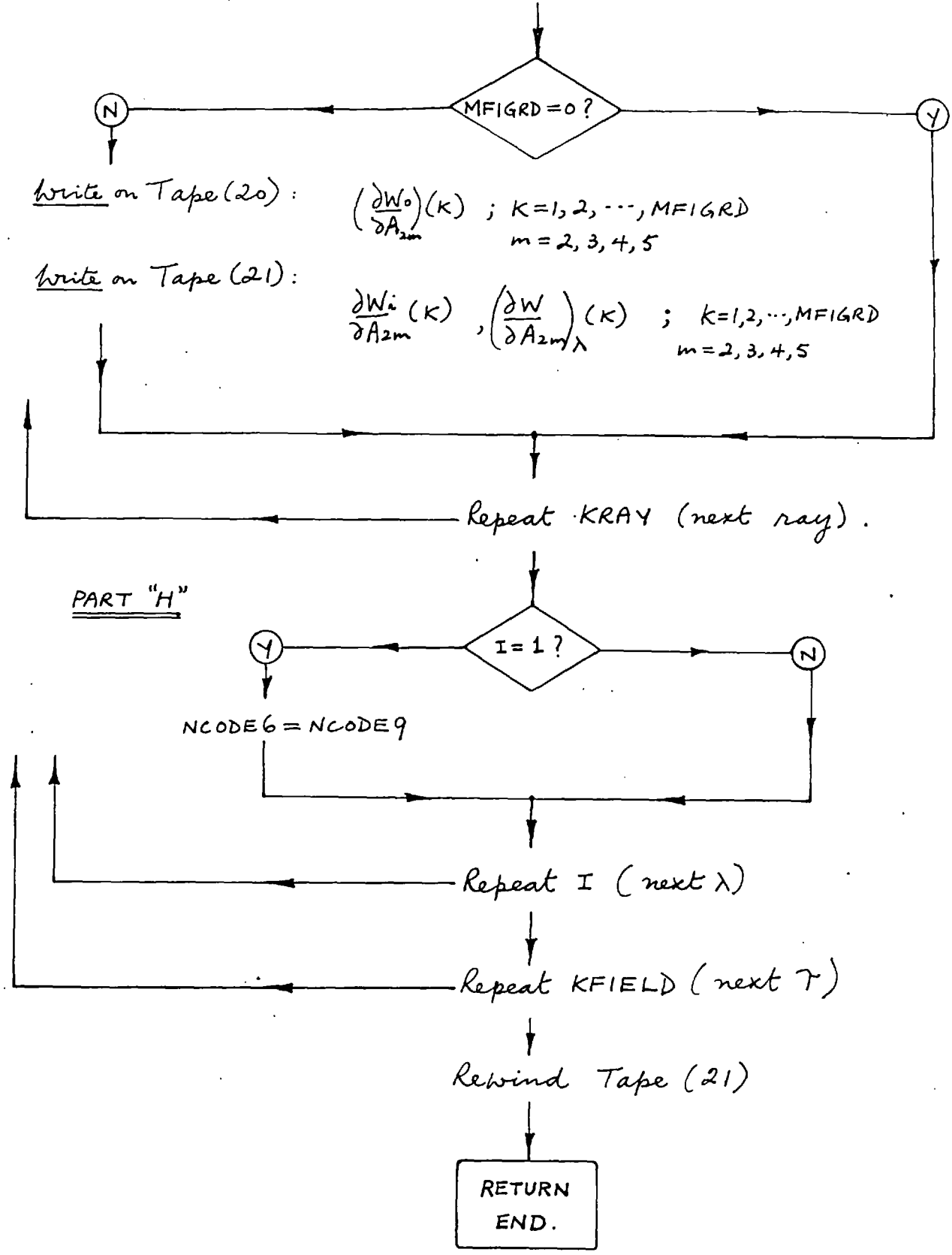
write on Tape (21): — for PRINT-OUT.

$$\begin{cases} W_{\text{extra-axial}}(KK1), \frac{\partial W}{\partial ({}_0W_{20})_{\bar{\lambda}}}(KK1), \frac{\partial W}{\partial (W_{11})_{\bar{\lambda}}}(KK1), \frac{\partial W}{\partial a_s}(KK1), \frac{\partial W}{\partial b_s}(KK1) \\ \frac{\partial W}{\partial \bar{y}_s}(KK1), \delta H_s'(KK1), \delta G_s'(KK1), \underline{x}'(KK1), \underline{y}'(KK1), \underline{z}'(KK1), \\ x_s'(KK1), y_s'(KK1), z_s'(KK1), H_0'(KK1), G_0'(KK1), \\ \frac{\partial W_i}{\partial p}(K), \left(\frac{\partial W}{\partial p}\right)_{\lambda}(K) \quad \begin{cases} K = 1, 2, \cdots, NOSURF. \\ p = c, z \end{cases} \\ \frac{\partial W_i}{\partial \mu}(K1), \left(\frac{\partial W}{\partial \mu}\right)_{\lambda}(K1); \quad K1 = 1, 2, \cdots, (NOSURF+1) \end{cases}$$



write on Tape (20): $\left(\frac{\partial W_0}{\partial Q}\right)(K)$; $K = 1, 2, \cdots, MQDRIC.$

write on Tape (21): $\frac{\partial W_i}{\partial Q}(K), \left(\frac{\partial W}{\partial Q}\right)_{\lambda}(K)$; $K = 1, 2, \cdots, MQDRIC.$


MFIGRD = 0 ?
N
Y
Write on Tape (20): $\left(\frac{\partial W_0}{\partial A_{2m}}\right)(K)$; K = 1, 2, ···, MFIGRD
m = 2, 3, 4, 5
Write on Tape (21): $\frac{\partial W_i}{\partial A_{2m}}(K)$, $\left(\frac{\partial W}{\partial A_{2m}}\right)_\lambda(K)$; K = 1, 2, ···, MFIGRD
m = 2, 3, 4, 5
Repeat KRAY (next ray).
PART "H"
I = 1 ?
Y
N
NCODE6 = NCODE9
Repeat I (next λ)
Repeat KFIELD (next τ)
Rewind Tape (21)
RETURN
END.

(VII-viii) Output and discussion

In the output subroutine, the system data, the ray and pupil data, the image space quantities and the data stored on tapes (2) and (21) are all printed out in sequence. An example of the computer output is enclosed at the end of this section. The system chosen is a 16 m.m. microscope objective with six surfaces, four of which are figured in order that the programme may be properly tested. The number of rays chosen are A=B=2, S=T=1. Three wavelengths have been used; and for each wavelength, the extra-axial ray is traced with τ=1, the values of (a,b,c) being those provided by the pupil exploration output for the same system - see the computer output example given at the end of Chapter 4.

The computer programme described in this chapter requires a total of over 2,600 I.B.M punched cards, excluding the input data cards. This programme has been under development for about two years, and the computer output has been carefully checked on desk calculator and also with results obtained from the Elliot 803B computer.

The machine time required for the accompanying example was approximately two minutes, using only the object programme deck. The machine time will naturally be proportionally longer when more surfaces, rays, wavelengths and fields than that mentioned above are involved. The use of eight magnetic tapes in the programme impedes slightly the speed of computation, but the use of the tapes has been carefully designed to operate

with the minimum number of rewinds to economise the machine time.

It has been briefly described in Chapter 1 that most lens designers calculate the ratio of increment $\delta W/\delta p$ by making small finite changes, δp's, to the system and retrace the rays; $\delta W/\delta p$ is then used as an approximation for the derivative $\partial W/\partial p$. The method described in this chapter has the following important advantages :

a) it calculates the true parameter derivatives, $\partial W/\partial p$'s , and these give improved predictions[2] of the aberrations resulting from any constructional change as compared with $\delta W/\delta p$'s.

b) the computation of the parameter derivatives is based on a single set of ray-trace, and forms an integral part of the ray tracing schedule; it eliminates the time wastage which is associated with the re-tracing of rays after each δp has been made. The net result is a gain in speed and this gain increases with the complexity of the system (e.g. zoom lenses).

REFERENCES

(1) H.H. Hopkins, Wave Theory of Aberrations, Oxford (1950)

(2) I.B.M. Research Report (H.H. Hopkins), Optical Calculation V (1963).

General References see Chapter 6

The author wishes to express his thanks to Mr. G.T. Church for his assistance in checking this computer output.

SYSTEM DATA

	MEAN	WAVELENGTH =	5461.000000000
	ADDITIONAL	WAVELENGTH =	6563.000000000
	ADDITIONAL	WAVELENGTH =	4861.000000000

SURFACE NO. = 1 C = 0. MAX. SEMI-APERTURE = 1.78600000 Q = 0.00001

SPECIAL FIGURING MTHAX = 2.500000
FIGURING COEF. = 0.000000999 A2
FIGURING COEF. = 0.000000999 A4
FIGURING COEF. = 0.000000999 A6
FIGURING COEF. = 0.000000999 A8

MEAN	WAVELENGTH E =	1.000000000	DELTA E = 0.
ADDITIONAL	WAVELENGTH E =	1.000000000	DELTA E = 0.
ADDITIONAL	WAVELENGTH E =	1.000000000	DELTA E = 0.

SURFACE NO. = 2 C = 0.15953249986 MAX. SEMI-APERTURE = 2.05560001 Q = 0.

D = 1.200000003

MEAN	WAVELENGTH E =	1.629899994	DELTA E = 0.01751
ADDITIONAL	WAVELENGTH E =	1.620619997	DELTA E = -0.00928
ADDITIONAL	WAVELENGTH E =	1.638219994	DELTA E = 0.00823

SURFACE NO. = 3 C = -.15774935% MAX. SEMI-APERTURE = 2.33700001 Q = 0.00001

D = 2.400000006

SPECIAL FIGURING MTHAX = 2.800000
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999

MEAN	WAVELENGTH E =	1.521019993	DELTA E = 0.00359
ADDITIONAL	WAVELENGTH E =	1.516369998	DELTA E = -0.00447
ADDITIONAL	WAVELENGTH E =	1.524159996	DELTA E = 0.00392

SURFACE NO. = 4 C = 0. MAX. SEMI-APERTURE = 4.15100002 Q = 0.

D = 14.055999994

MEAN	WAVELENGTH E =	1.000000000	DELTA E = 0.
ADDITIONAL	WAVELENGTH E =	1.000000000	DELTA E = 0.
ADDITIONAL	WAVELENGTH E =	1.000000000	DELTA E = 0.

SURFACE NO. = 5 C = 0.08425599% MAX. SEMI-APERTURE = 4.32900006 Q = 0.00001

D = 1.599999994

SPECIAL FIGURING MTHAX = 4.800000
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999

MEAN	WAVELENGTH E =	1.629899994	DELTA E = 0.01751
ADDITIONAL	WAVELENGTH E =	1.621419997	DELTA E = -0.00928
ADDITIONAL	WAVELENGTH E =	1.638219994	DELTA E = 0.00823

SURFACE NO. = 6 C = -0.07792300C MAX. SEMI-APERTURE = 4.50999999 Q = 0.00001

SURFACE NO. = 6 THIS IS STOP

D = 3.900000012

SPECIAL FIGURING MTHAX = 5.300000
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999
FIGURING COEF. = 0.000000999

MEAN	WAVELENGTH E =	1.523039993	DELTA E = 0.00859
ADDITIONAL	WAVELENGTH E =	1.516769998	DELTA E = -0.00447
ADDITIONAL	WAVELENGTH E =	1.524459996	DELTA E = 0.00392

IMAGE SPACE

MEAN	WAVELENGTH E =	1.000000000	DELTA E = 0.
ADDITIONAL	WAVELENGTH E =	1.000000000	DELTA E = 0.
ADDITIONAL	WAVELENGTH E =	1.000000000	DELTA E = 0.

RAY AND PUPIL DATA

ELLIPTICAL PUPIL

FIELD X1.00000000

WAVELENGTH = 5461.00 A A = 0.4928660 B = 0.4832300 C = 0.25k1k04

WAVELENGTH = 6563.00 A A = 0.4928660 B = 0.4832300 C = 0.25k1k04

WAVELENGTH = 4861.00 A A = 0.4928660 B = 0.4832300 C = 0.25k1k04

HCIX = 1.71k16000 UCIX = -0.2000000 HSAMIX = -1.0233050 UBARCIX = 0.

NO. OF A RAYS = 2 NO. OF C RAYS = 2 NO. OF S RAYS = 1 NO. OF T RAYS = 1

USE LINEAR SIZE FORMULAE

RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	5461.00 A	SURF. NO. =	2
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	5461.00 A	SURF. NO. =	3
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	6563.00 A	SURF. NO. =	2
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	6563.00 A	SURF. NO. =	3
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	6563.00 A	SURF. NO. =	5
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	6563.00 A	SURF. NO. =	6
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	4861.00 A	SURF. NO. =	4
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	4861.00 A	SURF. NO. =	5
RAY HT. EXCEEDS MAX. VALUE, F = 1 CASE	FIELD = 1.0000000	WAVELENGTH =	4861.00 A	SURF. NO. =	6

2ND. FOCAL LENGTH = 10.1025752 DIST. 2ND. P PLANE FROM LAST SURF.= 18.2170550

1ST FOCAL LENGTH = -14.2125754 DIST. 1ST. P PLANE FROM 1ST. SURF.= 11.7381288

LINEAR MAGNIFICATION = -.7799 TOTAL THICK = 385.0435505

PRIMARY ABERRATIONS ARE IN PRACTICAL UNITS

SIm 2.73458716	SIm = -1.4557960	SIm =	1.3451555	SIm =	2.62067510	SIm =	-0.17202177

SPHERICAL DISTORTION OF PUPIL = -.8709

WAVELENGTH NO. = 1	CC =	-1.6610796	CT = 6.09399408
WAVELENGTH NO. = 2	CC =	-1.5807551	CT = -0.03638826

MEAN WAVELENGTH N = 1.48178 N' = 0.286525

5461.000000000 A

DATA AT 1ST SURF.

W DASH = 5.12044452	PARAXIAL RAY HT. =	4.47705160	L DAR DASH =	-0.06317499	HT. LIM =	0.12722940
YIELD = 1.00000000	HBARDASH =	-.28385956	HSAMDASH =	0.28795736	DISTORTION =	0.00450837

DISTORTION DERIVATIVES IN PERCENTAGE

SURF.	DD/DC ($\frac{\partial D}{\partial C}$)	DD/DC ($\frac{\partial D}{\partial C}$)	DD/DB ($\frac{\partial D}{\partial B}$)	DD/DC ($\frac{\partial D}{\partial C}$)	DD/DAN ($\frac{\partial D}{\partial A_N}$)	DD/DAN ($\frac{\partial D}{\partial A_N}$)	DD/DAN ($\frac{\partial D}{\partial A_N}$)	DD/DAN ($\frac{\partial D}{\partial A_N}$)
1	-2.73852	-161.67103	0.	0.	11.11037	1.08424	0.09404	0.00775
2	1.47357	35.05912	34.72145	0.	0.	0.	0.	0.
3	-4.69251	264.96262	-331.85880	0.07336	-2.45102	-0.58951	-0.08291	0.
4	-0.43647	-929.40959	0.	0.	0.	0.	0.	0.
5	0.69120	174.99651	-302.19578	0.05893	1.40378	0.10039	0.00635	-0.00016
6	-1.30688	938.03688	190.40705	-0.41358	4.49780	0.79830	0.06712	0.
IMAGE SPACE		0.						

PARAMETER DERIVATIVES ARE IN TERM OF WAVELENGTH (= 0.00054610)

SURF.	DW20/DC ($\frac{\partial W_{20}}{\partial C}$)	DW20/DC	DW42/DC ($\frac{\partial W_{42}}{\partial C}$)	DW31/DZ ($\frac{\partial W}{\partial Z}$)	DW31/DC	DW31/DE
	DBARSQ20/DC ($\frac{\partial \bar{B}^2_{20}}{\partial C}$)	DBARSQ42/DC	DBARSQ20/DE	DW/DBD($\frac{\partial W}{\partial \bar{B}\bar{D}}$)	DW/DBC	DW/DBE ($\frac{\partial W}{\partial \bar{B}\bar{E}}$)
1	27.743155	1694.619263	0.	0.	-2023.279480	0.
	-0.	623.929688	0.	-426.943920	-19015.279688	0.
2	-12.046809	-367.542671	-370.353509	6.694825	391.716431	573.307526
	.389615	-134.370166	-132.753771	208.211328	4285.454346	6776.968750
3	46.941957	-2781.518066	3295.385788	-21.144561	2422.061523	-1294.793945
	-4.108434	-527.266357	591.814844	-648.658015	34355.003750	-15267.984668
4	4.693785	9383.161133	0.	5.564478	-412.632397	0.
	1.661047	6.492623	0.	-91.728977	-142971.537891	0.
5	-.677652	-1713.779633	-1333.754532	-3.736346	17.360059	4.733407
	-.172060	-0.043963	-2.622653	135.810375	26314.946289	20510.924463
6	13.574044	-9170.492432	2727.079285	15.914567	-541.707909	40.436359
	-0.492214	-7.722320	-4.907297	-261.435032	149190.640313	-45111.964844
IMAGE SPACE		0.				0.

AXIAL RAY DATA NO. OF A RAYS = 2 RAY NO. J = 1 X = 0. Y = 1.000000

ABERRATIONS ARE IN WAVELENGTH UNITS (= 0.00054610) W = 0.848 DW5-DASH = $\frac{\partial W_5}{\partial \bar{H}^2 \bar{\rho}^2}$ = 4.4620

DW/ DCBAR(MEAN WAVEL) = -1835.33592 DW/ DAN = $\frac{\partial W}{\partial A_N}$ = -4.46660

SURF.	DW/DZ ($\frac{\partial W}{\partial Z}$)	DW/DC	DW/DE	DW/DC ($\frac{\partial W}{\partial C}$)	DW/DAN ($\frac{\partial W}{\partial A_N}$)	DW/DAN	DW/DA6
	DW/DA8	DW/DA10	DW/DF	DW/DC	DW/DC	DW/DE	
1	14.146163	892.496709	0.	0.	97.557902	28.199841	
	8.151375	2.356216	0.294051	41.708380	0.		
2	-8.263339	-232.278819	-309.322910	0.	0.	0.	
	.	0.	-0.733159	-27.810745	-23.158239		
3	49.619751	-2914.872168	1427.198212	16.985931	-486.102813	-356.359219	
	-20.295350	-151.517001	2.734914	-147.049833	149.908362		
4	5.117407	9979.812378	0.	0.	0.	0.	
	0.	0.	0.365392	57.841980	0.		
5	-2.260026	-2162.992371	-1506.213211	-6.277231	-137.292582	-111.663045	
	-00.819285	-73.965830	-1.441363	-845.458344	-172.462006		
6	15.557770	-10529.206055	2977.942200	20.162511	-491.926903	-356.140533	
	-257.831233	-186.665039	2.613704	-932.293688	244.308426		
IMAGE SPACE		0.				0.	

AXIAL RAY DATA NO. OF A RAYS = 2 RAY NO. J = 2 X = 0. Y = 0.707107

ABERRATIONS ARE IN WAVELENGTH UNITS (= 0.00054610) W = 0.442 DW5-DASH = -1.7186

DW/ DCBAR(MEAN WAVEL) = -914.49765 DW/ DAN = 1.21446

SURF.	DW/DZ	DW/DC	DW/DE	DW/DC	DW/DAN	DW/DA6
	DW/DA8	DW/DA10	DW/DF	DW/DC	DW/DC	DW/DE
1	14.146163	892.496709	0.	0.	97.557902	28.199841
	8.151375	2.356216	0.294051	41.708380	0.	
2	-8.263339	-232.278819	-309.322910	0.	0.	0.
	.	0.	-0.733159	-27.810745	-23.158239	
3	49.619751	-2914.872168	1427.198212	16.985931	-486.102813	-356.359219
	-20.295350	-151.517001	2.734914	-147.049833	149.908362	
4	5.117407	9979.812378	0.	0.	0.	0.
	0.	0.	0.365392	57.841980	0.	
5	-2.260026	-2162.992371	-1506.213211	-6.277231	-137.292582	-111.663045
	-00.819285	-73.965830	-1.441363	-845.458344	-172.462006	
6	15.557770	-10529.206055	2977.942200	20.162511	-491.926903	-356.140533
	-257.831233	-186.665039	2.613704	-932.293688	244.308426	
IMAGE SPACE		0.				0.

AXIAL RAY DATA NO. OF A RAYS = 2 RAY NO. J = 3 X = 0. Y = 0.

ABERRATIONS ARE IN WAVELENGTH UNITS (= 0.00054610) W = 0. DW5-DASH = 0.

DW/ DCBAR(MEAN WAVEL) = 0. DW/ DAN = 0.

SURF.	DW/DZ	DW/DC	DW/DE	DW/DC	DW/DAN	DW/DA6
	DW/DA8	DW/DA10	DW/DF	DW/DC	DW/DC	DW/DE
1	14.146163	892.496709	0.	0.	97.557902	28.199841
	8.151375	2.356216	0.294051	41.708380	0.	
2	-8.263339	-232.278819	-309.322910	0.	0.	0.
	.	0.	-0.733159	-27.810745	-23.158239	
3	49.619751	-2914.872168	1427.198212	16.985931	-486.102813	-356.359219
	-20.295350	-151.517001	2.734914	-147.049833	149.908362	
4	5.117407	9979.812378	0.	0.	0.	0.
	0.	0.	0.365392	57.841980	0.	
5	-2.260026	-2162.992371	-1506.213211	-6.277231	-137.292582	-111.663045
	-00.819285	-73.965830	-1.441363	-845.458344	-172.462006	
6	15.557770	-10529.206055	2977.942200	20.162511	-491.926903	-356.140533
	-257.831233	-186.665039	2.613704	-932.293688	244.308426	
IMAGE SPACE		0.				0.

DATA FROM PUPILT

FIELD	COLOUR	HSRAD	HOBAR	DHBASD	BADBND	BPERMC
		DHSRAD	PIDASH	DYBASD	WAPUPIL	HOBARD
1.0000	5461.000000	0.287957	0.286525	0.	1.000000	1.000000
		0.	1.000437	0.	-0.000085	0.287817

DATA FROM PUPILT

FIELD	COLOUR	HSRAD	HOBAR	DHBASD	BNDBND	BPERMC
		DHSRAD	PIDASH	DYBASD	WAPUPIL	HOBARD
1.0000	5461.000000	0.287957	0.286525	0.000106	1.000001	0.999631
		0.000106	1.000120	0.000446	-0.000084	0.287923

DATA FROM PUPILT

FIELD	COLOUR	HSRAD	HOBAR	DHBASD	BNDBND	BPERMC
		DHSRAD	PIDASH	DYBASD	WAPUPIL	HOBARD
1.0000	5461.000000	0.287957	0.286525	0.000106	1.000001	0.999631
		0.000106	1.000120	0.000446	-0.000084	0.287923

T-RAY DATA

FIELD = 1.000000 WAVELENGTH = 5461.000000 RTRAD = 0.12139691 ROWEN = 0.9857677

UNITS HERE ARE NOT IN WAVELENGTHS , DWDE ETC. NEED NOT BE TESTED FOR DISPERSION.

SURF.	DWT/DC	DWT/DC	DWT/DZ	DWT/DZ
	DWT/DE	DWT/DE	DWT/DC	DWT/DC
	DWT/DA	DWT/DA	DWT/DA6	DWT/DA6
	DWT/DA8	DWT/DA8	DWT/DA10	DWT/DA10
1	-0.637772545	0.217320329	0.007736635	0.001223636
	-0.122902036	0.417115377	-0.	0.
	-0.013394747	0.810678449	0.003563695	0.000841070
	-0.000726871	0.000000269	0.000021802	0.000007429
2	0.412678566	-0.431332776	-0.000854283	0.000503313
	0.155535065	-0.047684247	0.	0.
	0.	0.	0.	0.
	0.	0.	0.	0.
3	0.563664134	-0.102466668	0.001607728	-0.000715476
	0.180753991	-0.029476964	-0.000207119	0.000039417
	0.006926276	-0.001259558	0.000052925	-0.000045941
	0.000000233	-0.000001483	-0.	0.
4	2.604355168	0.574490842	0.004254691	0.000933659
	0.712093611	0.167129248	0.	0.
	0.	0.	0.	0.
	0.	0.	0.	0.
5	-0.484274253	-1.124366559	-0.003975942	-0.003176459
	-0.856122126	0.215007607	-0.000166420	-0.000044691
	-0.083961200	-0.030315596	-0.000283688	-0.000072660
	-0.000018045	-0.000004819	0.000000431	0.000000112
6	-2.944926679	-0.850111544	0.013251114	0.004531864
	-0.533086350	-0.132264733	0.001160713	0.000340067
	-0.023985159	-0.006917014	-0.000206326	-0.000062633
	-0.000161402	-0.000046437	-0.	-0.

IMAGE SPACE DWO/DE = -0.192.52462 -0.055635735

	WAVELENGTH =	5461.000000000	TOTAL NO. OF X RAYS =	2	(Axial Ray Case)
RAY NO.	HDASH (Y)	HSCH-DASH (Y')	DELTAN (nY')	YCAXIL (Y axial)	YCAXIL (Y axial)
1	0.0024136883	0.286525369	-0.012139269	1.000130492	0.01435487
2	-1.000938542	0.286525369	0.000333583	0.726655126	0.00715264
	WAVELENGTH =	4563.000000000	TOTAL NO. OF X RAYS =	2	
RAY NO.	HDASH	HSCH-DASH	DELTAN	YCAXIL	ZCAXIL
1	0.001644452	0.286525036	-0.000833587	1.000242190	0.01434703
2	-0.000875969	0.286525036	0.000233942	0.726461739	0.00715581
	WAVELENGTH =	4861.000000000	TOTAL NO. OF X RAYS =	2	
RAY NO.	HDASH	HSCH-DASH	DELTAN	YCAXIL	ZCAXIL
1	0.0043107736	0.2866431166	-0.0021553106	1.0007158138	0.01437123
2	-0.0010429891	0.2866431166	0.0005518833	0.7268631809	0.00715470

FIELD = 1.000000 WAVELENGTH = 5461.000000000 TOTAL NO. OF X RAYS = 2 TOTAL NO. OF S RAYS = 1 (Entire - axial Ray Case)

ABERRATIONS AND DERIVATIVES ARE IN WAVELENGTH UNITS (= 0.0005460998)

J-TH. RAY = 1 WAVE ABERRATION = 1.3789A111

DW/DW20 (MEAN WAVE) = -845.1629559955 DW/DW1 (MEAN WAVE) = -1237.801727295

DW/DS1 = 0. DW/DS2 = 1.285775518 DW/DS3RAYS = 3.464235107

ADDITIONAL DATA FOR CHECKING ONLY UNITS ARE NOT IN WAVELENGTHS

YDASH=	-0.000011180	DSDASH=	-0.	-0.539610685	XDASH= -0.
YZDASH=	0.875965473	ZSDASH=	-0.011220443	0.000279278	XSDASH= -0.
YSDASH=	0.473963528	ZSDASH=	-0.031133727	PUPILX= 0.888425931	PUPILY= 0.479032414
		HOBARD=	0.287006319	DSDASH= -0.	

SURF. DW/DC (W_{040}) DW0/DZ (W_{131}) DW/DC (W_{222}) DW0/DC (W_{220})

DW/DA (W_{040}) DW0/DA6 DW/DA6 DW0/DA6

BELOW ARE FOR CHECKING ASRAYS ONLY

	DWDC (W_{040})	DWDZ	DWDE (W_{311})	DWDC
	DWDA (W_{040})	DWDA6	DWDA8	DWDA10
	DYSDC	DYSDC (W_{222})	DYSDZ	DYSDZ
	DYSDE (W_{222})	DYSDE (W_{220})	DYSDC	DYSDC
	DYSDA	DYSDA6	DYSDA8	DYSDA10
1	65.84350B995	1.035110533.	0.	0.
	41.831675581	4.060936412	0.355908297	0.029478698
	0.024581334	0.012606259	-0.213685379	0.
	0.001233636	0.000130566	0.000052398	0.000001110
	0.	-0.746791653	-0.	0.025040343
	0.	-0.134713145	0.	0.
	-0.003731905	-0.003163585	-0.000037034	-0.000002101
2	-19.75547136	-1.872836798	-15.08478792	0.
	0.	0.	0.	0.
	-0.025656748	-0.004821238	-0.010197922	0.
	0.	0.	0.	0.
	-0.	0.180361230	0.	0.000735201
	0.	0.023694205	0.	0.
	0.	0.	0.	0.
3	-98.885491362	1.615131152	-435.0730C9C9	1.436126128
	-46.835057321	-8.630214047	-1.572036266	-0.315773549
	-0.386868794	0.013606128	0.109752052	0.000642553
	-0.020856349	-0.004212253	-0.000852868	-0.000172444
	-0.	0.976835614	-0.	-0.020983706
	-0.	0.143124173	0.	-0.000389568
	-0.	-0.	0.	0.
	0.012074672	0.000441104	0.000031327	0.
4	451.6070175??	0.065213256	0.	0.
	0.	0.	0.	0.
	4.2841450667	0.004131201	-0.263986404	0.
	0.	0.	0.	0.023042299
	-0.	7.699550271	-0.	0.
	-0.	-0.548877058	0.	0.
	0.	0.	0.	0.
5	-3572.207780004	-1.433330134	-1672.336990356	-4.415585424
	-93.215022926	-72.267677204	-51.015362740	-15.820847509
	-0.956605181	-0.055150754	-0.446537094	-0.072752468
	-0.054338310	-0.039722033	-0.027871010	-0.019561186
	0.	-1.585116514	-0.	-0.015860151
	0.	2.778762068	0.	-0.000548206
	0.	0.	0.	-0.
	-0.001242798	-0.000925481	-0.000058852	0.000001423
6	-455.019195557	1.835544676	1829.286784459	19.237304419
	-364.047504425	-257.224853516	-169.988405225	-111.480837186
	-4.805917024	0.012483381	1.318492287	0.011330693
	-0.215187815	-0.141895428	-0.092942904	-0.060859429
	0.	-9.936490686	-0.	0.044869321
	0.	-1.724594493	-0.	0.003953446
	0.	0.	0.	0.
	-0.002055923	-0.000762164	-0.000054485	-0.

IMAGE SPACE DWO/DE = 0. WAVELENGTHS

DWDE = -0.631534107 DXSDE = 0. DYSDE = -0.649999507

FIELD = 1.000000 WAVELENGTH = 5461.000000000 TOTAL NO. OF X RAYS = 1 TOTAL NO. OF S RAYS = 1

ABERRATIONS AND DERIVATIVES ARE IN WAVELENGTH UNITS (= 0.0005460998)

J-TH. RAY = 2 WAVE ABERRATION = 0.8359816A

DW/DW20 (MEAN WAVE) = -421.3193364455 DW/DW1 (MEAN WAVE) = -876.1778310469

DW/DS1 = 0. DW/DS2 = 2.2A1105550 DW/DS3RAYS = 3.126572814

ADDITIONAL DATA FOR CHECKING ONLY UNITS ARE NOT IN WAVELENGTHS

DXSDA=	-1.201192A6A	DSDASA=	-0.	CULEX=	-0.291553554A XDASH= 0.
YDASH=	0.476644A579	ZSDASH=	-0.013437251	0.000413260	XSDASH= -0.
YSDASH=	0.878940734	ZSDASH=	-0.023137725	PUPILX=	0.8504525923 PUPILY= 0.479032414
		HOBARD=	0.286406462	CDTEX=	-0.

SURF. DW0/DC DW0/DZ DW0/DE DW0/DC

DW0/DA DW0/DA6 DW0/DA8 DW0/DA10

BELOW ARE FOR CHECKING ASRAYS ONLY

	DWDC	DWDZ	DWDE	DWDC
	DWDA	DWDA6	DWDA8	DWDA10
	DYSDC	DYSDC	DYSDZ	DYSDZ
	DYSDE	DYSDE	DYSDC	DYSDC
	DYSDA	DYSDA6	DYSDA8	DYSDA10
1	19.302638531	0.411875654	0.	0.
	23.164450177	2.402556433	0.215138726	0.018675444
	-0.005549445	0.007432744	-0.048891035	0.
	-0.002350229	-0.000171930	-0.000011258	-0.000000730
	0.	-0.816476910	-0.	0.019677392
	0.	-0.066195869	0.	0.
	0.	0.	0.	0.
	-0.015744954	-0.001153664	-0.000133544	-0.000010907
2	-5.352539897	-0.224594350	-7.264439867	0.
	0.	0.	0.	0.
	0.005544275	-0.002207321	-0.000302538	0.
	-0.	0.	0.	0.
	-0.	0.892076430	0.	0.021331291
	-0.	0.096086549	0.	0.
	0.	0.	0.	0.
	-0.	0.	0.	0.
3	-30.981987476	0.477621627	-455.402949730	0.219313504
	-10.595834732	-0.576789193	-0.133652312	-0.001910850
	-0.099173613	0.007078071	0.001168473	0.000074267
	-0.002437855	-0.000192072	-0.000014402	-0.000001055
	-0.	0.588506840	-0.	-0.021764509
	-0.	0.216166552	0.	-0.000245959
	-0.	-0.	0.	0.
	0.007271156	0.000226495	0.000006828	0.
4	163.110059738	0.039276482	-0.	0.
	0.	0.	0.	0.
	2.429648881	0.002161810	-0.127439208	0.
	-0.	0.003267060	-0.	0.000091554
	-0.	-0.557839262	0.	0.
	0.	0.	0.	0.
5	-122.001934052	-0.331467978	-1131.184588845	-1.387604834
	-32.503335026	-15.043310195	-8.477877805	-2.705951589
	-0.314584705	-0.008526383	-0.074430276	-0.000810046
	-0.019455198	-0.002833562	-0.010513878	-0.001674437
	-0.	-1.264496179	-0.	-0.012803897
	-0.	2.184782808	0.	-0.000432715
	-0.	0.	0.	-0.
	-0.010177582	-0.000728761	-0.000046335	0.000001125
6	-152.508295859	0.480585815	121.039718528	0.814377246
	-134.902164383	-83.703303337	-26.986331325	-11.386210844
	-2.701554772	0.007852457	0.755236611	0.004285167
	-0.085244603	-0.035791367	-0.034803929	-0.006110430
	-0.	-7.786038473	-0.	0.038316790
	-0.	-1.354436182	-0.	0.003143953
	-0.	0.	0.	0.
	-0.044132158	-0.005195259	-0.000432916	-0.

IMAGE SPACE DW0/DE = 0. WAVELENGTHS

DWDE = -0.360552058 DXSDE = 0. DYSDE = -0.516459231

CHAPTER 8

The evaluation of the merit function, its parameter derivatives and the optimisation procedure

(VIII-i) Introduction

1) Notations for the changes in parameters.

For convenience, the changes in constructional parameters will be denoted by p_r, r=1,2,3,...,R. The first q parameters will be used to denote changes in the image position and image heights, corresponding to the shifts of the centres of the $\bar{\lambda}$ reference sphere. Thus, if the number of extra-axial image fields are denoted by NFIELD, then

$$p_1 = (\delta_o W_{20})_{\bar{\lambda}}$$

$$p_2 = (\delta W_{11})_{\bar{\lambda}, \tau_1}$$

$$p_3 = (\delta W_{11})_{\bar{\lambda}, \tau_2}$$

.....

$$p_q = (\delta W_{11})_{\bar{\lambda}, \tau_{q-1}} \qquad q = (\text{NFIELD} + 1)$$

The parameters p_{q+1}, p_{q+2}, ..., p_R denote the vertex curvatures C of surfaces, their axial positions z, the refractive indices μ, and the aspheric factors Q and A_4, A_6, A_8, A_{10}, taken in this order. The wave-aberration parameter derivatives required for the procedures of this chapter are the derivatives

$\partial W_o/\partial p_r$ for the current "initial" system, which will have been computed and stored by the computer programme described in the previous chapter.

2) Target values

In the optimisation procedure to be described later, the effective values of the wave-aberration along the j^{th} ray, W_j, and the distortion $(\delta\bar{H}_o/\bar{H}_o)_{\bar{\lambda}}$ will be taken to mean the net values departing from their respective target values W_j^* and $(\delta\bar{H}_o/\bar{H}_o)_{\bar{\lambda}}^*$. Thus, the effective values are :

$$W_j = (W_j)_{\text{ray trace}} - W_j^* \qquad (8.1)$$

$$\left(\frac{\delta \bar{H}_o}{\bar{H}_o}\right) = \left(\frac{\delta\bar{H}_o}{\bar{H}_o}\right)_{\text{ray trace}} - \left(\frac{\delta\bar{H}_o}{\bar{H}_o}\right)^* \qquad (8.2)$$

Similarly, for the gaussian properties, we define the effective values

$$\left(\frac{\delta\bar{u}'}{\bar{u}'}\right) = \frac{\bar{u}' - (\bar{u}')^*}{\bar{u}'} \qquad (8.3)$$

$$\left(\frac{\delta\eta'}{\eta'}\right) = \frac{\eta' - (\eta')^*}{\eta'} \qquad (8.4)$$

where $(\bar{u}')^*$ and $(\eta')^*$ are the desired values of $\bar{u}'$ and η' respectively. In practice, the quantities on the right hand sides of (8.2), (8.3) and (8.4), relating to the size and distortion of the image, are required to be minimised for $\lambda = \bar{\lambda}$

only.

The effective values of $(\delta_o W_{20})_{\bar{\lambda}}$ is the longitudinal shift of focus between the image plane of the current 'initial' system and the desired (target) image position; the latter being defined by the target values (h')* and (u')*. Thus we have

$$(u')^* = \frac{(h')^*}{(R')^*} \tag{8.5}$$

where (h')* is the target value of the paraxial ray height at the exit pupil of the current 'initial' system; this exit pupil is at an axial distance (R')* from the target image plane. If h' and R' are corresponding quantities for the current 'initial' system, we have

$$u' = \frac{h'}{R'} \tag{8.6}$$

and the effective value of $(\delta_o W_{20})_{\bar{\lambda}}$ is then given by

$$(\delta_o W_{20})_{\bar{\lambda}} = \frac{n'}{2} \cdot (h')^2 \cdot \left[\frac{1}{R'} - \frac{1}{(R')^*} \right] \tag{8.7}$$

Substituting (8.5) and (8.6) in (8.7), we have

$$(\delta_o W_{20})_{\bar{\lambda}} = \frac{n'}{2} \cdot \left\{ \frac{h'}{(h')^*} \right\} \cdot \left[(h')^* u' - h' (u')^* \right] \tag{8.8}$$

3) Chromatic aberrations and target values.

As described in Chapter 6, section (VI-i), the wave-aberration calculated for each wavelength is referred to a reference sphere at its own exit pupil, but centred on the $\bar{\lambda}$ image point. Since the λ ray is traced with the same entrance pupil co-ordinates as the corresponding ray for $\bar{\lambda}$, the chromatic aberration for the λ ray will be given by

$$\partial W_\lambda = W_\lambda - W_{\bar{\lambda}} \qquad (8.9)$$

If the target value of ∂W_λ is ∂W_λ^*, the effective value of ∂W_λ for the current 'initial' system is defined as

$$\partial W_\lambda = \left[\partial W_\lambda\right]_{\text{ray trace}} - \partial W_\lambda^* \qquad (8.10)$$

If, for the extra-axial ray case, the ray co-ordinates employed are as defined in equation (5.9), the chromatic aberrations may be resolved into its even and odd parts, parallel to that described for the wave aberrations in Chapter 5 (see equations 5.11 and 5.13). Thus

$$(\partial U_\lambda)_j = \frac{(\partial W_\lambda)_j + (\partial W_\lambda)_{j+N}}{2} \qquad (8.11)$$

$$(\partial V_\lambda)_j = \frac{(\partial W_\lambda)_j - (\partial W_\lambda)_{j+N}}{2} \qquad (8.12)$$

and corresponding to (8.10), the effective values of the even and odd chromatic aberrations are defined as :

$$(\partial U_\lambda)_j = \left[(\partial U_\lambda)_j\right]_{\text{ray trace}} - (\partial U_\lambda)_j^* \tag{8.13}$$

$$(\partial V_\lambda)_j = \left[(\partial V_\lambda)_j\right]_{\text{ray trace}} - (\partial V_\lambda)_j^* \tag{8.14}$$

where $(\partial U_\lambda)_j^*$ and $(\partial V_\lambda)_j^*$ are the respective target values.

The parameter derivative of ∂W_λ (see equation 8.9) is given by

$$\frac{\partial}{\partial p}(\partial W_\lambda) = \left(\frac{\partial W}{\partial p}\right)_\lambda - \left(\frac{\partial W}{\partial p}\right)_{\bar{\lambda}} \tag{8.15}$$

where $p = p_{q+1}, p_{q+2}, \ldots, p_R$, and similarly

$$\frac{\partial}{\partial p}(\partial U_\lambda) = \left(\frac{\partial U}{\partial p}\right)_\lambda - \left(\frac{\partial U}{\partial p}\right)_{\bar{\lambda}} \tag{8.16}$$

$$\frac{\partial}{\partial p}(\partial V_\lambda) = \left(\frac{\partial V}{\partial p}\right)_\lambda - \left(\frac{\partial V}{\partial p}\right)_{\bar{\lambda}} \tag{8.17}$$

where

$$\frac{\partial U}{\partial p} = \frac{\left(\frac{\partial W}{\partial p}\right)_j + \left(\frac{\partial W}{\partial p}\right)_{j+N}}{2} \tag{8.18}$$

$$\frac{\partial V}{\partial p} = \frac{\left(\frac{\partial W}{\partial p}\right)_j - \left(\frac{\partial W}{\partial p}\right)_{j+N}}{2} \tag{8.19}$$

are the derivatives of the even and odd parts of W.

(VIII-ii) Definition of the merit function Ψ.

An introductory account of the use of a merit function and the use of the method of damped least-squares in optical design has been given in Chapter 1. Thus in equation (1.26) the merit function ψ was defined as

$$\psi = \sum_{i=1}^{m} W_i^2 + S^2 \sum_{j=1}^{n} c_j^2 p_j^2 \qquad (8.20)$$

In the present work, it is proposed to use the variance E of the wave-aberrations as the criterion of image quality (see Chapter 5, section V-iii). The merit function Ψ to be employed here is defined as the sum of the image quality merit function Φ, the boundary condition function χ, and the damping term

$$S_r^2 \sum_{r=1}^{R} C_r^2 p_r^2$$

where S_r and C_r are weighting factors similar to those which have been described in Chapter 1, equations (1.25) and (1.26). Thus

$$\Psi = \Phi + \chi + S_r^2 \sum_{r=1}^{R} C_r^2 p_r^2 \qquad (8.21)$$

It is convenient here to recapitulate the definition for C_r, which has the same significance as C_j in equations (8.20)

(1.25) and (1.26), that is

$$c_r = \left[\sum_j \left(\frac{\partial W_j}{\partial P_r} \right)^2 \right]^{\frac{1}{2}}_{r>q} \qquad (8.22)$$

$$= 1 \quad \text{if} \quad r \leqslant q$$

so that each of the system parameter p_r $(r>q)$ is weighted differently according to the 'sensitivity' of that particular parameter on the aberrations.

Referring now to the image quality terms in the merit function Φ in (8.21), the criterion for these will be the variances E of the wave-aberrations. For the current 'initial' system, equation (5.33) gives for the variance of the wave-aberrations of an axial pencil the expression

$$E_{axial} = \sum_{i=1}^{A} \sum_{j=1}^{A} P(i,j)\, W_i W_j \qquad (8.23)$$

and for each extra-axial image field position τ, equations (5.31) and (5.32) give for E_{even} and E_{odd} the expressions

$$E_{even} = \sum_{i=1}^{N} \sum_{j=1}^{N} Q(i,j)\, U_i U_j \qquad (8.24)$$

$$E_{odd} = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} R(i,j)\, V_i V_j \qquad (8.25)$$

where W_j, U_j and V_j are now the effective values, that is with the respective target values subtracted out. The values

of the coefficients P(i,j), Q(i,j) and R(i,j) are those tabulated in Chapter 5.

To account for the fractional distortion corresponding to each value of τ, the term

$$\left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right)^2_{\lambda = \bar{\lambda}}$$

for each extra-axial point is included in the image quality merit function Φ. Once again, the effective values of $(\delta \bar{H}_o'/\bar{H}_o)$ are used, as defined in equation (8.2).

The actual form of the function Φ will depend on the number of rays traced for $\bar{\lambda}$ and $\lambda \neq \bar{\lambda}$; that is, on the values of (2A+S) and (2B+T). Thus for the special case when (2B+T)=(2A+S), the image quality merit function Φ for the current 'initial' system will be defined as a sum of the above terms, each being multiplied by a suitable weighting factor, as follows :

$$\Phi = \sum_{\lambda} (B_{axial})_\lambda \cdot (E_{axial})_\lambda + \sum_{\lambda} \sum_{\tau} C_{\lambda,\tau} (E_{even})_{\lambda,\tau}$$

$$+ \sum_{\lambda} \sum_{\tau} D_{\lambda,\tau} (E_{odd})_{\lambda,\tau} + \sum_{\tau} \alpha_\tau \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right)^2_\tau \qquad (8.26)$$

where the weighting factors associated with the wavelength λ are :

$(B_{axial})_\lambda$ - for axial image

$C_{\lambda,\tau}$ - for the even-aberrations, corresponding to the extra-axial image field τ.

$D_{\lambda,\tau}$ - for the odd-aberrations and field τ.

The weighting factor α_τ is associated with the distortion corresponding to the extra-axial image field τ, and $\lambda = \bar{\lambda}$.

When (2B+T) < (2A+S), that is, when the number of rays traced for $\lambda \neq \bar{\lambda}$ are less than that for $\bar{\lambda}$, the image quality function Φ for the current initial system will be defined in a general form as follows :

$$\Phi = (B_{axial})_{\bar{\lambda}} \cdot (E_{axial})_{\bar{\lambda}} + \sum_{\tau} C_{\bar{\lambda},\tau} (E_{even})_{\bar{\lambda},\tau}$$

$$+ \sum_{\tau} D_{\bar{\lambda},\tau} (E_{odd})_{\bar{\lambda},\tau} + \sum_{\tau} \alpha_\tau \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right)_\tau^2$$

$$+ \sum_{\ell=1}^{L} F_\ell (\partial W_\lambda)_\ell^2 + \sum_{m=1}^{M} G_m (\partial U_\lambda)_m^2 + \sum_{m=1}^{M'} H_m (\partial V_\lambda)_m^2 \qquad (8.27)$$

where the weighting factors B_{axial}, C, D, α are as defined above and correspond to the case when $\lambda=\bar{\lambda}$. The other weighting factors are :

F_ℓ - for the chromatic aberration of the ℓ^{th} axial ray the rays are numbered with $\ell=1,2,\ldots,L$ where L = (NOWAVE - 1).B.

NOWAVE is the total number of wavelengths involved Thus for λ_1, the first wavelength other than the mean $\bar{\lambda}$, the B *axial* rays are numbered with , , ℓ=1,2,...,B. For the next λ, i.e λ_2, ℓ=B+1,B+2, ...,2B. etc.

G_m - for the even part of the chromatic aberration of the m^{th} extra-axial ray; the rays are numbered with m=1,2,...,M where M = NFIELD.(NOWAVE - 1).(B+ $\frac{T+1}{2}$). NFIELD is the total number of extra-axial fields.

B = no. of aperture rays
T = no. of skew rays
} $\lambda \neq \bar{\lambda}$.

Thus, for the 1st. field and $\lambda=\lambda_1$, m=1,2,..., (B + $\frac{T+1}{2}$). The next (B+ $\frac{T+1}{2}$) rays are associated with the same field but $\lambda=\lambda_2$, etc.. After all the rays corresponding to the last λ have been counted, the next (B+ $\frac{T+1}{2}$) rays will then be for the 2nd field and $\lambda=\lambda_1$ etc.

H_m - For the odd parts of the chromatic aberration of the m^{th} extra-axial ray; the rays are numbered with m=1,2,...,M' where M'= NFIELD.(NOWAVE-1) x (B+ $\frac{T-1}{2}$). Thus, for the 1st field and $\lambda=\lambda_1$, m= 1,2,..., (B+ $\frac{T-1}{2}$). The next (B+ $\frac{T-1}{2}$) rays are associated with the same field but $\lambda=\lambda_2$, etc., so that the counting procedure is similar to that

described above.

The chromatic aberrations (∂W_λ) , (∂U_λ) and (∂V_λ) in (8.27) take the effective values as defined in equations (8.10), (8.13) and (8.14).

The boundary condition function χ in (8.21) will depend on the displacement of the gaussian image plane $(\delta_o W_{20})_{\bar{\lambda}}$ - see equation (8.8) - and the effective value of $(\delta\eta'/\eta')$ or $(\delta\bar{u}'/\bar{u}')$ - see equations (8.3) and (8.4). To ensure χ to be not less than zero, we use the following definition for the boundary condition function :

$$\chi = Z \cdot (\delta_o W_{20})^2_{\bar{\lambda}} + Y \left(\frac{\delta\eta'}{\eta'} \right)^2_{\bar{\lambda}} \qquad (8.28)A$$

or, alternatively (Z and Y are weighting factors)

$$\chi = Z \cdot (\delta_o W_{20})^2_{\bar{\lambda}} + Y \left(\frac{\delta\bar{u}'}{\bar{u}'} \right)^2_{\bar{\lambda}} \qquad (8.28)B$$

when the angular size of the image has to be conserved.

(VIII-iii) Parameter derivatives of the merit function Ψ.

Referring to equation (8.21), the expression for the total merit function Ψ may be differentiated with respect to the parameter p_s, $1 \leqslant s \leqslant R$, giving

$$\frac{\partial\Psi}{\partial p_s} = \frac{\partial\Phi}{\partial p_s} + \frac{\partial\chi}{\partial p_s} + 2\, G_s^2\, C_s^2\, p_s \qquad (8.29)$$

The conditions for Ψ to be a minimum is that the R.H.S. of equation (8.29) should be zero, that is

$$\frac{\partial \Phi}{\partial p_s} + \frac{\partial \chi}{\partial p_s} + 2 S_s^2 C_s^2 p_s = o \qquad (8.30)$$

with s=1,2,...,R. giving R damped normal equations with R unknowns, the latter being the parameter changes $p_1, p_2, \ldots p_R$. When these parameter changes are introduced into the system, the wave-aberrations W_j will also change, resulting in new values of the variances E_{axial}, E_{even} and E_{odd}, and the chromatic aberrations ∂W_λ, ∂U_λ and ∂V_λ. To examine in detail the consequences of all these new values on the values of Φ and $\partial\Phi/\partial p$ would incur lengthy algebraic expressions. For convenience, therefore, we consider a typical case with Φ represented by equation (8.27); moreover, we consider only the new values of E and ∂W_λ corresponding to the axial image case; the extra-axial cases will later be generalised from these results. Thus for the special case considered, the new value of Φ becomes

$$\Phi = \left\{ B_{axial} \sum_{i=1}^{\Lambda} \sum_{j=1}^{\Lambda} P(i,j) \left[W_i + \sum_{r=1}^{R} \frac{\partial W_i}{\partial p_r} \cdot p_r \right] \cdot \left[W_j + \sum_{r=1}^{R} \frac{\partial W_j}{\partial p_r} \cdot p_r \right] \right\}_{\bar{\lambda}}$$

$$+ \sum_{\tau} \alpha_\tau \left[\left(\frac{\delta \bar{H}'_o}{\bar{H}_o} \right) + \sum_{r=q+1}^{R} \frac{\partial}{\partial p_r} \left(\frac{\delta \bar{H}'_o}{\bar{H}_o} \right) p_r \right]^2_{\tau,\bar{\lambda}}$$

$$+ \sum_{\ell=1}^{L} F_\ell \left\{ (\partial W_\lambda) + \sum_{r=q+1}^{R} \left[\frac{\partial}{\partial p_r} (\partial W_\lambda) \right] . p_r \right\}_\ell^2 \qquad (8,31)$$

where $\frac{\partial W_i}{\partial p_r}$ and $\frac{\partial W_j}{\partial p_r}$ are the values of $\frac{\partial W_o}{\partial p}$ computed in Chapter 7 for the axial rays i and j; $\frac{\partial}{\partial p_r} (\partial W_\lambda)$ is as defined in (8.15).

The parameter derivative of Φ, from (8.31), is

$$\frac{\partial \Phi}{\partial p_s} = B_{axial} \sum_{i=1}^{A} \sum_{j=1}^{\Lambda} P(i,j) \left\{ \frac{\partial W_i}{\partial p_s} \left[W_j + \sum_{r=1}^{R} \frac{\partial W_j}{\partial p_r} . p_r \right] + \left[W_i + \sum_{r=1}^{R} \frac{\partial W_i}{\partial p_r} . p_r \right] \frac{\partial W_j}{\partial p_s} \right\}_{\bar{\lambda}}$$

$$+ \sum_{\tau} \alpha_\tau . 2 \left[\left(\frac{\delta \bar{H}'_o}{\bar{H}_o} \right) + \sum_{r=q+1}^{R} \frac{\partial}{\partial p_r} \left(\frac{\delta \bar{H}'_o}{\bar{H}_o} \right) p_r \right]_\tau \left\{ \frac{\partial}{\partial p_s} \left(\frac{\delta \bar{H}'_o}{\bar{H}_o} \right)_\tau \right\}_{s \geqslant (q+1)}$$

$$+ \sum_{\ell=1}^{L} F_\ell . 2 \left\{ (\partial W_\lambda) + \sum_{r=q+1}^{R} \left[\frac{\partial}{\partial p_r} (\partial W_\lambda) \right] p_r \right\}_\ell . \left[\frac{\partial}{\partial p_s} (\partial W_\lambda)_\ell \right]_{s \geqslant q+1}$$

(8.32)

We now introduce the following notation :

$$b(s,r) = \sum_{i=1}^{A} \sum_{j=1}^{A} P(i,j) \left\{ \frac{\partial W_i}{\partial p_s} . \frac{\partial W_j}{\partial p_r} + \frac{\partial W_i}{\partial p_r} . \frac{\partial W_j}{\partial p_s} \right\} \qquad (8.33)$$

$$b(s,o) = \sum_{i=1}^{A}\sum_{j=1}^{A} P(i,j) \left\{ \frac{\partial W_i}{\partial p_s} \cdot W_j + W_i \cdot \frac{\partial W_j}{\partial p_s} \right\} \qquad (8.34)$$

$$b(o,o) = \sum_{i=1}^{A}\sum_{j=1}^{A} P(i,j) \left\{ W_i W_j + W_i W_j \right\}$$

or

$$b(o,o) = 2 \sum_{i=1}^{A}\sum_{j=1}^{A} P(i,j) \left\{ W_i W_j \right\} \qquad (8.35)$$

and for $s \geqslant q+1$, $r \geqslant q+1$, we define

$$e(s,r) = 2 \left\{ \frac{\partial}{\partial p_s} \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right) \right\} \left\{ \frac{\partial}{\partial p_r} \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right) \right\} \qquad (8.36)$$

$$e(s,o) = 2 \left\{ \frac{\partial}{\partial p_s} \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right) \right\} \cdot \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right) \qquad (8.37)$$

$$e(o,o) = 2 \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right)^2 \qquad (8.38)$$

$$f_\ell(s,r) = 2 \left\{ \frac{\partial}{\partial p_s} (\partial W_\lambda) \right\}_\ell \left\{ \frac{\partial}{\partial p_r} (\partial W_\lambda) \right\}_\ell \qquad (8.39)$$

$$f_\ell(s,o) = 2 \left\{ \frac{\partial}{\partial p_s} (\partial W_\lambda) \right\}_\ell (\partial W_\lambda)_\ell \qquad (8.40)$$

$$f_\ell(o,o) = 2 (\partial W_\lambda)_\ell^2 \qquad (8.41)$$

The expression for $\frac{\partial \Phi}{\partial p_s}$ given in (8.32) may now be written

$$\frac{\partial \Phi}{\partial p_s} = B_{axial} \left\{ b(s,o) + b(s,1)p_1 + \ldots\ldots + b(s,R)p_R \right\}_{\bar{\lambda}}$$

$$+ \left\{ \sum_{\tau} \alpha_{\tau} \left[e(s,o) + e(s,q+1)p_{q+1} + e(s,q+2)p_{q+2} + \ldots + e(s,R)p_R \right]_{\tau} \right.$$

$$\left. + \sum_{\ell=1}^{L} F_{\ell} \left[f(s,o) + f(s,q+1)p_{q+1} + f(s,q+2)p_{q+2} + \ldots \right.\right.$$

$$\left.\left. \ldots + f(s,R)p_R \right]_{\ell} \right\}_{s \geqslant q+1} \qquad (8.42)$$

The value of Φ for the initial system without the parameter changes (i.e. $p_1 = p_2 = \ldots = p_R = o$) may now be written as

$$\Phi = B_{axial} \sum_{i=1}^{A} \sum_{j=1}^{A} P(i,j)\, W_i W_j + \sum_{\tau} \alpha_{\tau} \left(\frac{\delta \bar{H}_o'}{\bar{H}_o} \right)^2$$

$$+ \sum_{\ell=1}^{L} F_{\ell} (\partial W_{\lambda})_{\ell}^2 \qquad (8.43)$$

Using the notation introduced above, we have

$$\Phi_o = \tfrac{1}{2} B_{axial} \left[b(o,o) \right] + \tfrac{1}{2} \sum_{\tau} \alpha_{\tau}\, e(o,o)_{\tau}$$

$$+ \tfrac{1}{2} \sum_{\ell=1}^{L} F_{\ell}\, f_{\ell}(o,o) \qquad (8.44)$$

The parameter derivative of Φ_o for the initial system is

$$\left(\frac{\partial\Phi}{\partial p_s}\right)_o = B_{axial}\ .b(s,o) + \left[\sum_{\tau}\alpha_{\tau}\ e(s,o)_{\tau}\right]_{s\geqslant q+1}$$

$$+\left[\sum_{\ell=1}^{L} F_{\ell} f_{\ell}(s,o)\right]_{s\geqslant q+1} \qquad (8.45)$$

We now consider the general expression for $\partial\Phi/\partial p_s$ which may be generalised from (8.42) by the inclusion of the additional extra-axial terms. For this purpose, additional notation is required, and this is introduced as follows :

(a) Corresponding to b(s,r) in equation (8.33) for the axial image case, we use c(s,r) and d(s,r) for the extra-axial even and odd parts of the wave-aberrations respectively, where, for $\lambda=\bar{\lambda}$,

$$c(s,r)_{\tau} = \sum_{i=1}^{N}\sum_{j=1}^{N} Q(i,j)\left\{\frac{\partial U_i}{\partial p_s}\cdot\frac{\partial U_j}{\partial p_r} + \frac{\partial U_i}{\partial p_r}\cdot\frac{\partial U_j}{\partial p_s}\right\}_{\tau} \qquad (8.46)$$

$$d(s,r)_{\tau} = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} R(i,j)\left\{\frac{\partial V_i}{\partial p_s}\cdot\frac{\partial V_j}{\partial p_r} + \frac{\partial V_i}{\partial p_r}\cdot\frac{\partial V_j}{\partial p_s}\right\}_{\tau} \qquad (8.47)$$

For the special cases when r=o, and also when r=o, s=o, the quantities in { } of (8.46) and (8.47) reduce to similar expressions as those defined in (8.34) and (8.35).

(b) Corresponding to the axial chromatic aberration notation $f_\ell(s,r)$ in equation (8.39), we use $g_m(s,r)$ and $h_m(s,r)$ for the extra-axial even and odd parts of the chromatic aberrations. Thus, for $s \geqslant q+1$, $r \geqslant q+1$, we define

$$g_m(s,r) = 2\left\{\frac{\partial}{\partial p_s}(\partial U_\lambda)\right\}_m \cdot \left\{\frac{\partial}{\partial p_r}(\partial U_\lambda)\right\}_m \qquad (8.48)$$

$$h_m(s,r) = 2\left\{\frac{\partial}{\partial p_s}(\partial V_\lambda)\right\}_m \cdot \left\{\frac{\partial}{\partial p_r}(\partial V_\lambda)\right\}_m \qquad (8.49)$$

For the special cases when r=o, and also when r=o, s=o, the quantities in { } of (8.48) and (8.49) reduce to similar expressions as those defined in (8.40)and (8.41).

Using the above notation, the general expression for $\partial\Phi/\partial p_s$ as derived from (8.42) may now be written as

$$\frac{\partial\Phi}{\partial p_s} = \Bigg\{ B_{axial} \cdot b(s,o) + \sum_\tau C_\tau\, c(s,o)_\tau + \sum_\tau D_\tau d(s,o)_\tau + \Bigg[\sum_\tau \alpha_\tau\, e(s,o)_\tau + \sum_{\ell=1}^{L} F_\ell f_\ell(s,o) + \sum_{m=1}^{M} G_m\, g_m(s,o) + \sum_{m=1}^{M'} H_m\, h_m(s,o) \Bigg]_{s \geqslant q+1} \Bigg\}$$

$$+ \sum_{r=1}^{R} \Bigg\{ B_{axial}\, b(s,r) + \sum_\tau C_\tau\, c(s,r)_\tau + \sum_\tau D_\tau\, d(s,r)_\tau + \gamma \Bigg\} p_r \qquad (8.50)$$

where

$$\gamma = \left[\sum_{\tau} \alpha_{\tau}\, e(s,r)_{\tau} + \sum_{\ell=1}^{L} F_{\ell}\, f_{\ell}(s,r) + \sum_{m=1}^{M} G_m\, g_m(s,r) + \sum_{m=1}^{M'} H_m\, h_m(s,r) \right], \text{ if } \begin{cases} s \geqslant (q+1) \\ r \geqslant (q+1) \end{cases} \qquad (8.51)$$

and

$$\gamma = o \quad \ldots\ldots\ldots \text{ if } \begin{cases} s \leqslant q \\ r \leqslant q \end{cases}$$

The value of $\partial\Phi/\partial p_s$ in equation (8.30) is that given in (8.50) above. The generalised formula for Φ corresponding to the initial system (with $p_1 = p_2 = \ldots = p_R = o$) may be obtained from (8.44) as

$$\Phi_o = \tfrac{1}{2} \left\{ B_{axial} \cdot b(o,o) + \sum_{\tau} C_{\tau}\, c(o,o)_{\tau} + \sum_{\tau} D_{\tau}\, d(o,o)_{\tau} + \sum_{\tau} \alpha_{\tau}\, e(o,o)_{\tau} + \sum_{\ell=1}^{L} F_{\ell}\, f_{\ell}(o,o) \sum_{m=1}^{M} G_m\, g_m(o,o) + \sum_{m=1}^{M'} H_m\, h_m(o,o) \right\} \qquad (8.52)$$

and the parameter derivative of Φ_o may be generalised from (8.45) as

$$\left(\frac{\partial\Phi}{\partial p_s}\right)_o = B_{axial}\cdot b(s,o) + \sum_{\tau} C_\tau\, c(s,o)_\tau + \sum_{\tau} D_\tau\, d(s,o)_\tau$$

$$+\left[\sum_{\tau}\alpha_\tau\, e(s,o)_\tau + \sum_{\ell=1}^{L} F_\ell f_\ell(s,o) + \sum_{m=1}^{M} G_m g_m(s,o)\right.$$

$$\left. + \sum_{m=1}^{M'} H_m\, h_m\,(s,o)\right]_{s\geqslant q+1} \qquad (8.53)$$

Referring now to the boundary condition function χ as defined in (8.28)A, when the parameter changes $p_1, p_2, p_3, \cdots, p_R$ are introduced into the system, the value of χ becomes

$$\chi = Z\cdot\left[\delta_o W_{20} + \sum_{r=q+1}^{R}\frac{\partial_o W_{20}}{\partial p_r}\cdot p_r\right]^2_{\bar{\lambda}} + Y\left[\left(\frac{\delta\eta'}{\eta'}\right) + \sum_{r=q+1}^{R}\frac{1}{\eta'}\left(\frac{\partial\eta'}{\partial p_r}\right)p_r\right]^2_{\bar{\lambda}} \qquad (8.54)$$

and the parameter derivative for χ is

$$\frac{\partial\chi}{\partial p_s} = Z.\,2.\left[\delta_o W_{20} + \sum_{r=q+1}^{R}\frac{\partial_o W_{20}}{\partial p_r}\cdot p_r\right]_{\bar{\lambda}}\cdot\left(\frac{\partial_o W_{20}}{\partial p_s}\right)_{\bar{\lambda}} \qquad (8.55)$$

$$+ Y.2.\left[\left(\frac{\delta\eta'}{\eta'}\right) + \sum_{r=q+1}^{R}\frac{1}{\eta'}\left(\frac{\partial\eta'}{\partial p_r}\right)p_r\right]_{\bar{\lambda}}\cdot\left(\frac{1}{\eta'}\,\frac{\partial\eta'}{\partial p_s}\right)_{\bar{\lambda}}$$

with s = q+1, q+2, ...,R. We now introduce the following notation for the case $\lambda = \bar{\lambda}$:

$$z(s,r) = 2 \left(\frac{\partial_o W_{20}}{\partial p_s}\right) . \left(\frac{\partial_o W_{20}}{\partial p_r}\right) \tag{8.56}$$

$$z(s,o) = 2 \left(\frac{\partial_o W_{20}}{\partial p_s}\right) . \delta_o W_{20} \tag{8.57}$$

$$z(o,o) = 2 (\delta_o W_{20})^2 \tag{8.58}$$

$$y(s,r) = 2 \left(\frac{1}{\eta'}, \frac{\partial \eta'}{\partial p_s}\right) . \left(\frac{1}{\eta'}, \frac{\partial \eta'}{\partial p_r}\right) \tag{8.59}$$

$$y(s,o) = 2 \left(\frac{1}{\eta'}, \frac{\partial \eta'}{\partial p_s}\right) \left(\frac{\delta \eta'}{\eta'}\right) \tag{8.60}$$

$$y(o,o) = 2 \left(\frac{\delta \eta'}{\eta'}\right)^2 \tag{8.61}$$

(8.55) may now be written as

$$\frac{\partial \chi}{\partial p_s} = Z. z(s,o) + Y. y(s,o) + \sum_{r=q+1}^{R} \left[Z.z(s,r) + Y.y(s,r) \right] p_r \tag{8.62}$$

for $s \geqslant q+1$; and for the initial system

$$\chi_o = \tfrac{1}{2} \left[Z.z(o,o) + Y.y(o,o) \right] \tag{8.63}$$

$$\left(\frac{\partial \chi}{\partial p_s} \right)_o = Z.z(s,o) + Y.y(s,o) \tag{8.64}$$

The value of $\frac{\partial \chi}{\partial p_s}$ in equation (8.30) is that given in (8.62) above.

(VIII - iv) Optimisation procedure.

It is proposed that the optimisation procedure should be carried out in two stages:

(a) The image quality merit function should be optimised first w.r.t. the focal shift parameters. The conditions required to be satisfied are

$$\frac{\partial \Phi}{\partial p_s} = 0 \quad \text{for} \quad s = 1, 2, \ldots, q.$$

Applying these conditions to equation (8.50), we then have q equations with the parameter changes $p_1, p_2, \ldots, p_q$ as the unknowns. The values of the coefficients b(s,r), c(s,r) and d(s,r) are calculated for s=1,2,...,q and r=o,1,2,..., q using the formulae given earlier; the values of W and $\partial W/\partial p$ required are those computed in Chapter 7. The solution of the q equations gives the best image plane and image heights for $\bar{\lambda}$, corresponding to the particular values of the weighting factors chosen for (8.50).

The new values of W_j to be used for the next stage are then given by

$$W_j + \sum_{s=1}^{q} \frac{\partial W_j}{\partial p_s} \cdot p_s$$

(b) Using these new values of W_j, we re-calculate the values of U_j, V_j, ∂U_λ, ∂V_λ and ∂W_λ. A complete set of coefficients b(s,r), c(s,r) and d(s,r) are then found for s=o,1,2,...,R and

r=o, 1, 2,...,R. Also required are the values of the coefficients $e(s,r)_T$, $f_\ell(s,r)$, $g_m(s,r)$ and $h_m(s,r)$ for $q < r \leqslant R$ and $q < s \leqslant R$, and for (r=o, s=o) and r=o. The coefficients calculated for (r=o, s=o) and r=o will be used for computing Φ_o, $(\partial\Phi/\partial p)_o$, χ_o and $(\partial\chi/\partial p_s)_o$ as defined in (8.52), (8.53), (8.63) and (8.64) respectively.

Referring now to equation (8.30), the conditions for the merit function Ψ to be a minimum are

$$\frac{\partial\Psi}{\partial p_s} = \frac{\partial\Phi}{\partial p_s} + \frac{\partial\chi}{\partial p_s} + 2 S_s^2 C_s^2 p_s = o \qquad (8.65)$$

where s=1,2,...,R and the values of $\partial\Phi/\partial p_s$ and $\partial\chi/\partial p_s$ are computed from formulae (8.50) and (8.62) respectively, using the new values of b(s,r) etc. as described above. From (8.65), the damped normal equations to be solved are

$$\frac{\partial\Psi}{\partial p_s} = K(s,o) + K(s,1)p_1 + \ldots + [K(s,s)+2S_s^2C_s^2]\, p_s + \ldots + K(s,R)p_R = 0 \qquad (8.66)$$

for s= 1,2,...,R; and where, from (8.50) and (8.62)

$$K(s,r) = B_{axial}\ .b(s,r) + \sum_T C_T\, c(s,r)_T + \sum_T D_T\, d(s,r)_T + \gamma$$

where,

$$\gamma = \sum_{\tau} \alpha_{\tau}\, e(s,r)_{\tau} + \sum_{\ell=1}^{L} F_{\ell} f_{\ell}(s,r) + \sum_{m=1}^{M} G_m\, g_m(s,r)$$

$$+ \sum_{m=1}^{M'} H_m\, h_m(s,r) + Z\, z(s,r) + Y\, y(s,r)$$

for $s = q+1, q+2, \ldots. R$ and $r=o, q+1, q+2, q+3, \ldots, R$.

$\gamma=o$ if $s = 1,2,\ldots.q$ and $r = 1,2,\ldots.,q$.

Thus if the weighting factors in $K(s,r)$ and the damping factors S_s are given, the R damped normal equations (8.66) may be solved to give the values of the parameter changes $p_1, p_2, p_3, \ldots, p_R$ to be introduced into the system. In practice, it will be necessary to impose additional boundary conditions to control the values of the edge thickness (for positive element) and of centre thickness (for negative element), the range of refractive indices, as well as the magnitudes of the parameter changes (see 'General requirements of a design programme', Chapter 1).

Discussion

The merit function Ψ and the optimisation procedure described in this chapter show how the weighted variances E of the wave-aberrations may be employed as the criterion of image quality in automatic lens design. One important feature of the method is that the use of the universal coefficients P(i,j), Q(i,j) and R(i,j) - see equations (8.23) to (8.25), and also Chapter 5 - facilitates rapid calculations of E and $\frac{\partial E}{\partial p}$. The computation procedure for the coefficients K(s,r) is relatively simple once the wave-aberrations W and the aberration derivatives $\frac{\partial W}{\partial p}$ for the current initial system are found. When parameter changes have been introduced into the system, W and $\frac{\partial W}{\partial p}$ must then be re-computed for this new system, followed by the coefficients K(s,r).

It may be pointed out that, as a special case, the universal coefficients P(i,j), Q(i,j) and R(i,j) may be set equal to unity for i=j, and zero for $i \neq j$; this will then give for the value of each E the sum of squares of the effective wave-aberrations of the rays. With minor modifications to the image quality merit function Φ (equation 8.27) for this special case, the problem of minimising Φ is then reduced to what is essentially the same as that described in Chapter 1, equation (1.23).

Special Supplement

1. In Chapter 2, we have introduced the following pupil variables:

$$x'_s = x_s \qquad \text{(see equation 2-17)}$$

and

$$y_s = \frac{Y}{h_s} = \frac{Y}{(\bar{N}h_T)} = y_T = y'_T = \frac{Y'}{(\bar{N}h'_T)} \qquad \text{(see equation 2-25)}$$

which are normalised by $h_s = \bar{N}h_T$. In general

$$y'_s = \frac{Y'}{h'_s} \neq y'_T$$

This is because when $h_s = \bar{N}h_T$, the expression for the anamorphotic ratio φ' from (2-20) is

$$\varphi' = \frac{h'_s}{(\bar{N}' h'_T)}$$

so that $h'_s \neq \bar{N}'h_T'$ unless φ' is unity.
The exact relation between y'_s and y'_T is

$$y'_T = y'_s \cdot \varphi'$$

The output from the computer programme described in Chapter 7 gives the values of φ', y'_s and x'_s. Since the values of x_s and y_s $(=y_T)$ are known from the starting data, one may readily calculate the values of y'_T and compare these with the corresponding values of y_s. Some numerical examples are presented below:

F/3.5 Triplet = 18° pencil

Peripheral Rays

x_s	$y_s = y_T$	x'_s	y'_T
0.000	0.835	0.000	0.834
0.714	0.591	0.712	0.590
1.010	0.000	1.013	-0.002
0.714	-0.591	0.719	-0.596
0.000	-0.835	0.000	-0.840

Watson 16 m.m. Para Objective

Extra-axial pencil

x_s	$y_s = y_T$	x'_s	y'_T
0.000	0.679	0.000	0.676
0.000	0.555	0.000	0.553
0.000	0.392	0.000	0.389
0.969	0.000	0.968	0.008

Watson 4 m.m. 0.65 N.A. Flat-field objective

Extra-axial pencil

x_s	$y_s = y_T$	x'_s	y'_s	y'_T
0.000	0.871	0.000	0.859	0.873
0.977	0.000	0.975	0.009	0.009
0.000	-0.871	0.000	0.851	-0.865

$$\rho' = 1.016$$

2. It has been described in Chapter 5 how the Strehl intensity ratio I may be calculated once the wave aberrations W_j are known. A numerical example is given below, illustrating the simplicity of the procedure involved in calculating the value of I for the axial image corresponding to the case number (2) in Table 5-1, with A = 3, S = 0. Thus, from (5-33) the value of E_{axial} is given by

$$E_{axial} = \sum_{i=1}^{A} \sum_{j=1}^{A} P(i,j)\, Wi\, Wj$$

and, for A = 3,

$$\begin{aligned} E_{axial} = {} & P(1,1)W_1^2+P(1,2)W_1W_2+P(1,3)W_1W_3 \\ & +P(2,1)W_2W_1+P(2,2)W_2^2+P(2,3)W_2W_3 \\ & +P(3,1)W_3W_1+P(3,2)W_3W_2+P(3,3)W_3^2 \end{aligned}$$

For a 16 m.m. microscope objective the wave aberrations W_j corresponding to the rays j = 1,2,3 are

$j = 1 \qquad W_1 = 0.081\lambda$

$j = 2 \qquad W_2 = 0.110\lambda$

$j = 3 \qquad W_3 = -0.023\lambda$

Substituting these W_j values in the expression for E_{axial} above and using the P(i,j) coefficients tabulated on P.168, we have the value of E_{axial} as 0.004, 917

The value of I may be calculated from the Maréchal's formula, that is,

from $$I = \left[1 - 2\pi^2 E_{axial}\right]^2$$

and for this particular example, I = 0.8153.

Since this is greater than 0.8, the Maréchal's formula is valid.

General Summary

Until recently most lens designers have tended to assess image quality indirectly in terms of the aberrations rather than directly in terms of the influence of the aberrations on the diffraction image itself. It follows from Maréchal's tolerance theory that in designing high quality systems, one should aim at minimising the mean squared deviation E of the wave surface from the spherical form, and this approach has been adopted in the present work.

A brief account has been given of the traditional design methods, the automatic design techniques and the advantages of using the techniques of modern image assessment to give the criterion of image quality. In the present work, the Strehl intensity ratio I was employed as the index of quality and a detailed investigation of its properties was described in Chapter 3. It was shown how I could be computed, using a numerical integration method, when the aberration coefficients were known.

By using canonical pupil variables, the wave-aberrations of an optical system may be expressed as a function of the entrance pupil co-ordinates (see Chapter 2). A pupil exploration computer programme was described in detail in Chapter 4, in which the pupil shape was specified in terms of the entrance pupil co-ordinates. This computer programme is of practical importance since it will locate the effective

central ray of an extra-axial pencil as well as the surface number (s) which is responsible for the vignetting; the pupil shape itself was represented by a best-fitting ellipse which was subsequently scaled to be a circle. In a large number of cases tested, it has been found possible to get a good representation of the pupil in this way. The wave-aberrations may then be expressed as a polynomial in this equivalent circular pupil co-ordinates; the ray pattern remains the same for all image fields. The advantage of the procedure lies in reducing all cases to the problem of a circular aperture, and the resulting possibility of calculating the variance E (both axial and extra-axial images) for any system very simply once the aberrations of the rays are known. This simplification follows directly from the use of universal coefficients P(i,j), Q(i,j) and R(i,j) and these were described in detail in Chapter 5, and a comprehensive set of tables of these coefficients was also presented.

In Chapter 6, an account was given of wave-aberration derivatives [Hopkins - private communication]. A computer programme for computing these derivatives was described in detail in Chapter 7, and an example of the computer results was also presented.

A useful criterion of image quality for the whole field is the sum of suitably weighted values of the variances E for a typical set of image points; a merit function based on

this was described in Chapter 8. An optimisation procedure was also proposed which required the merit function to be minimised w.r.t. the different design parameters. For this purpose, the universal coefficients P(i,j) etc. and the aberration derivatives described in Chapter 5 and 7 respectively were needed. The use of canonical pupil co-ordinates thus makes possible the use of a diffraction- based image criterion in automatic lens design.

ACKNOWLEDGEMENTS

I should like to express my deepest gratitude and sincere thanks to Dr. H.H. Hopkins for suggesting the subject of this thesis and for his advice, guidance, criticisms and for reading the manuscript.

I also wish to thank Miss. J.M. Drewitt and Mr. G.J. Church of Messrs. W. Watson and Sons, Ltd., Mr. N.J. Davidson of Imperial College, Mr. R. Cressy and Mr. I. Harold of the I.B.M. Data Centre, London, for their helpful discussions, and particularly Mr. G.J. Church for his assistance in checking some of the numerical results.

It is a pleasure to acknowledge the assistance of Dr. J.S. Wilczynski of the I.B.M. Research Centre, U.S.A., through whom computing facilities were made available for most of this work.

The provision by the Ministry of Aviation for the Bursary under which most part of this work was carried out is also gratefully acknowledged.

W.B.K. FEB. 1966.

W B King