Off-Axis Aberration Coefficients

P.J. Sands

Errata (as of January 1971)

P. 26	Eq.(4.4)	should read $q = P_yR_y + P_zR_z$.
P. 29	lines 19 and 20	replace $\underset{\sim}{Z}_1$ by $\underset{\sim}{V}_1$.
P. 32	line 23	replace Ⓗ by Ⓜ.
P. 33	line 2	omit subscript B on $Ⓒ_B$.
	line 15	$m = (m_y, m_z)$ should read $\underset{\sim}{m} = (m_y, m_z)$.
P. 36	line 10	omit subscript A from $\underset{\sim}{Y}_A$.
P. 37	line 8	replace $\underset{\sim}{I}_{Ax}$ by I_{Ax} .
	line 13	replace n by $\vec{n}$.
	line 15	replace $\underset{\sim}{I}$ by I_A .
P. 39	Eq.(10.7)	should read $\Delta\underset{\sim}{Y} = \bar{R}_B P(\underset{\sim}{I} + \underset{\sim}{I}_B)$.
P. 41	Eq.(11.7)	replace I_B by $\underset{\sim}{I}_B$.
P. 44	Eq.(13.1)	replace $\underset{\sim}{V}.\underset{\sim}{V}$ by $\alpha_B^2 \underset{\sim}{V}.\underset{\sim}{V}$.
P. 56	Eq.(19.1)	replace $s_b T_z$ by $z_b T_z$.
P. 70	line 10	the signs of G_{ya} and G_{za} should both be -
P. 73	lines 7 and 9	replace $\underset{\sim}{G}_i$ by $\Delta\underset{\sim}{G}_i$.
P. 75	Eq.(24.12)	replace $G_{\mu\nu\tau ai}^{(n)}$ by $g_{\mu\nu\tau ai}^{(n)}$, μ_k' by $\underset{\sim}{\mu}_k'$.
P. 76	Eqs.(25.4)	subscript y to be omitted from g_y .
	Eqs.(25.5)	in equation for V_j , δ_y and δ_v should read δ_{yj} and δ_{vj} .
P. 78	line 11	the first Q should be $\underset{\sim}{Q}$.
	Eq.(26.4)	$\underset{\sim}{q}^{(n)}$ should be $\hat{\underset{\sim}{q}}^{(n)}$.
	line 18	Y and V should be $\underset{\sim}{Y}$ and $\underset{\sim}{V}$.

P. 80	Eqs.(27.2)	the equation for ρ_1 is $\rho_1 = -c[\lambda_1^{-1} + \underset{\sim}{n}_B \cdot \underset{\sim}{V} - c\underset{\sim}{Y} \cdot (\underset{\sim}{V}+\underset{\sim}{V}_B)]$.
P. 91	table 30/1	the last term in the equation for $\hat{t}_{q1}$ is $\hat{s}_{q1}\ \hat{p}_{\psi 1}$.
P. 93	table 31/1 line 6	the second equation is $\hat{t}_{x14} = \hat{p}_{x1}\ \hat{s}_{x8}$.
P. 96	table 32/1	in the equation for $t_{\tau 8}$, replace $\bar{S}_B$ by $\bar{R}_B$.
P.111	Eqs.(40.4)	each G should be $\underset{\sim}{G}$.
	Eqs.(40.6,7)	in these three equations replace I_{By} by $\underset{\sim}{I}_B$.
P.267	line 1	replace "surface" by "transfer".
P.272	Eq.(89.2)	$(\sigma^2-\mu\tau)$ should read $(\sigma^2-\mu\tau)^{\frac{1}{2}}$.
P.327	line 12	replace N_{y1} by $\underset{\sim}{N}_1$.
	line 19	replace Sp(5) by Sb(5) .
P.162	Table 59/2	in the second equation, replace T_{By} by S_{By} .

To whom it may concern:

(i) The work in this thesis has been generalised to arbitrary plane symmetric surfaces and a paper reporting this has been submitted for publication.

(ii) A computer listing of an improved version of the programme of Appendix F is available on request from

Dr P.J. Sands,
Department of Theoretical Physics,
Faculty of Science,
Australian National University,
P.O. Box 4,
Canberra. A.C.T. 2601.
Australia.

OFF-AXIS ABERRATION COEFFICIENTS

by

P. J. SANDS

A thesis submitted for examination for the degree of Doctor of Philosophy, Australian National University

Department of Theoretical Physics
School of General Studies
Australian National University

April 1967

STATEMENT OF ORIGINALITY

I declare that the work embodied in this thesis was carried out entirely by myself.

Peter John Sands

ACKNOWLEDGEMENTS

It is with pleasure that I thank my supervisor, Professor H. A. Buchdahl, for his interest in this work and for the many discussions we had relating to various aspects of the theory and its applications. My grateful thanks go to Miss J. Flint for her painstaking typing of the stencils for the manuscript and for doing the duplicating. I cannot speak too highly of her handling of the many complicated equations appearing in the text.

My thanks are also due to my wife, Ritva, for assistance with the plotting of graphs and proof-reading; to Mr. P. Logan for preparing the line drawings for the graphs; and to the Visual Aids Unit for photographically reproducing the spot diagrams and pages 243-260. The computations were carried out on the IBM 360/50 at the A.N.U. Computer Centre and the originals for the spot diagrams were produced on the IBM 1620 plotter at Mt. Stromlo Observatory.

P. J. Sands

CONTENTS

I. INTRODUCTION

(a) The study of the aberrations of an optical system is, in effect, the study of the behaviour of some suitably defined functions representing the extent to which rays, after traversing the system, fail to pass through assigned points in the image space. In particular the Taylor series of these functions may be studied, the coefficients of which, the so-called aberration coefficients, are somehow to be calculated from the system parameters (i.e., the parameters specifying the constitution of the system). The radius of convergence of these series is, in general, unknown and varies from system to system. The aberration coefficients have sometimes been determined approximately by tracing a set of selected rays and fitting polynomials of a pre-assigned order to the results of these traces.[1] The point of view adopted here is that the aberration coefficients themselves shall be exact. To determine these coefficients it is of course necessary to have algebraic formulae expressing them either directly or indirectly in terms of the system parameters. It is the determination of these formulae with which this thesis is, in part, concerned.

Very little work has been done in this respect for systems other than symmetric systems, that is, systems which are both axially symmetric and are invariant under reflection in any plane containing the axis of symmetry, or optical axis, of the system. However, the work of Smith[2] and Wynne[2] may be consulted. For symmetric systems arbitrary rays have been traditionally referred to the ray which lies along the

---

1 See, for example, Herzberger (1958) p.385.

2 Smith (1928a,b) and Wynne (1954). See also §b on p.3.

optical axis, in which case the subsequent usefulness of the series representing the aberration functions is restricted to a certain neighbourhood of the axis. The axial theory of Buchdahl[3] is an example of a theory by means of which the exact aberration coefficients up to the seventh and higher orders may, in practice, be calculated for symmetric systems. It is a matter of experience that when angles of incidence or refraction become sufficiently large no reliance can be placed on the aberrations predicted using the Taylor series for aberration functions in the neighbourhood of the optical axis.[4] Depending very much on the class of system, this occurs at relatively low apertures and field angles. The problem as to how the reliability of the predictions may be predetermined receives little attention in the literature although Ford[4] investigated this in relation to the usefulness of aberration coefficients.

It is clear that it is impossible to predict the aberrations of rays in the outer regions of the field of systems of even quite moderate maximum field angle if arbitrary rays are always referred to the optical axis. This may be overcome by introducing the concept of the base-ray: the base-ray is defined to be a ray which has been accurately traced through the system and to which arbitrary rays may be referred in a suitable manner. The base-ray may, in principle, be chosen arbitrarily. Granted that the aberration functions can be expanded about the base-ray (i.e., in terms of the variables specifying arbitrary rays

---

[3] H. A. Buchdahl: "Optical Aberration Coefficients". O.U.P. 1954. This monograph will hereafter be referred to by the letter M. Equations in M will be referenced by placing, in parenthesis, the letter M followed by the equation number as it appears in M.

[4] See Ford (1962); Buchdahl (1959) §3, (1960) §7.

with respect to the base-ray) the resulting aberration coefficients can be used to predict the aberrations of rays which lie in a neighbourhood of the base-ray. In this manner regions of the field away from the axis of the system can be satisfactorily covered. Naturally the new series for the aberration functions will give reliable predictions only within a certain neighbourhood of the base-ray. It is expected that a study of the reasons for the breakdown of the new series will give additional information relating to the analogous problem in the neighbourhood of the axis, and may even make it possible to give more definite rules of thumb relating to the convergence of the series for the aberration functions.

(b) A theory whereby the exact aberration coefficients defined with respect to an arbitrary base-ray may be calculated will be called a <u>basal theory</u>; in particular, the theory presented in this thesis will be called <u>the basal theory</u>. It is apposite to mention here the work of Smith, Wynne and Weinstein. Smith[2] has investigated first order imagery for a general asymmetric system using the angle characteristic for a single surface as his starting point. He derives the laws for combining the second order parts of the characteristic functions of two systems and thus constructs the characteristic function of the complete system, correct to the second order (i.e., <u>first</u> order as regards displacement). In particular he illustrates his results by a discussion of spectacle lenses and the human eye. As a continuation of his earlier work, Smith[5] considers first order (parabasal) imagery in the neighbourhood of a skew ray traversing an asymmetrical system. Again using the angle characteristic of a single refracting surface, he derives the formulae representing

---

5 Smith (1930).

the coordinates specifying the ray after refraction in terms of those before refraction at a single surface. With suitable formulae for transferring from one surface to the next, rays may be traced through the system in much the same way as parabasal rays may be traced using equations (13.9,10) of this thesis. The coordinates of the ray in the image space may be expressed linearly in terms of those in the object space using sixteen coefficients, of which ten only are independent. Wynne[2] has derived from the wave front aberrations the third order ray aberrations and primary chromatic ray aberrations of an anamorphotic system of cylindrical lenses. It is to be noted that the aberrations of even order vanish identically for such systems. Weinstein[6] has derived expressions for the wave aberration of a pencil of rays about a principal ray up to the fourth order in the aperture (i.e., the third order for ray intersection aberrations). His base-ray (a principal ray) is accurately traced, and meridional and sagittal fans in the first order neighbourhood of the principal ray are "traced" to determine the data required to compute the aberration coefficients. The theory presented in this thesis goes beyond this since pencils inclined to the base-ray are also covered. The systematic development of the theory makes it possible to calculate coefficients of order exceeding the third.

Although it is anticipated that, as it stands, the basal theory presented in this thesis will be applied principally to an investigation of the performance of wide angle symmetric systems in the outer regions of the field, the theory is considerably more general. For instance, the aberration functions arise essentially from refractions at each

---

[6] Weinstein (1949, 1950).

surface and transfer from one surface to the next. In considering both the refraction and the transfer, the explicit assumption that the system is symmetric is not necessary. The symmetric nature of the system is hidden, so to speak, in the trace of the base-ray and is quite irrelevant to the development of the theory. Consequently, if the surfaces are assumed to be spherical, the theory will apply to systems of spherical surfaces suffering from arbitrarily large decentrings in a given plane. Such a system is scarcely of more than academic interest, and of more practical interest is a general system of surfaces which are not necessarily rotationally symmetric. Although the basal theory developed here assumes the surfaces to be spherical, the general principles hold for aspheric surfaces and the generalization of the theory to such systems is indicated in Appendix D. In particular, the first order coefficients of such a system are obtained. Granted this generalization, the theory could be applied to the outer regions of the field of a wide angle symmetric system incorporating aspherical surfaces so as to extend the region of the field in which reliable predictions of the aberrations of finite pencils of rays can be made.

In order to gain insight into the most general case, only a plane symmetric system with a base-ray lying in the plane of symmetry is considered. A plane of symmetry of any system is called a <u>meridional plane</u>. The theory is thus a basal theory with a meridional base-ray. The base-ray could be any ray in the meridional plane of an arbitrary plane symmetric system or a ray in a meridional plane of a symmetric system. It is further assumed that the surfaces are spherical. Although a system of spherical surfaces with their centres restricted to a plane but suffering from large decentrings is, as has already been remarked, largely academic, the theory is developed as if the system were not (axially) symmetric. The artificiality of the resulting system is irrelevant in view of the fact that it is only a matter of detail and not of principle to generalize to arbitrary surfaces.

Weakly decentred systems have received considerable attention in the literature. For example, to mention but a few, Conrady, Maréchal, Stephan and Kiuti[7] express the aberration due to decentring as series in the decentring parameters. For weak decentring the second and higher powers of these parameters may be neglected. Epstein[7] considers the aberration due to decentring by tracing rays through the centred portions of a system, the decentring effects arising from transfer from one centred portion to the next.

(c) The sign conventions employed throughout this thesis are those of coordinate geometry and <u>not</u> the traditional conventions of optics. In references to the axial theory it is assumed that these conventions are in fact used in M. Thus, in the notation of M, the direction cosines of a ray are (α,β,γ) and $\underline{\underline{e}}$ has the components α,β and γ (cf. M§2). V and W, the corresponding direction tangents, are β/α , γ/α. Further the coordinates of the object point O are now (ℓ_0, H_{y1}, H_{z1}) (cf. M§1). In order to convert the equations of M to these sign conventions it is sufficient to note that the changes arise only from the changes in the signs of $\underline{V}$ and $\underline{H}$. All other changes can, with a little care, be deduced from these. In addition anti-clockwise turning is reckoned positive in

---

7 Conrady (1919), Maréchal (1949), Epstein (1949), Stephan (1950), Kiuti (1951).

the measurement of angles, and the angles of incidence and refraction[8] are measured <u>from</u> the normal to the surface (at the point of incidence) <u>to</u> the ray.

Rays are specified with respect to the base-ray by their <u>canonical variables</u> $\underset{\sim}{Y} = (Y,Z)$, $\underset{\sim}{V} = (V,W)$ such that the first pair define a point on the ray and the last pair give the direction of the ray (§5). $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are, of course, quasi-linear (§25 and M§9b). In addition, rays are specified in the object space by their <u>paracanonical coordinates</u> $\underset{\sim}{S}$ and $\underset{\sim}{T}$, that is, some combination of the canonical variables at the first surface, in particular by their SPC (see §6 and M§12,13). For the base-ray all four of $\underset{\sim}{Y}$, $\underset{\sim}{V}$, $\underset{\sim}{S}$ and $\underset{\sim}{T}$ are zero. In order to take the maximum advantage of any simplifications inherent in the plane symmetry of the system, the variables and coordinates are chosen so that under a reflection in the plane of symmetry, Y, V, S_y and T_y remain unchanged whereas Z, W, S_z and T_z change in sign.

The aberration of a ray may be defined in several ways (§7b). Imagery is not ideal in the first order (parabasal) neighbourhood of the base-ray and thus the definition of the aberration which one naturally employs in the axial theory requires modification. Two possible choices are as follows: first, having assigned some suitable image plane Ⓗ, the "ideal image" may be defined to be the point in Ⓗ specified by $\underset{\sim}{h}_k' = (m_{yk}'H_{y1}, m_{zk}'H_{z1})$ and the aberration defined to be $\underset{\sim}{\epsilon}_k' = \underset{\sim}{H}_k' - \underset{\sim}{h}_k'$. (Note, the magnifications m_{yk}' and m_{zk}' may be distinct.) Alternatively - and this is especially useful for extreme wide angle systems - define $\underset{\sim}{\epsilon}_k' = \underset{\sim}{H}_k' - \underset{\sim}{H}_{Pk}'$ where $\underset{\sim}{H}_{Pk}'$ is the image height in Ⓗ of the principal

---

8 Although refraction only is explicitly referred to, reflection can be treated by the usual artifice of setting N' equal to -N.

ray from the object. (A principal ray is *any* ray through the centre of the diaphragm. See §62.) It is simplest to regard $\underset{\sim}{H}_k{}'$ as the fundamental quantity rather than $\underset{\sim}{\epsilon}_k{}'$ and to determine the coefficients of the Taylor series for $\underset{\sim}{H}_k{}'$ in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ - that is, the coefficients in the *paracanonical expansion* (§24a) of $\underset{\sim}{H}_k{}'$ - to which those of $\underset{\sim}{\epsilon}_k{}'$ are simply related, irrespective of the definition of $\underset{\sim}{\epsilon}_k{}'$.

The theory is of course based on the laws of refraction from which the behaviour of the canonical variables under refraction may be derived (§10). The simplicity of the resulting equations naturally depends on the choice of canonical variables and has a strong bearing on the relative simplicity of the algebra of the theory but not on the general principles involved. As a first approximation to the behaviour of the rays traversing the system the dominant terms of the exact equations are considered. These give rise to *parabasal optics* (Chapter III), the analogue of paraxial optics. In the literature the parabasal region is usually defined to be the neighbourhood of the base-ray such that second and higher powers of the canonical variables may be neglected. However, unless considerable caution is exercised in the application of this definition, inconsistent results are obtained (§18). Essentially, the second and higher order terms *must not be neglected* as soon as they appear; rather, the exact equations describing the required result must be obtained to a *higher order* before any approximations are made. As examples of correct results obtained in this manner are (i) the sagittal focal line is *not* in general normal to the base-ray, and will be so if and only if a certain *second order* coefficient vanishes (§21), and (ii) a straight line in the meridional plane and normal to the base-ray is imaged by meridional rays into a straight line which is again not necessarily normal to the base-ray (§20). It is usually stated[9] that both these lines are normal to the base-ray.

Two invariants $\underset{\sim}{g}$ may be defined for arbitrary pairs of parabasal rays (§16). However, if one of these rays is not parabasal, $\underset{\sim}{g}$, now denoted by $\underset{\sim}{G}$, is in fact not an invariant. Since $\underset{\sim}{G}$ reduces to an invariant in the parabasal region, it is said to be a quasi-invariant. If $\underset{\sim}{G}$ happens to be an invariant in the case of a particular system, then parabasal optics accurately describes the imagery of the system. However, given the parabasal equations and the quasi-invariant $\underset{\sim}{G}$ it is possible to obtain the exact equations describing the passage of rays through the system (i.e., the exact equations for $\underset{\sim}{Y}$ and $\underset{\sim}{V}$) simply by replacing $\underset{\sim}{S}$ and $\underset{\sim}{T}$ wherever they occur by $\underset{\sim}{S} + \underset{\sim}{\delta}_y$ and $\underset{\sim}{T} + \underset{\sim}{\delta}_v$, where $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ are expressed simply in terms of the surface or refraction increments $\Delta\underset{\sim}{G}$ to $\underset{\sim}{G}$ (§25). The essential problem is thus the determination of $\underset{\sim}{G}$. Moreover, the augmented image height (§22 and §7) can be expressed as the sum over surfaces of the refraction increments $\Delta\underset{\sim}{G}$ and an out-of-focus term which is almost invariably small (§23,24). $\underset{\sim}{H}$ may therefore be expressed as a sum of contributions by the individual surfaces. In particular the paracanonical coefficients of $\underset{\sim}{H}$ (the coefficients of its Taylor series in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$) are expressed as a sum of surface contributions. This feature is highly desirable since it is just a knowledge of the contributions from the individual surface to the final aberration which makes the axial theory so useful.

The surface increments to the quasi-invariants are expressible in terms of the canonical variables before refraction at the surface in

---

[9] Synge (1937) §9; Luneburg (1964) §35; Born and Wolf (1964) §46.

question through two auxiliary quantities R and X (§26). The expansion of $\Delta\underset{\sim}{G}$ proceeds in three stages. First, the surface expansions (i.e., the expansion in terms of the canonical variables before refraction at a surface) of various quantities, in this case R and P(= RX), are formed (§30-34). This essentially involves solving quadratic equations for R and X as series in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. Second, as a purely formal step $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ are (at the surface in question) initially taken as zero. This makes $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ linear functions of $\underset{\sim}{S}$ and $\underset{\sim}{T}$, and the surface expansions of both R and P are transformed into expansions in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ by substituting these linear functions in place of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ (§36). The resulting expansion is in general not exact and is termed the pseudo-expansion. Finally, to obtain the exact, or paracanonical, expansion of R and P it is sufficient to replace $\underset{\sim}{S}$ and $\underset{\sim}{T}$ in the pseudo-expansion by $\underset{\sim}{S} + \underset{\sim}{\delta}_y$ and $\underset{\sim}{T} + \underset{\sim}{\delta}_v$ respectively (§37). The increments $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ are dependent upon the $\Delta\underset{\sim}{G}$ at preceding surfaces. Thus the paracanonical coefficients for R and P are also dependent upon the $\underset{\sim}{g}$-coefficients (the paracanonical coefficients of $\Delta\underset{\sim}{G}$) at the preceding surfaces. These $\underset{\sim}{g}$-coefficients are also necessary in order to express the $\underset{\sim}{g}$-coefficients at the surface under consideration in terms of the r- and p-coefficients. It is evident therefore that in order to determine the $\underset{\sim}{g}$-coefficients at the $\underset{\sim}{j}$th surface it is necessary to know the corresponding coefficients at the preceding $\underset{\sim}{j}$-1 surfaces. However, at the first surface the $\underset{\sim}{g}$-coefficients are trivially zero. Hence the process has a starting point and the $\underset{\sim}{g}$-coefficients may be determined in an iterative fashion by proceeding from surface to surface through the system (see also §28,35). Moreover, the $\underset{\sim}{g}$-coefficients up to the ($\underset{\sim}{n}$-1)th order at the first $\underset{\sim}{j}$-1 surfaces are required in order to compute the $\underset{\sim}{n}$th order $\underset{\sim}{g}$-coefficients at the $\underset{\sim}{j}$th surface. Hence the latter could in fact be obtained iteratively by proceeding order by order, the first order $\underset{\sim}{g}$-coefficients being zero since $\underset{\sim}{G}$ is a quasi-invariant.

Because of the symmetry of the system, H_{yk}' and H_{zk}' are respectively even and odd in the pair of variables S_z and T_z. It follows that a typical term in the expansion of either is of the form coeff $\times\ S_y^{\,n-\mu} T_y^{\,\mu-\nu} S_z^{\,\nu-\tau} T_z^{\,\tau}$, where, in view of the symmetry properties of $\underset{\sim}{H}_k'$, ν is even in H_{yk}' and odd in H_{zk}'. The <u>order</u> of the term is the value of <u>n</u>. It is evident that there are two first order, six second order and ten third order terms in H_{yk}' and two first order, four second order and ten third order terms in H_{zk}' (§24a). In general there will be a similar number of terms in the aberration functions. This number is to be compared with the two first order and the six third order terms of the image height in the axial theory. The great variety of coefficients can be conveniently classified by generalizing the classifications of Steward and Buchdahl[10] to cover terms of all orders (§49). The terms "primary", "secondary" and "tertiary" are used in the context of the axial theory only, and as usual are synonomous with the third, fifth and seventh orders respectively.

Just as in the axial theory, certain identities exist between the aberration coefficients, more explicitly between the $\underset{\sim}{G}$-coefficients. First, the <u>extremal identities</u>, are obtained along the lines of Buchdahl[11] from the existence of an extremal condition (Fermat's principle), associated with the point characteristic, say (Chapter VII). Second, if the system is in fact rotationally symmetric additional identities must exist since the imposition of rotational symmetry reduces the number of degrees of freedom of the system and this is naturally reflected

---

10 Steward (1958) §16; Buchdahl (1958) §7.

11 Buchdahl (1965) §9.

in the number of independent coefficients associated with imagery of any order. In this case the additional identities, the rotation identities (§57-59), are obtained from a generalization of the invariant E_x^* of the axial theory (M§4b) and take a quite remarkable form (Table 59/2). If the characteristic function of a symmetric system is determined with respect to an arbitrary base-ray, it will be formally indistinguishable from the characteristic function of an asymmetrical system. However, in view of the rotational symmetry, certain identities between the coefficients of the characteristic function must be satisfied. These identities are exactly equivalent to the rotation identities and can be determined from a differential equation which the characteristic function (qua function of the basal coordinates) must satisfy when it is that of a symmetric system (see §57c).

(d) It is often stated in the literature[12] that it is impossible to predetermine the entrance pupil for pencils of rays traversing a symmetric system without tracing a large number of rays through the system and examining each for vignetting. However, this is not the case (even for plane symmetric systems) once the $\underset{\sim}{G}$-coefficients are known. Given the $\underset{\sim}{G}$-coefficients, the coordinates of $\underset{\sim}{Y}_d$ (the diaphragm coordinates) of the point of intersection of any ray with the diaphragm may be expressed as series in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ (§63). It can be shown (§61) that these series may be inverted. Thus, expressing $\underset{\sim}{S}$ in terms of $\underset{\sim}{Y}_d$ and $\underset{\sim}{T}$, it is possible to determine (i) the boundary and area of the entrance pupil (§64), and (ii) the coordinates $\underset{\sim}{S}$ of the principal ray of the pencil (§65). It

---

12 Herzberger (1958) p.105; Stavroudis and Sutton (1965) p.2.

is to be noted that these are <u>not</u> trial and error methods (cf. Gardner[13]); the required quantities are given directly once the $\underset{\sim}{G}$-coefficients are known and no ray tracing (apart from the base-ray) is required. With $\underset{\sim}{S}$ now known for principal rays as a function $\underset{\sim}{T}$ it is possible to express the image height of principal rays as a series in $\underset{\sim}{T}$, the coefficients of which are all of the distortion type (§66). The aberration may be referred to the principal ray by subtracting these coefficients from the corresponding coefficients in $\underset{\sim}{H}$. Alternatively - and in retrospect is is felt that this offers the most practical advantages - the members of a pencil of rays could be specified by $\underset{\sim}{T}$ and $\bar{\underset{\sim}{S}}$ where $\bar{\underset{\sim}{S}} = \underset{\sim}{S} - \underset{\sim}{S}_p$, and $\underset{\sim}{S}_p$ is the value of $\underset{\sim}{S}$ for the principal ray of the pencil. Then, rays are referred to the centre of the entrance pupil associated with the given pencil, and $\underset{\sim}{H}$ may be expressed in terms of $\bar{\underset{\sim}{S}}$ and $\underset{\sim}{T}$, in which case the image height of the principal ray is given by the distortion terms alone. The aberration is thus referred to the principal ray by neglecting the distortion terms. Similar results, differing in detail, may be derived in the context of the axial theory. However, it is expected that results such as those above are likely to be of most use in the analysis of wide angle systems, that is, in the context of the basal theory. The illumination of the image is given when the image height of the principal ray and the area of the entrance pupil are known as functions of field angle[14]. Consequently, the illumination may be estimated from the basal coefficients.

(e) It is evident that the principles of the theory are very similar to those of the axial theory. Differences in detail arise, however, from

---

13 Gardner (1947) part VII.

14 For example, see Reiss (1948) equation 17.

several sources. First, the system is not rotationally symmetric and consequently the number of coefficients of any order is greatly increased, and many of the equations lose their symmetrical appearances: y- and z-components, in fact, require separate treatment. Second, the base-ray does not pass through the system undeviated (compare a ray along the axis of a symmetric system). Consequently there are many quantities associated with the base-ray, the so-called pseudo-parameters (§2,11), which do not take simple values like zero or unity. Because of this proliferation of parameters, the task of finding simple forms - such as the Seidel form of the axial theory - for the intrinsic contributions (pseudo-coefficients) of even the second order is one of extreme complexity, and it is doubtful whether the final equations would be of any advantage over those presented here. All manner of coefficients have to be left in turn as functions of other coefficients until some quantities can be expressed simply in terms of the system parameters (e.g., the $\hat{p}$- and $\hat{r}$-coefficients, §32,33). In view of this, the task of computing the lowest order coefficients of $\Delta \underset{\sim}{G}$ in the basal theory (i.e., the second) is considerably heavier than the corresponding task in the axial theory, that is, the computation of the primary (third order) coefficients. Third, the presence of the out-of-focus terms (§22) in the expression for $\underset{\sim}{H}$ in terms of the quasi-invariants implies that both the $\underline{a}$- and $\underline{b}$-components of the $\underline{n}$th order $\underset{\sim}{G}$-coefficients are required in order to determine the exact value of the $\underline{n}$th order $\underset{\sim}{h}$-coefficients. The out-of-focus term arises from the non-coincidence of the meridional and sagittal foci of the base-ray. It is to be noted that the terms of order higher than the first must be included in the out-of-focus term otherwise the resulting $\underset{\sim}{h}$-coefficients would not be exact. Compare this with the usual method of treating small displacements of the image plane, where only the first order terms of $\underset{\sim}{V}$ are included (M§36).

Although the general principles involved in the basal theory are very similar to those of M (by design, the basal theory closely parallels the axial theory), all the results required are derived anew in a more general manner. No knowledge of the contents of M is required, although such knowledge would naturally be a great aid to the full comprehension of the theory. The mathematics employed in the derivation of the equations leading up to the computation of the coefficients is elementary, consisting essentially of manipulations of series and identities in four variables.

(f) It has already been mentioned that the simplicity of the derivations above depends on the definitions of the canonical variables. So do the practical aspects of the theory. Perhaps the most important practical aspect of the axial theory is that the contributions by the various surfaces to the final aberration coefficients are known. This is because the final aberration is expressed simply in terms of the quasi-invariants. If, in either the axial or the basal theories, the image plane is not normal to the x-axis of the coordinate systems associated with the surfaces, the aberration is not expressed simply in terms of the quasi-invariants (consider the analysis of §22 based on (7.4)) and it is, in fact, doubtful if surface increments could be <u>usefully</u> defined in this case. Second, suppose that the transfer increments to the quasi-invariants were not identically zero (for instance, if rotation of coordinates were associated with transfer), then the final aberration coefficients would consist of surface contributions as well as the essentially redundant transfer increments.

At this point it is appropriate to consider some of the possible choices of coordinate systems associated with each surface and by means of which rays are referred to the base-ray. For symmetry reasons the z-axis is always to be taken normal to the plane of symmetry. First there are the so-called hamiltonian coordinates (§84) in which the x-axis is along the base-ray and two coordinate systems are associated with each surface - one in the object space and one in the image space. It is evident that transfer from the image space of one surface to the object space of the next is simply a translation of coordinates. Hence there are no transfer increments to the g-coefficients. However, the image plane is not in general normal to the base-ray and the image height is therefore not linearly dependent upon the quasi-invariants and consequently the surface contributions to the aberration coefficients cannot be obtained. Moreover, the association of two coordinate systems with each surface introduces considerable complications to the algebra of the theory, especially to the quadratic equation for R and its subsequent solution. As a second possibility normal coordinates (§85) may be considered. Here the x-axis of the coordinate system associated with the surface is along the normal to the surface at the point of intersection of the base-ray. There is only a single coordinate system associated with each surface and transfer from one surface to the next requires a rotation of coordinates. Hence there are transfer increments to the g-coefficients. Again, the image plane is not necessarily normal to the x-axis, that is, parallel to the tangent plane at the point of incidence of the base-ray with the last surface, and the surface increments are unknown.

It is evident that if at all possible, the coordinate systems must satisfy the following requirements: (i) the coordinate axes of each coordinate system must be respectively parallel and (ii) the x-axis must be normal to the chosen image plane. These are satisfied by the translated coordinates employed in the thesis. It is to be noted that the object plane is not necessarily parallel to the image plane. However, this is a matter of detail rather than principle since the use of non-linear paracanonical coordinates[15] may be invoked.

(g) Throughout the text considerable emphasis is at times placed on matters related to conventions of notation. It is important in work such as this that the notation employed should be as consistent and systematic as possible. The advantages of this far outweigh the disadvantages of having a complex symbolism, especially in as much as the rules governing the use of the various affixes are clearly set out (e.g., Table 29/1). At first sight the mass of affixes attached to various kinds of coefficients is rather bewildering - there are in general two more subscripts attached to the kernel symbols of the coefficients of the basal theory than to the corresponding symbols of the axial theory. However, affixes are uniquely characterized by their "rank" and a set of formal rules for their interpretation is given (§29). Appendix H lists and identifies all the symbols which occur in the text. The principles governing the notation of M have been adopted, although certain changes have been made in the interest of consistency. Script letters (here denoted by a ◯ around the plain type equivalent) exclusively refer to geometrical concepts associated with the system and rays traversing the system. Many quantities denoted by German or Greek type in M are now represented by Roman type.

---

15 See footnote 1, §7. Note that the non-linear nature of the coordinates refers to the expressions for $\underset{\sim}{Y}_1$, $\underset{\sim}{V}_1$ in terms of the paracanonical coordinates $\underset{\sim}{S}$, $\underset{\sim}{T}$ and that in the parabasal region $\underset{\sim}{Y}_1$, $\underset{\sim}{V}_1$ are as usual linear in $\underset{\sim}{S}$, $\underset{\sim}{T}$.

Examples of the need for, and benefits of, a consistent and systematic notation are as follows: first, although the basic equations occur in asymmetric pairs it is possible through the introduction of a simple notation and associated conventions to combine two, and at times four, such equations into a single equation and greatly compress others without obscuring their basic character. Second, the need to distinguish clearly between the surface-, pseudo-, and paracanonical expansions and the many coefficients associated with these should be obvious. Third, in order to distinguish between basal quantities and the corresponding axial quantity, a subscript A is placed on the latter. If an axial quantity pertains to the base-ray, a subscript B is attached. With very rare exceptions each symbol or combination of symbols and affixes has a unique meaning. Attention has been paid to the generalisation of the notation to the case of a non-meridional base-ray and it is believed that no major modifications would be required, nor would any inconsistencies develop when such a generalisation is made.

(h) In order to test the quality of the predictions of aberrations using the basal coefficients numerical work was carried out on a wide angle system of full field 200° designed by Havlicék.[16] The use of an extreme wide angle system ensured that a wide choice of base-rays was available in regions of the field where the axial theory could not possibly produce a reasonable representation of the aberrations. (The tertiary predictions break down between 30° and 35° half-field for this system. See fig.2.) Moreover, a wide range of dominant aberration-types presents itself in this case - from large spherical aberration near the axis to large coma and curvature of field in the outer regions

---

16 See §69 and Havlicék (1951).

of the field (e.g., see §74) - and the aberrations associated with the pupil planes are large, so that the results of Chapter X may be used to advantage. Since the aberration is referred to the principal ray two features are required of the coefficients: (i) that they correctly reproduce the coordinates and image heights of principal rays, and (ii) that the aberrations referred to the principal ray (which represent the blur of the image), should be accurately predicted from the coefficients. The last of these requirements is probably the most important.

Base-rays were chosen to be principal rays corresponding to half-fields of 47.7°, 64.1° and 78.6° and the basal coefficients with respect to each of these rays were determined on an IBM 360/50 computer. The third order predictions of the coordinates of principal rays were tested in two ways (§75). First, the predicted coordinates were compared graphically with those obtained by tracing rays back from the centre of the diaphragm. Second, the rays with the predicted coordinates were traced so as to determine their diaphragm coordinates which were expressed as a fraction of the radius of the diaphragm. When this was a fraction of a percent the predicted coordinates were regarded as satisfactorily representing principal rays. This was the case for each base-ray over a range of field angles somewhat larger than the range over which satisfactory third order predictions of the aberrations were obtained.

The predictions of the aberrations referred to the principal ray were tested by comparing graphically the predicted and corresponding traced aberrations. For each base-ray meridional and sagittal fans, spot diagrams and aperture curves were constructed and compared. An aperture curve is the curve of light obtained in the image plane by replacing the circular diaphragm by an annular one whose radius corresponds to the required aperture. Normally the construction of aperture curves would involve tracing a large number of rays and determining by trial and error which of these pass through the required annular aperture. However, the basal coefficients enable the coordinates of rays passing through an annular aperture of radius ρ to be expressed simply in terms of ρ (§78).

An analysis of the agreement between corresponding predicted and traced aberrations reveals that the predictions break down badly once S_y or S_z exceed a value corresponding to the full aperture of the pencil about the base-ray. As the field angle of the pencil increases the entrance pupil moves away from the axis, and this movement is by no means negligible. Hence, for pencils inclined to a base-ray, the point $\underset{\sim}{S} = 0$ does not correspond to the centre of the entrance pupil. For pencils such that $T_y > 0$ (< 0) the maximum (minimum) value of S_y exceeds in magnitude the corresponding value for the pencil about the base-ray. It is these large values which correspond to the rays whose aberrations are badly predicted. Even for field angles such that the prediction of the flare of the spot diagram is very poor the prediction of the core is quite good (see fig.15). It is possible that by expressing $\underset{\sim}{H}$ in terms of $\underset{\sim}{T}$ and $\bar{\underset{\sim}{S}}$ the quality of the predictions might be improved. Although vignetting was determined with sufficient accuracy by the first and second order terms only (§77), it was absolutely necessary that the third order terms be included in the aberration. It appears as though the predictions of the aberrations of pencils inclined to the base-ray may be considerably improved by inclusion of the fourth order terms and, if so, that the main cause of improvement would probably be due to the coefficients of fourth order third degree astigmatism. Other terms, including fifth order fifth degree astigmatism and fifth order first degree coma, will act as small corrections.

(i) If the system is symmetric the basal coefficients can formally be expressed explicitly in terms of the axial coefficients and coordinates $\underset{\sim}{S}_B$ and $\underset{\sim}{T}_B$ of the base-ray. Replace the axial coordinates $\underset{\sim}{S}_A$ and $\underset{\sim}{T}_A$ in the series for the aberration function by $\underset{\sim}{S} + \underset{\sim}{S}_B$ and $\underset{\sim}{T} + \underset{\sim}{T}_B$ and expand with respect to $\underset{\sim}{S}$ and $\underset{\sim}{T}$. The basal coefficients then appear as <u>infinite</u> series in $\underset{\sim}{S}_B$ and $\underset{\sim}{T}_B$, the coefficients of which are simple combinations of the axial aberration coefficients. By truncating these series at, say, the tertiary terms and comparing the basal coefficients determined from these truncated series with those determined exactly, it should be possible to extract information relating to the radius of convergence of the axial series for the aberration.

Neither chromatic aberrations nor the derivatives of the basal coefficients with respect to the system parameters are considered at all. The main problem to be overcome here is that if either the colour of the light or a parameter of the system is changed, the base-ray is altered and hence the numerical values of the pseudo-parameters will vary. By contrast, for a symmetric system the changes of the parameters are always such that the system remains symmetric. The base-ray - the ray along the optical axis - is therefore unaltered.

II. TRANSLATED COORDINATES

1. Specification of Plane Symmetric Systems

The optical system Ⓚ with which this thesis is concerned is an arbitrary member of a class of optical systems, each member of which possesses a plane of symmetry Ⓜ and is constructed from homogeneous isotropic media of refractive indices N separated by k spherical boundaries or surfaces Ⓕ of curvature c. The jth surface of Ⓚ encountered by an arbitrary ray from some object is called the jth surface and denoted by $Ⓕ_j$. Immediately prior to reflection or refraction at the jth surface rays are travelling in the object space of $Ⓕ_j$, and immediately after reflection or refraction at $Ⓕ_j$ they are in the image space of $Ⓕ_j$. The refractive index of the medium in the object space of $Ⓕ_j$ is denoted by N_j. Thus, if a prime attached to a quantity Q associated with a particular surface denotes the value of Q in the image space of that surface,

$$N_j' = N_{j+1} \quad .$$

In general, if Q is associated with the jth surface, Q is denoted by Q_j. Subscripts j and k are exclusively reserved for indicating particular surfaces of Ⓚ, with k referring to the last surface. If Q has additional subscripts attached to the right of the kernel symbol Q, the surface indicator is at the extreme right of these (see §29).

In order to prescribe Ⓚ uniquely, the locations of the centres C_j of the $Ⓕ_j$ must be specified with respect to C_1. To do this it is convenient to introduce in Ⓜ an arbitrary straight line Ⓐ, called the pseudo-axis of Ⓚ, and a corresponding coordinate system $Ⓒ_0$ whose origin is on Ⓐ, whose x-axis lies along Ⓐ and whose y-axis is in Ⓜ. Accordingly, the coordinates of C_j in $Ⓒ_0$ are $(x_{Cj}, y_{Cj}, 0)$ and the structure of Ⓚ is now given by the 4$\underline{k}$-1 numbers c_j, N_j j=1,2,...,$\underline{k}$, N_k' and $x_{Cj}-x_{C1}$, $y_{Cj}-y_{C1}$ j=2,3,...,$\underline{k}$. These are called the (proper) parameters of Ⓚ and uniquely determine its optical properties. By varying any one of these parameters the optical properties of Ⓚ are altered, thus Ⓚ will have 4$\underline{k}$-1 degrees of freedom.

Either for the purpose of controlling the speed of the lens, or for controlling undesirable aberrations by limiting the aperture of pencils of rays traversing the system, a diaphragm Ⓓ will be associated with the system. It will be assumed that the diaphragm is circular and that there are no foci of the object in or near the plane of Ⓓ. Neither this last assumption nor the existence of a diaphragm are essential to the derivation of the equations leading to the computation of the basal aberration coefficients but are necessary for the analysis of Chapter 10.

2. The Base-Ray and Pseudo-Parameters

To discuss the optical behaviour of Ⓚ quantitatively it is convenient to introduce the concept of the base-ray. The base-ray $Ⓡ_B$ of Ⓚ is any ray, lying in the plane of symmetry of Ⓚ, which has been accurately traced[1] through Ⓚ. Arbitrary rays are referred to $Ⓡ_B$ (see §5a) and consequently the aberration coefficients obtained in this thesis make it possible for the performance of Ⓚ to be studied for arbitrary rays in the neighbourhood of $Ⓡ_B$. It is evident that

---

[1] See Appendix C for a discussion on tracing rays through arbitrary plane symmetric systems.

if the base-ray is changed[2] Ⓚ will, as far as the basal theory is concerned, behave as a different system. Consequently suitable specifications for $Ⓡ_B$ are included in the parameters of Ⓚ. This requires two additional parameters and Ⓚ has, formally, 4k+1 degrees of freedom. In the development of the basal theory it will be assumed that certain quantities referring to the base-ray are accurately known from the trace of the base-ray - for example, the angles of incidence and refraction I_{Bj}, I_{Bj}', and the points of incidence $Ⓟ_{Bj}$ of $Ⓡ_B$ with $Ⓕ_j$. Such quantities are referred to as pseudo-parameters and distinguished by the appearance of a subscript B to the right of, but adjacent to, the kernel symbol (see §29). Evidently the pseudo-parameters are functions of the 4k+1 parameters of Ⓚ and hence do not represent additional degrees of freedom. It proves convenient to use simultaneously pseudo-parameters which are simply related to one another (see §11).

3. Meridional, Sagittal and Basal Rays

The plane of symmetry Ⓜ of Ⓚ is called the meridional plane of Ⓚ and any ray lying in Ⓜ is said to be a meridional ray. Due to symmetry reasons, "meridional" is a property of a ray throughout Ⓚ - any ray lying in Ⓜ in one of the media constituting Ⓚ will lie in Ⓜ in all other media. $Ⓡ_B$, itself a meridional ray, is broken into linear segments $Ⓟ_{Bj}Ⓟ_{Bj+1}$ by the surfaces $Ⓕ_j$. In any of the k+1 distinct optical media comprising Ⓚ, a sagittal plane may be defined to be the plane

---

2 It is to be emphasized that the base-ray is itself arbitrary. The only reason why it was restricted to Ⓜ was to take account of any simplifications due to the plane symmetry of Ⓚ.

normal to Ⓜ but containing the segment of Ⓡ$_B$ in that medium. In particular, the sagittal planes in the object and image spaces of Ⓕ$_j$ are Ⓢ$_j$ and Ⓢ$_j'$ respectively. If a ray lies in the sagittal plane in a particular medium it is said to be a *sagittal ray* in that medium. "Sagittal" is not a property of a ray as a whole - if any ray is sagittal in Ⓢ$_j$ it is not necessarily sagittal in Ⓢ$_j'$ (see also §17). Any point lying on the base-ray is said to be a *basal point* and rays passing through a basal point are *basal rays*. Evidently rays in Ⓜ or in any sagittal plane are basal rays. In this case the basal point through which the rays pass may be *virtual* in the sense of a virtual object or image.

A *meridional fan* is a collection of meridional rays from a point object (*meridional object*) in Ⓜ. Similarly, a *sagittal fan* is a collection of sagittal rays (in the object space of Ⓚ as a whole) from a sagittal point object in Ⓢ$_1$. (e.g., see §17,18,20 and 21.) It is sometimes convenient to generalize the concept of sagittal to refer to a plane normal to Ⓜ and containing a ray of particular importance, for example, a principal ray (see §76a).

4. Some Conventions Relating to Notation

Suppose $\vec{Q}$ is a vector in some cartesian coordinate system, for example, the normal $\vec{n}$ to the surface Ⓕ. $\vec{Q}$ represents the three numbers (Q_x, Q_y, Q_z). Frequently it is sufficient to consider only the y- and z-components of $\vec{Q}$, in which case it is convenient to write

$$\underset{\sim}{Q} = (Q_y,\ Q_z) \quad . \tag{4.1}$$

The above notation is extended to cover two quantities q_y and q_z which are not necessarily the components of a physically meaningful vector but none the less are unequivocally associated with the y- and z-components, respectively, of such a vector. Thus

$$\underset{\sim}{q} = (q_y, q_z) \quad .$$

Quantities such as $\underset{\sim}{Q}$ and $\underset{\sim}{q}$ are called <u>two-vectors</u> in distinction to <u>vectors</u> such as $\vec{n}$ and <u>scalars</u> such as R (defined in §8). If $\underset{\sim}{P}$, $\underset{\sim}{Q}$ and $\underset{\sim}{R}$ are three two-vectors and q a scalar, the equations

$$\underset{\sim}{Q} = \underset{\sim}{P}\underset{\sim}{R} \quad , \quad q = \underset{\sim}{P}\cdot\underset{\sim}{R} \quad , \tag{4.2}$$

are respectively equivalent to the pair of equations

$$Q_y = P_y R_y \quad , \quad Q_z = P_z R_z \quad , \tag{4.3}$$

on the one hand, and

$$q = P_y R_y + Q_z P_z \tag{4.4}$$

on the other. It is felt that no confusion need arise with the notations of (4.2) since it should be obvious which product is being used from the nature of other terms occurring in the equations in which such a product appears. At any rate the scalar product is written as a normal scalar product between two-vectors. These notations frequently enable pairs of equations to be combined into a single equation and other equations to be written succintly.

5. <u>Translated Coordinates and Canonical Variables</u>

(a) At each surface set up a coordinate system $\textcircled{C}_j$ obtained by translating $\textcircled{C}_0$ to a new origin at the point of incidence $\textcircled{P}_{Bj}$ of $\textcircled{R}_B$ with $\textcircled{F}_j$. Clearly the coordinate axes of the $\textcircled{C}_j$ are mutually parallel, the x-axes are parallel to $\textcircled{A}$ and the y-axes are in $\textcircled{M}$. The coordinate systems Ⓒ$_j$ will henceforth be known as the translated coordinate system associated with the jth surface. Collectively, the Ⓒ$_j$ will be termed translated coordinates. The planes $x_j = 0$ are known simply as the x_j-planes. Except when confusion is likely to arise the subscript j (j=1, 2,...,k) will, in future, be suppressed. In Ⓒ an arbitrary ray Ⓡ may be specified some point $\vec{Y}$ on the ray and the direction cosines $\vec{\beta} = (\alpha,\beta,\gamma)$ of the ray. (Note: The sign conventions used in this thesis are those of coordinate geometry and not the traditional ones of geometrical optics; see p.6 §c.) A convenient choice for the initial point $\vec{Y}$ of the ray Ⓡ is the point of intersection of Ⓡ with the x-plane. Thus $\vec{Y} = (0,Y,Z)$. Define the direction tangents $\underset{\sim}{V}_A$ of Ⓡ to be

$$\underset{\sim}{V}_A = \underset{\sim}{\beta}/\alpha \quad . \tag{5.1}$$

Clearly Ⓡ may be uniquely specified by the four numbers $\underset{\sim}{Y} = (Y,Z)$, $\underset{\sim}{V}_A = (V_A,W_A)$. However, it is convenient to introduce variables $\underset{\sim}{V}$ defined by

$$\underset{\sim}{V} = \underset{\sim}{V}_A - \underset{\sim}{V}_B \tag{5.2}$$

where $\underset{\sim}{V}_B = (\beta_B/\alpha_B,0)$ are the direction tangents of Ⓡ$_B$. The canonical variables of Ⓡ at the jth surface are $\underset{\sim}{Y}_j$, $\underset{\sim}{V}_j$ and at the first surface in particular $\underset{\sim}{Y}_1$, $\underset{\sim}{V}_1$ are called the canonical coordinates of Ⓡ and uniquely specify the ray. The canonical variables of the base-ray are all zero. Analagous definitions and equations hold for the canonical variables after refraction at the jth surface.

(b) Suppose that the canonical variables of the ray Ⓡ are required in some coordinate system Ⓒ* other than Ⓒ. Ⓒ and Ⓒ* are assumed to be related by a rotation about the z-axis so that the direction cosines of Ⓡ$_B$ in Ⓒ* are $\vec{\beta}_B$*. The transformation from Ⓒ to Ⓒ* is of the form

$$x^* = ax + by \quad , \quad y^* = -bx + ay \quad , \quad z^* = z \tag{5.3}$$

where

$$a = \alpha_B\alpha_B^* + \beta_B\beta_B^* \quad , \quad b = \alpha_B^*\beta_B - \beta_B^*\alpha_B \quad . \tag{5.4}$$

The canonical variables in Ⓒ* are defined in an analagous fashion to those in Ⓒ. Thus

$$V^* = \beta^*/\alpha^* - V_B^* = (\alpha_B/\alpha_B^*)^2 V/[1+(\alpha_B b/\alpha_B^*)V]$$

$$W^* = (\alpha_B/\alpha_B^*)W/[1+(\alpha_B b/\alpha_B^*)V] \quad .$$

Ⓡ will intersect the x*-plane in the point $x^* = 0$, $\underset{\sim}{Y}^* = \underset{\sim}{y}^*$. Thus,

$$x = -by/a \quad , \quad Y^* = y/a \quad , \quad Z^* = z \quad . \tag{5.5}$$

The equation of Ⓡ in Ⓒ is

$$\underset{\sim}{y} = \underset{\sim}{Y} + x(\underset{\sim}{V}+\underset{\sim}{V}_B) \tag{5.6}$$

and thus, from (5.6),

$$y = Y/[1+(V+V_B)b/a] \quad , \quad x = -bY/[a+bV_B+bV] \quad .$$

From (5.6) and the equations preceding these

$$Y^* = (\alpha_B/\alpha_B^*)\chi Y \quad , \quad Z^* = \chi[Z+(\alpha_B b/\alpha_B^*)(ZV-YW)] \quad ,$$

$$V^* = (\alpha_B/\alpha_B^*)^2\chi V \quad , \quad W^* = (\alpha_B/\alpha_B^*)\chi W \quad , \tag{5.7}$$

$$\chi = [1+(\alpha_B b/\alpha_B^*)V]^{-1} \quad . \tag{5.8}$$

If the origin is also transferred, to (d, y_0, z_0) say, $\underset{\sim}{Y}$ must be replaced by $\underset{\sim}{Y} - \underset{\sim}{y}_0 + d\underset{\sim}{V} + d\underset{\sim}{V}_B$. By means of these equations any results obtained for translated coordinates may be referred to arbitrary coordinates. (See also §85.) Furthermore, any coefficients relevant to the new coordinates can be expressed in terms of those relevant to the original coordinates (see §19). Use will be made of these results when discussing parabasal imagery (§20,21).

In view of the symmetry of Ⓚ the optical properties of Ⓚ and of systems of rays traversing Ⓚ must be invariant under reflection in Ⓜ. Thus, under reflection in Ⓜ Y and V are invariant whereas Z and W change sign. Hence Y and V are <u>even</u>[1] functions of Z_1 and W_1, Z and W are <u>odd</u>[1] functions of Z_1 and W_1. Any two-vector which transforms under reflection in Ⓜ like the canonical variables is said to be <u>proper two-vector</u>.

6. Paracanonical Coordinates

(a) It is often convenient in practice to specify a ray by some combination of its canonical coordinates. Such a situation occurs when the object is an extended element of some surface $Ⓗ_1$ or when vignetting can be effectively represented by some aperture in the object space of Ⓚ. These combinations of $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$ are called <u>paracanonical coordinates</u> and may be non-linear (e.g., when $Ⓗ_1$ is a plane surface not parallel to the x-plane). It is natural to choose the paracanonical coordinates so that they reflect the plane symmetry of Ⓚ. If $\underset{\sim}{S} = (S_y,S_z)$ and $\underset{\sim}{T} = (T_y,T_z)$ are the paracanonical coordinates to be used in place of $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$, S_y and T_y must be even functions of Z_1 and W_1, S_z and T_z odd functions of Z_1 and W_1. This ensures that $\underset{\sim}{Y}$ and $\underset{\sim}{Z}$ have the same symmetry properties with respect to $\underset{\sim}{S}$ and $\underset{\sim}{T}$ as they have with respect to $\underset{\sim}{Y}_1$ and $\underset{\sim}{Z}_1$. Consequently, in the case when $\underset{\sim}{S}$ and $\underset{\sim}{T}$ are linear functions of $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$, write

---

[1] A quantity is said to be even or odd in a pair of variables if under a simultaneous reversal of sign of these variables it remains unchanged in sign or changes sign respectively.

$$\underset{\sim}{S} = \underset{\sim}{\sigma}\underset{\sim}{Y}_1 + \underset{\sim}{\bar{\sigma}}\underset{\sim}{V}_1 \quad , \quad \underset{\sim}{T} = \underset{\sim}{\tau}\underset{\sim}{Y}_1 + \underset{\sim}{\bar{\tau}}\underset{\sim}{V}_1 \quad . \tag{6.1}$$

The eight constants $\underset{\sim}{\sigma}, \underset{\sim}{\bar{\sigma}}, \underset{\sim}{\tau}$ and $\underset{\sim}{\bar{\tau}}$ are subject only to the condition

$$\underset{\sim}{g} \equiv \underset{\sim}{\sigma}\underset{\sim}{\bar{\tau}} - \underset{\sim}{\bar{\sigma}}\underset{\sim}{\tau} \neq 0 \quad , \tag{6.2}$$

and completely characterize the set of paracanonical coordinates being used. Since $\underset{\sim}{g} \neq 0$, (6.1) may be inverted:[1]

$$\underset{\sim}{Y}_1 = (\underset{\sim}{\bar{\tau}}\underset{\sim}{S} - \underset{\sim}{\bar{\sigma}}\underset{\sim}{T})/\underset{\sim}{g} \quad , \quad \underset{\sim}{V}_1 = (-\underset{\sim}{\tau}\underset{\sim}{S} + \underset{\sim}{\sigma}\underset{\sim}{T})/\underset{\sim}{g} \quad . \tag{6.3}$$

Four rays of particular importance (§15) are the so-called <u>meridional and sagittal a- and b-rays</u>, whose paracanonical coordinates are given in Table 6/1. If the paracanonical coordinates are in fact the canonical coordinates (i.e., $\underset{\sim}{\sigma} = \underset{\sim}{\bar{\tau}} = 1$, $\underset{\sim}{\bar{\sigma}} = \underset{\sim}{\tau} = 0$) the $\underline{a}$- and $\underline{b}$-rays are called $\underline{p}$- and $\underline{q}$-rays respectively. <u>Note</u>: Paracanonical coordinates are defined in the object space only and there are no "paracanonical variables" (see also footnote 1, §19).

<u>TABLE 6/1</u> <u>Paracanonical Coordinates of a- and b-rays</u>

	S_y	S_z	T_y	T_z
meridional $\underline{a}$-ray	1	0	0	0
sagittal $\underline{a}$-ray	0	1	0	0
meridional $\underline{b}$-ray	0	0	1	0
sagittal $\underline{b}$-ray	0	0	0	1

---

[1] $\underset{\sim}{Q}/\underset{\sim}{q}$ is to be understood to mean $\underset{\sim}{Q}\underset{\sim}{\bar{q}}$ where $\underset{\sim}{\bar{q}} = (1/q_y \; , \; 1/q_z)$.

(b) A class of paracanonical coordinates of particular importance are the so-called special (linear) paracanonical coordinates or SPC. Suppose vignetting by Ⓚ can be represented by an aperture Ⓔ in the plane $x_1 = p$ (see also §61). If E_B is the basal point of Ⓔ, define $\underset{\sim}{Y}_E$ to be the y- and z-coordinates of the point of intersection of a ray with Ⓔ, referred to local coordinates in Ⓔ such that the origin is E_B and the y-axis lies in Ⓜ. Thus

$$\underset{\sim}{Y}_E = \underset{\sim}{Y}_1 + p\underset{\sim}{V}_1 \quad . \tag{6.4}$$

Further, let $Ⓗ_1$ be the plane $x_1 = \ell_1$ and 0_1 any point in $Ⓗ_1$. If 0_{B1} is the basal point of $Ⓗ_1$, define $\underset{\sim}{H}_1$ as the coordinates of 0_1 referred to a coordinate system in $Ⓗ_1$ whose origin is 0_{B1} and y-axis is in Ⓜ. Thus

$$\underset{\sim}{H}_1 = \underset{\sim}{Y}_1 + \ell_1\underset{\sim}{V}_1 \quad . \tag{6.5}$$

($\underset{\sim}{H}_1$ is called the object height of 0_1, see §7a.) Granted these definitions, the SPC of the ray from the object 0_1 and intersecting Ⓔ at $\underset{\sim}{Y}_E$ are defined by

$$\underset{\sim}{S} = \underset{\sim}{\sigma}Y_E \quad , \quad \underset{\sim}{T} = \underset{\sim}{\tau}\underset{\sim}{H}_1 \quad , \tag{6.6}$$

whence

$$\bar{\underset{\sim}{\sigma}} = p\underset{\sim}{\sigma} \quad , \quad \bar{\underset{\sim}{\tau}} = \ell_1\underset{\sim}{\tau} \quad .$$

Furthermore, it is evident from (6.4,5) that the y- and z-components of each of $\underset{\sim}{\sigma},\underset{\sim}{\tau},\bar{\underset{\sim}{\sigma}}$ and $\bar{\underset{\sim}{\tau}}$ are equal, that is, $\sigma_y = \sigma_z$, etc. . In this case $\underset{\sim}{\sigma},\underset{\sim}{\tau},\bar{\underset{\sim}{\sigma}}$ and $\bar{\underset{\sim}{\tau}}$ are written $\sigma,\tau,\bar{\sigma}$ and $\bar{\tau}$ respectively. In future, irrespective of the nature of the paracanonical coordinates, the "y" and "z" will be omitted from the components of $\underset{\sim}{\sigma},\underset{\sim}{\tau},\bar{\underset{\sim}{\sigma}},\bar{\underset{\sim}{\tau}}$ and $\underset{\sim}{g}$. They may be readily restored if required since g_y etc., multiplies or divides only quantities with a subscript "y", g_z etc., multiplies or divides only quantities with a subscript "z".

7. Image Heights and Aberrations

(a) The optical system Ⓚ is used to image an object of finite extent on the object surface $Ⓗ_1$ onto some image surface Ⓗ. It is natural that these surfaces should be consistent with the symmetry of Ⓚ and accordingly any object or image surface is assumed to be invariant under reflection in Ⓜ. Let Ⓡ be an arbitrary ray from the point object 0_1 in $Ⓗ_1$. 0, the image of 0_1 in Ⓗ by the ray Ⓡ, is the point of intersection of Ⓡ with Ⓗ, and the corresponding image height $\underset{\sim}{H}$ is the displacement of 0 from the basal point 0_B of Ⓗ. The definition of object height is analagous and consistent with the definition of $\underset{\sim}{H}_1$ used in §6. The image 0 is uniquely specified by Ⓗ and $\underset{\sim}{H}$. Note: the image surface of $Ⓕ_j$ can be treated as the object surface of $Ⓕ_{j+1}$ provided the only rays which are considered to emanate from some point of its object surface are just those rays which originate from the original object 0_1 and pass through the point in question. Under these conditions, the terms image and object are interchangeable. (cf. the case of the axial theory where ideal images are uniquely defined. All rays from a given object will pass through the (ideal) image and no qualifications such as the above are required in order to consider the image by one surface as the object for the next.)

Suppose Ⓗ is a plane which is not necessarily parallel to the x-plane. Let $Ⓒ^*$ be a cartesian coordinate system in Ⓗ and whose origin is 0_B, whose y-axis lies in Ⓜ and whose z-axis is normal to Ⓗ. Then, if $\vec{y} = (x,y,z)$ is a point on the ray Ⓡ

$$\underset{\sim}{y} = \underset{\sim}{Y} + x\underset{\sim}{V} + x\underset{\sim}{V}_B \tag{7.1}$$

In $Ⓒ^*$ the coordinates of $\vec{y}$ are

$$x^* = \alpha^*(x-\ell) + \beta^*(y-\ell V_B) \ , \ y^* = -\beta^*(x-\ell) + \alpha^*(y-\ell V_B) \ , \ z^* = z \quad , \qquad (7.2)$$

where in Ⓒ$_B$ the coordinates of 0_B are $(\ell,\ell V_B,0)$ and the normal to Ⓗ is $(\alpha^*,\beta^*,0)$. In order that (x,y,z) be a point on Ⓗ, $x^* = 0$, that is,

$$x = \ell(1+V^*V_B) - V^*y \quad . \qquad (7.3)$$

Thus, from (7.1,2)

$$H_y = y^* = (Y+\ell V)/(\alpha^*+\beta^*V_B+\beta^*V) \quad ,$$

$$H_z = z^* = Z + \ell W - \beta^*H_yW \quad . \qquad (7.4)$$

(Compare the above with the discussion leading to (5.7).) In the special case when Ⓗ is parallel to the x-plane, $\alpha^* = 1$, $\beta^* = 0$, and (7.4) becomes

$$\underset{\sim}{H} = \underset{\sim}{Y} + \ell\underset{\sim}{V} \quad . \qquad (7.5)$$

(b) Ⓗ$_1$ will be stigmatically imaged onto Ⓗ if and only if for any object 0_1 of object height $\underset{\sim}{H}_1$, the corresponding image height is given by

$$\underset{\sim}{H} = \underset{\sim}{m}\underset{\sim}{H}_1 \quad , \qquad (7.6)$$

where $m = (m_y,m_z)$ is some <u>magnification</u> which is the same for all rays from 0_1 but not necessarily independent of $\underset{\sim}{H}_1$. That is,

$$\underset{\sim}{m} = \underset{\sim}{m}(\underset{\sim}{H}_1) \quad ,$$

and the image is said to suffer from <u>distortion</u>. If $\underset{\sim}{m}$ is in fact independent of $\underset{\sim}{H}_1$ imagery will be <u>distortion free</u>. In that case different values for m_y and m_z will not be considered as constituting distortion since their dissimilarity corresponds simply to a difference in scale along the y- and z-axes. Any deviation from (7.6) is regarded as an <u>aberration</u>; the aberration $\underset{\sim}{\epsilon}$ is defined by

$$\underset{\sim}{\epsilon} = \underset{\sim}{\epsilon}(\underset{\sim}{S},\underset{\sim}{T}) = \underset{\sim}{H} - \underset{\sim}{m}\underset{\sim}{H}_1 \qquad (7.7)$$

where $\underset{\sim}{S}$ and $\underset{\sim}{T}$ are the paracanonical coordinates of rays from O_1. If the image is distortion free, $\underset{\sim}{\epsilon}$ is given by (7.7) provided $\underset{\sim}{m}$ is taken to be independent of $\underset{\sim}{H}_1$. <u>Distortion</u> is then by definition the aberration which depends only on the object. However, in applications of the basal theory to symmetric systems, the distortion will generally be large and unavoidable (e.g., in wide-angle systems) or irrelevant in the sense that it can be allowed for after the image has been formed by Ⓚ.[1] In such cases the general definition (7.7) of the aberration will be used, or preferably, if the image height $\underset{\sim}{H}_P(\underset{\sim}{H}_1)$ of the principal ray from the object can be predetermined (see §65) the aberration is defined by

$$\underset{\sim}{\epsilon} = \underset{\sim}{H} - \underset{\sim}{H}_P(\underset{\sim}{H}_1) \quad . \tag{7.8}$$

The aberration is then said to be <u>referred to the principal ray</u> (see §66).

(c) The simplicity of (7.5) as compared with (7.4) suggests that it is desirable to work with image planes normal to the x-axis. In general, however, the image and object planes need not be parallel, in which case Ⓐ (and hence all the x-axes) could be taken normal to either $Ⓗ_1$ or $Ⓗ_k'$. If the former, SPC can be used with profit, whereas (7.4) must be used in the image space and the surface contributions to the quasi-invariants (§22-24) are not simply related to the aberration (cf. §22a and 51b). On the other hand, if Ⓐ is taken normal to $Ⓗ_k'$, (7.5) can be used but not SPC unless the object is infinitely distant or only a single, fixed, point object is being considered. Curved object surfaces and object planes not normal to the x-axis are most conveniently

---

[1] See for instance Hill (1924) and Gardner and Washer (1948).

treated using non-linear paracanonical coordinates[2] with the pseudo-axis taken normal to the image plane. This method ensures that all the equations developed in this thesis for the case of parallel object and image spaces remain valid and applicable in the more general case. Moreover, the full power of analyses of the surface increments to the final aberration is readily available for design work. Accordingly, formulae for the aberration coefficients will be derived under the assumption that the object and image spaces are plane and parallel. Curved image surfaces may be treated in a manner analagous to that used when a symmetric system has curved image surfaces.

8. The Law of Refraction

Let $\vec{\beta}$ and $\vec{\beta}'$ be the direction cosines of the ray Ⓡ before and after refraction at the surface Ⓕ and let $\vec{n}$ be the inwardly directed unit normal to Ⓕ at the point of incidence Ⓟ of Ⓡ with Ⓕ. According to the law of refraction $\vec{n}$, $\vec{\beta}$ and $\vec{\beta}'$ are coplanar and hence linearly dependent. Moreover, if I and I' are the angles of incidence and refraction of Ⓡ, measured from $\vec{n}$ to Ⓡ, then $\Delta(N \sin I) = 0$, where, for any quantity Q, $\Delta Q_j = Q_j' - Q_j$ is the change in Q upon refraction at

---

2 If non-linear paracanonical coordinates are used $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$ will be expressed as series in $\underset{\sim}{S}$ and $\underset{\sim}{T}$. The effect of this is that the increments $\underset{\sim}{\delta}_{yj}$ and $\underset{\sim}{\delta}_{vj}$ in (25.6) do <u>not</u> vanish at the first surface. The most convenient manner in which non-linear or paracanonical coordinates may be used is to suppose there exists a fictitious surface in the object space that introduces exactly the "aberrations", represented by $\underset{\sim}{\delta}_{y1}$ and $\underset{\sim}{\delta}_{v1}$, required to produce the series for $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$. The theory would then proceed as developed in the following chapters.

the jth surface. It follows that

$$\Delta N\vec{\beta} = (\Delta N \cos I)\vec{n} \quad .^{1} \tag{8.1}$$

Introduce two quantities $\bar{R}_A$ and R such that

$$\bar{R}_A = (\cos I' - k \cos I)/\alpha' = \bar{R}_B R \tag{8.2}$$

where $\bar{R}_B$ is the value of $\bar{R}_A$ for the base-ray. (Note: For the base-ray R = 1.) k = N/N' is the refractance of Ⓕ and need not be confused with the $\underline{k}$ denoting the last surface of the system since $\underline{k}$ appears only as a subscript or as a summation limit. (8.1,2) combined give

$$\vec{\beta}' - k\vec{\beta} = \bar{R}_B R\alpha'\vec{n} \quad . \tag{8.3}$$

9. Refraction Increments for $\underset{\sim}{V}_A$ and $\underset{\sim}{Y}_A$. Equation for $\bar{R}_A$.

(a) From (8.3)

$$1 - k(\alpha/\alpha') = \bar{R}_A n_x \quad ,$$

$$\underset{\sim}{V}_A' - k(\alpha/\alpha')\underset{\sim}{V}_A = \bar{R}_A \underset{\sim}{n} \quad . \tag{9.1}$$

Thus, eliminating $k\alpha/\alpha'$ between this pair of equations,

$$\Delta\underset{\sim}{V}_A = - \bar{R}_A(n_x\underset{\sim}{V}_A - \underset{\sim}{n}) = -\bar{R}_A\underset{\sim}{I}_A \tag{9.2}$$

where $\underset{\sim}{I}_A$ is defined by

$$\underset{\sim}{I}_A = n_x\underset{\sim}{V}_A - \underset{\sim}{n} \quad . \tag{9.3}$$

If the coordinates of the point of intersection Ⓟ of any ray Ⓡ with Ⓕ are $\vec{y} = (X,\underset{\sim}{y})$, then

$$\underset{\sim}{y} = \underset{\sim}{Y} + X\underset{\sim}{V}_A$$

---

1 $(\Delta AB)C$ means $(A'B'-AB)C$.

and since $\Delta\vec{y} = 0$

$$\Delta \underset{\sim}{Y} = - X\Delta \underset{\sim}{V}_A = \bar{R}_A X \underset{\sim}{I}_A \quad . \tag{9.4}$$

Equations (9.2,4) give the changes in $\underset{\sim}{V}_A$ and $\underset{\sim}{Y}$ for any ray upon refraction at the surface Ⓕ.

To find an interpretation for $\underset{\sim}{I}_A$, it is sufficient to note that

$$\alpha \underset{\sim}{I}_A = n_x \underset{\sim}{\beta} - \alpha \underset{\sim}{n}$$

has the form of the components of the vector product of $\vec{n}$ and $\vec{\beta}$. Define

$$\underset{\sim}{I}_{Ax} = n_z V_A - n_y W_A \quad . \tag{9.5}$$

Then

$$\vec{n} \times \vec{\beta} = (n_y\gamma - n_z\beta \ , \ n_z\alpha - n_x\gamma \ , \ n_x\beta - n_y\alpha)$$

$$= \alpha(-I_{Ax} \ , \ -I_{Az} \ , \ I_{Ay}) \quad .$$

From (8.1) it follows that

$$N'n \times \vec{\beta}' = N\vec{n} \times \vec{\beta} \quad .$$

Thus

$$\Delta N\alpha I_{Ax} = \Delta N\alpha \underset{\sim}{I} = 0 \quad . \tag{9.6}$$

(b) A quadratic equation can be determined expressing $\bar{R}_A$ in terms of $\underset{\sim}{Y}$, $\underset{\sim}{V}_A$ and the surface normal at the point of intersection of the ray with Ⓕ. From the definition of $\underset{\sim}{V}_A$ it follows that

$$\alpha^2 = 1/(1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A) \quad .$$

Thus, from (9.1)

$$(1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A)\ (1-\bar{R}_A n_x)^2 = k^2(1+\underset{\sim}{V}_A' \cdot \underset{\sim}{V}_A')$$

$$= k^2(1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A + 2\underset{\sim}{V}_A \cdot \Delta\underset{\sim}{V}_A + \Delta\underset{\sim}{V}_A \cdot \Delta\underset{\sim}{V}_A) \quad .$$

Substituting for $\Delta\underset{\sim}{V}_A$ according to (9.2), a quadratic equation for $\bar{R}_A$

is obtained:

$$\tau \bar{R}_A^{\,2} + 2\sigma \bar{R}_A + \mu = 0 \tag{9.7}$$

where

$$\tau = k^2 \underset{\sim}{I}_A \cdot \underset{\sim}{I}_A - (1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A) n_x^{\,2}$$

$$\sigma = -k^2 \underset{\sim}{I}_A \cdot \underset{\sim}{V}_A + (1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A) n_x$$

$$\mu = (k^2-1)(1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A) \quad . \tag{9.8}$$

Equations (9.7,8) are the analogues of (M62.4,5).

In the analysis of this section no particular form for the surface Ⓕ was assumed. It follows that the equations of this section are entirely general. They would form the starting point of any computation of the aberrations and aberration coefficients of any optical system consisting of homogeneous isotropic media. In Appendix C these equations are applied to a scheme for tracing rays through systems of arbitrary surfaces.

10. Refraction and Transfer Increments for the Canonical Variables

(a) If the normal to Ⓕ at $Ⓟ_B$ has the components $\vec{n}_B = (n_{Bx}, n_{By}, 0)$ in Ⓒ, the equation of the surface Ⓕ is

$$(c\vec{y}-\vec{n}_B)\cdot(c\vec{y}-\vec{n}_B) = 1 \quad , \tag{10.1}$$

where c is the curvature of Ⓕ and points on Ⓕ have the coordinates $\vec{y} = (X,y,z)$. The ordinates of points on Ⓕ are exceptionally denoted by lower case symbols (see §12). The abscisse X is of prime importance and $\underset{\sim}{y}$ may be eliminated in favour of X by using (5.6). Accordingly, the vector $\vec{n}$ is

$$\vec{n} = [n_{Bx}-cX \ , \ n_{By}-c(Y+XV_A) \ , \ -c(Z+XW_A)] \quad . \tag{10.2}$$

Define I_x and $\underset{\sim}{I}$ by

$$I_x = I_{Ax} - I_{Bx} \quad , \quad \underset{\sim}{I} = \underset{\sim}{I}_A - \underset{\sim}{I}_B \quad , \tag{10.3}$$

From (9.3) and (10.2) $\underset{\sim}{I}_A$ may be written for spherical surfaces as

$$\underset{\sim}{I}_A = c\underset{\sim}{Y} + n_{Bx}\underset{\sim}{V}_A - \underset{\sim}{n}_B \quad .$$

In this case

$$I_x = c(YW-ZV) - (n_{By}W+cZV_B) \quad , \quad \underset{\sim}{I} = c\underset{\sim}{Y} + n_{Bx}\underset{\sim}{V} \quad , \tag{10.4}$$

and from (9.6)

$$\Delta(N\alpha\underset{\sim}{I}) = -N\alpha_B\underset{\sim}{I}_B\Delta(\alpha/\alpha_B) \quad . \tag{10.5}$$

I_x plays no role in the theory but $\underset{\sim}{I}$ is of considerable importance.

From the definition of V, (5.2), and (9.2) the change in $\underset{\sim}{V}$ due to refraction is found to be

$$\Delta\underset{\sim}{V} = -\bar{R}_A\underset{\sim}{I}_A + \bar{R}_B\underset{\sim}{I}_B = -\bar{R}_B[(R-1)\underset{\sim}{I}_B+R\underset{\sim}{I}] \quad . \tag{10.6}$$

Similarly, from (9.4), it is found that

$$\underset{\sim}{Y} = \bar{R}_B P(\underset{\sim}{I}+\underset{\sim}{I}_B) \quad , \tag{10.7}$$

where P is defined by

$$P = RX \quad . \tag{10.8}$$

It is evident from (10.6,7) that if Z = W = 0 at the first surface, then Z = W = 0 at all surfaces. However, if Y = V = 0 at the first surface, in general $Y \neq 0$, $V \neq 0$ at subsequent surfaces. This verifies the remarks of §3 on the non-global nature of "sagittal" as a property of a ray.

(b) If Q is some quantity referring to a given surface the corresponding quantity at the next surface is denoted by Q_+. Moreover, evaluated along a given ray,

$$\nabla Q_j = Q_{j+1} - Q_j' \quad , \quad \nabla Q = Q_+ - Q' \quad , \tag{10.9}$$

represent the change in Q due to transfer from the image space of Ⓕ$_j$ to the object space of Ⓕ$_{j+1}$. Since Ⓒ and Ⓒ$_+$ are related by a translation

$$\nabla \underset{\sim}{V}_A = \nabla \underset{\sim}{V} = 0 \quad . \tag{10.10}$$

If the distance, measured along Ⓡ$_B$, between Ⓟ$_B$ and Ⓟ$_{B+}$ is d_B' , the separation of the x- and x_+-planes is $d' = d_B'\alpha_B'$ and a ray Ⓡ will intersect the x_+-plane in the point whose coordinates are

$$\bar{\underset{\sim}{y}}_+ = \underset{\sim}{Y}' + d'\underset{\sim}{V}_A'$$

in Ⓒ. The corresponding point $\bar{\underset{\sim}{y}}_{B+}$ for the base-ray is the origin for Ⓒ$_+$. Accordingly, the initial point of Ⓡ in Ⓒ$_+$ is

$$\underset{\sim}{Y}_+ = \bar{\underset{\sim}{y}}_+ - \bar{\underset{\sim}{y}}_{B+} = \underset{\sim}{Y}' + d'\underset{\sim}{V}' \quad . \tag{10.11}$$

Gathering together the equations (10.6,7,10,11), the refraction and transfer increments to the canonical variables are:

$$\begin{aligned} \Delta\underset{\sim}{Y} &= \bar{R}_B P(\underset{\sim}{I}+\underset{\sim}{I}_B) \quad , & \nabla\underset{\sim}{Y} &= d'\underset{\sim}{V}' \quad , \\ \Delta\underset{\sim}{V} &= -\bar{R}_B[(R-1)\underset{\sim}{I}_B+R\underset{\sim}{I}] \quad , & \nabla\underset{\sim}{V} &= 0 \quad . \end{aligned} \tag{10.12}$$

11. Pseudo-parameters

It is evident that the equations of the preceding sections involve certain quantities pertaining to the base-ray and which are not parameters of the system in the sense of §1. These are the so-called pseudo-parameters referred to in §2 and are, explicitly, $\vec{R}_B$, $\vec{n}_B$, $\vec{\beta}_B$, $\underset{\sim}{V}_B$, $\underset{\sim}{I}_B$, $\cos I_B$, $\cos I_B'$, d_B' or equivalently d', $\vec{\beta}_{Bk}'$ and $\underset{\sim}{V}_{Bk}'$. Note that d' is not a proper parameter since it depends on the choice of the base-ray whereas its analogue in the axial theory (d') is a proper parameter of a symmetric system. Henceforth the term parameter is to be understood to mean both the pseudo-parameters and the $4\underset{\sim}{k}+1$ proper parameters of the system. The results of §8,9 can be used to derive relations between the pseudo-parameters which prove of immense value in simplifying the algebra of the basal theory (e.g., in §13,31,32). They could be used to eliminate certain pseudo-parameters from the theory, but since this effectively replaces a single, frequently occurring symbol by a collection of symbols, no advantage either in the appearance of the equations or in the subsequent computation of the aberration coefficients accrues from such an elimination.

According to (8.2) $\bar{R}_B$ is given by

$$\bar{R}_B = (\cos I_B' - k \cos I_B)/\alpha_B' \quad , \tag{11.1}$$

where

$$\cos I_B = \vec{n}_B \cdot \vec{\beta}_B \quad , \quad \cos I_B' = \vec{n}_B \cdot \vec{\beta}_B' \quad . \tag{11.2}$$

Applied to Ⓡ$_B$ the law of refraction (8.3) is

$$N'\vec{\beta}_B' - N\vec{\beta}_B = N'\alpha_B'\bar{R}_B\vec{n}_B \quad . \tag{11.3}$$

If $\vec{I}$ is the unit normal to the plane of incidence, $\vec{I} \sin I = \vec{n} \times \vec{\beta}$. Whence, for the base-ray,

$$\sin I_B = \alpha_B I_{By} \tag{11.4}$$

where

$$\underset{\sim}{I}_B = n_{Bx}\underset{\sim}{V}_B - \underset{\sim}{n}_B \quad . \tag{11.5}$$

(11.4,5) are consistent with

$$\Delta N\alpha_B \underset{\sim}{I}_B = 0 \quad . \tag{11.6}$$

Finally

$$\Delta\underset{\sim}{V}_B = -\bar{R}_B\underset{\sim}{I}_B \quad . \tag{11.7}$$

All z-components of two-vectors referring to the base-ray are zero. However, by continuing to write $Q_B = (Q_{By}, 0)$ for such quantities, pairs of equations can be readily combined into a single equation. This gives a more symmetrical appearance to the equations concerned and, incidentally, puts many of them into a form that is valid even when the Ⓚ is not plane symmetric, for example, (10.12).

III. PARABASAL OPTICS

12. Parabasal Optics and the Parabasal Equations

Parabasal optics concerns itself with the behaviour of families of rays as described by the dominant terms of the exact equations. The equations of parabasal optics are therefore obtained from the exact equations by neglecting all terms of degree higher than the first in the canonical variables. In this process no high order terms must be neglected until the exact equation corresponding to the required parabasal equation has, in principle, been obtained. Inattention to this has meant that first order terms have in fact been neglected in certain parabasal equations; for example, in the study of parabasal imagery and focal lines (see §18-21). Any ray whose behaviour is described by the equations of parabasal optics is said to be a parabasal ray. In as much as they are the dominant terms of the exact equations, the equations of parabasal optics are in fact relevant to arbitrary rays, not merely to parabasal rays. Quantities referring to general rays are usually denoted by upper case type, in which case the corresponding lower case type represents the same quantity in parabasal optics.

13. Parabasal Ray Trace Equations

(10.12) express $\Delta\underset{\sim}{Y}$ and $\Delta\underset{\sim}{V}$ as functions of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ through the dependence of P, R and $\underset{\sim}{I}$ on the canonical variables. $\Delta\underset{\sim}{Y}$ and $\Delta\underset{\sim}{V}$ may be formally expressed as power series in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ and the dominant terms of these series are obtained below, making it possible for the parabasal coefficients to be computed (§15). The higher order terms are not required, rather the expansions of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are obtained in terms of $\underset{\sim}{S}$

and $\underset{\sim}{I}$ (e.g., §25). The dominant terms of $\Delta\underset{\sim}{Y}$ and $\Delta\underset{\sim}{V}$, more precisely, of $\underset{\sim}{Y}'$ and $\underset{\sim}{V}'$, follow most readily from (10.1,5). First it is necessary to expand α in terms of $\underset{\sim}{V}$:

$$\alpha = (1+\underset{\sim}{V}_A \cdot \underset{\sim}{V}_A)^{-\frac{1}{2}} = \alpha_B(1+2\alpha_B\beta_B V+\underset{\sim}{V}\cdot\underset{\sim}{V})^{-\frac{1}{2}}$$

$$= \alpha_B(1-\alpha_B\beta_B V) + 0(2) \quad . \qquad (13.1)$$

In general 0(m) denotes any expression containing only terms of degree equal to or greater than $\underline{m}$ in the canonical variables or coordinates. The coordinates $\vec{y} = (X, Y+XV_A, Z+XW_A)$ of the point of incidence Ⓟ of Ⓡ with Ⓕ are of the first order and consequently (10.1) can be written $\vec{y}\cdot\vec{n}_B = 0(2)$. Hence

$$X(n_{Bx}+n_{By}V_B) + n_{By}Y = X/\lambda_1 + n_{By}Y = 0(2) \quad ,$$

where

$$\lambda_1 = \alpha_B/\cos I_B \quad . \qquad (13.2)$$

Thus

$$X = -\lambda_1 n_{By} Y + 0(2) \quad . \qquad (13.3)$$

Since $\Delta\vec{y} = 0$ and $\Delta\vec{n} = 0$

$$\Delta(\lambda_1 Y) = 0(2) \quad , \quad \Delta Z = 0(2) \quad . \qquad (13.4)$$

In view of (13.1) the dominant terms of (10.5) become

$$\Delta N\alpha_B(\underset{\sim}{I}-\alpha_B\beta_B\underset{\sim}{I}_B V) = 0(2) \quad . \qquad (13.5)$$

From the second component of this it follows that

$$\Delta N\alpha_B(cZ+n_{Bx}W) = (\Delta N\alpha_B W + N'\alpha_B{}'c\bar{R}_B Z)n_{Bx} = 0(2) \quad , \qquad (13.6)$$

where the last equality is obtained from the first by means of (11.3) and (13.4). The first component of (13.5) yields

$$\Delta N\alpha_B[cY+(n_{Bx}-\alpha_B\beta_B I_{By})V] = 0(2) \quad .$$

Now, according to (11.2,5),

$$n_{Bx} - \alpha_B\beta_B I_{By} = (n_{Bx}\alpha_B+\beta_B n_{By})\alpha_B = \alpha_B \cos I_B$$

and hence the preceding equation may be written

$$\Delta(N\alpha_B{}^2\cos I_B V) + N'\alpha_B{}'c\bar{R}_B\lambda_1 Y = 0(2) \quad . \tag{13.7}$$

Define

$$N_y = N\alpha_B{}^3 \quad , \quad N_z = N\alpha_B \quad ,$$
$$k_y = N_y/N_y{}' \quad , \quad k_z = N_z/N_z{}' \quad . \tag{13.8}$$

From (13.4,6) the <u>parabasal ray trace equations</u>, expressing $\underset{\sim}{y}'$ and $\underset{\sim}{v}'$ in terms of $\underset{\sim}{y}$ and $\underset{\sim}{v}$, may be written down:

$$y' = (\lambda_1/\lambda_1{}')y \quad , \quad v' = k_y(\lambda_1{}'/\lambda_1)v - (c\bar{R}_B\lambda_1/\alpha_B{}'\cos I_B{}')y \quad ,$$
$$z' = z \quad , \quad w' = k_z w - c\bar{R}_B z \quad . \tag{13.9}$$

Since the transfer equations (10.12) are already linear in $\underset{\sim}{Y}'$ and $\underset{\sim}{V}'$ they will not require modification. Thus, for parabasal rays:

$$\underset{\sim}{v}_+ = \underset{\sim}{v}' \quad , \quad \underset{\sim}{y}_+ = \underset{\sim}{y}' + d'\underset{\sim}{v}' \quad . \tag{13.10}$$

By means of (13.9,10) parabasal rays may be traced through the system. Note that these equations are readily generalized to the case of plane symmetric, aspheric surfaces (Appendix D).

14. <u>The Parabasal Coefficients</u>

(a) The parabasal refraction invariant $(\lambda_1 y,z)$ for a parabasal ray may be interpreted as follows: Rotate Ⓒ about the z-axis until the x-axis is in the direction of $\vec{n}_B$. The resulting coordinate system is $Ⓒ_N$ (the so-called normal coordinate system associated with the surface, see Appendix B) in which canonical variables $\underset{\sim}{Y}_N$ and $\underset{\sim}{V}_N$ may be defined as in §5a. From (5.7) $Y_N = \lambda_1 Y + 0(2)$, $Z_N = Z + 0(2)$. To the first order in $\underset{\sim}{Y}_N$, $\underset{\sim}{V}_N$ the coordinates of Ⓟ in Ⓒ$_N$ are $(0, Y_N, Z_N)$. Hence, for parabasal rays, $\lambda_1 y$ and z are respectively the distances of Ⓟ from the plane normal to Ⓜ and containing $\vec{n}_B$ and from the plane of symmetry Ⓜ.

(b) (13.9,10) show that $\underset{\sim}{y}_+$ and $\underset{\sim}{v}_+$ are linear functions of $\underset{\sim}{y}$ and $\underset{\sim}{v}$ and hence, iteratively, of $\underset{\sim}{y}_1$ and $\underset{\sim}{v}_1$. This may be formally expressed by writing

$$\underset{\sim}{y}_j = \underset{\sim}{y}_{pj}\underset{\sim}{y}_1 + \underset{\sim}{y}_{qj}\underset{\sim}{v}_1 \quad , \quad \underset{\sim}{v}_j = \underset{\sim}{v}_{pj}\underset{\sim}{y}_1 + \underset{\sim}{v}_{qj}\underset{\sim}{v}_1 \quad , \qquad (14.1)$$

where

$$\underset{\sim}{y}_p = (y_p, z_p) \ , \ \underset{\sim}{v}_p = (v_p, w_p) \ , \ \underset{\sim}{y}_q = (y_q, z_q) \ , \ \underset{\sim}{v}_q = (v_q, w_q) \quad ,$$

are the <u>canonical parabasal coefficients</u> of Ⓚ and are functions of the parameters of Ⓚ. Since $\underset{\sim}{y}_1$ and $\underset{\sim}{v}_1$ are linearly related to $\underset{\sim}{s}$ and $\underset{\sim}{t}$ by (6.3),[1] (14.1) becomes

$$\underset{\sim}{y}_j = \underset{\sim}{y}_{aj}\underset{\sim}{s} + \underset{\sim}{y}_{bj}\underset{\sim}{t} \quad , \quad \underset{\sim}{v}_j = \underset{\sim}{v}_{aj}\underset{\sim}{s} + \underset{\sim}{v}_{bj}\underset{\sim}{t} \quad . \qquad (14.2)$$

The eight coefficients $\underset{\sim}{y}_a$, $\underset{\sim}{y}_b$, $\underset{\sim}{v}_a$, and $\underset{\sim}{v}_b$ are the <u>(paracanonical) parabasal coefficients</u> of Ⓚ. Of these the four relating to the y-components of (14.2), viz. y_a, y_b, v_a and v_b, are the <u>meridional parabasal coefficients</u>; those relating to the z-components, viz. z_a, z_b, w_a and w_b, are the <u>sagittal parabasal coefficients</u>. Moreover, those bearing a subscript <u>a</u> or <u>b</u> are <u>a</u>- or <u>b</u>- <u>parabasal coefficients</u> respectively. Characteristic of the plane symmetry of Ⓚ, y_j and v_j are independent of s_z and t_z; z_j and w_j are

---

1 Note the use of $\underset{\sim}{s}$ and $\underset{\sim}{t}$ in place of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. (6.3) is valid in the parabasal region even for non-linear paracanonical coordinates.

independent of s_y and t_y. Moreover, the meridional and sagittal coefficients are dissimilar, that is, in general $y_a \neq z_a$ etc. (Compare the case of the axial theory, M§5a.) The determinants of the transformations (14.2) are simply the products of the determinants of the transformations (6.3) and (13.9,10). Thus

$$(y|v)_j = (1/g)\prod_{i=1}^{j} k_{yi} = N_{y1}/N_{yj}g \quad ,$$

$$(z|w)_j = (1/g)\prod_{i=1}^{j} k_{zi} = N_{z1}/N_{zj}g \quad , \tag{14.3}$$

where for any two symbols f and g

$$(f|g)_j = f_{aj}g_{bj} - f_{bj}g_{aj} \quad . \tag{14.4}$$

As a result of (14.3) only six of the eight parabasal coefficients at any surface are mutually independent. Note: Although all the equations of this section were written in a form appropriate to the object space of Ⓕ the same formal equations are valid in the image space provided that a prime (') is attached to all symbols bearing a subscript j, for example v_{aj} becomes v_{aj}'.

15. Computation of the Parabasal Coefficients

Parabasal rays may be traced through the system by applying the ray trace equations (13.9,10). By tracing the a- and b-rays (§6a) the parabasal coefficients are obtained. Assume that the required base-ray has been traced and from the values of the associated pseudo-parameters compute at each surface the coefficients of y and v in (13.9). Suppose a set of paracanonical coordinates have been specified by σ, $\bar{\sigma}$, τ and $\bar{\tau}$, and that the rays to be traced are given by their values of s and t. By means of (6.3) the corresponding values of y_1 and v_1 are computed. y_2

and $\underset{\sim}{v}_2$ follow from these in two steps: first $\underset{\sim}{y}_1'$ and $\underset{\sim}{v}_1'$ are computed from (13.9) and second, transfer to the second surface is carried out by means of (13.10). Using the values of $\underset{\sim}{y}_2$ and $\underset{\sim}{v}_2$ a repetition of these two steps gives $\underset{\sim}{y}_3$ and $\underset{\sim}{v}_3$, and so on. In this manner $\underset{\sim}{y}_k'$ and $\underset{\sim}{v}_k'$ are eventually determined.

Now suppose that the (parabasal) meridional a-ray is traced in this manner. From Table 6/1, $s_y = 1$, $t_y = s_z = t_z = 0$ and hence, from (14.2),

$$y_j = y_{aj} \ , \ v_j = v_{aj} \ , \ z_j = w_j = 0 \quad .$$

Thus the meridional a-coefficients are determined simply and directly by tracing the meridional a-ray. Similarly, the meridional b-coefficients and the sagittal a- and b-coefficients are determined by tracing the meridional b-ray and the sagittal a- and b-rays respectively. This illustrates the usefulness of the a- and b-rays.

16. Parabasal Invariants

(a) A quantity Q is an optical invariant provided for all rays

$$\Delta Q = \nabla Q = 0 \tag{16.1}$$

at each surface. If Q is in general not an invariant but none the less satisfies (16.1) in the parabasal region, it is called a quasi-invariant and the corresponding parabasal quantity q is a parabasal invariant. The two identities (14.3) are intimately related with certain parabasal invariants associated with two arbitrary parabasal rays. The form of (14.3) suggests investigating quantities such as

$$g_y = f_y(\bar{y}v - \bar{v}y) \quad ,$$

where f_y is a combination of the parameters of Ⓚ and $\underset{\sim}{y}$, $\underset{\sim}{v}$ and $\bar{\underset{\sim}{y}}$, $\bar{\underset{\sim}{v}}$ are the canonical variables of two arbitrary parabasal rays Ⓡ$_y$ and $\bar{Ⓡ}_y$ respectively. It is required to find the condition that g_y be a parabasal invariant. For this to be the case

$$\nabla g_y = (\bar{y}_+ v_+ - \bar{v}_+ y_+)\nabla f_y + f_y'\nabla(\bar{y}v - \bar{v}y) = 0$$

for all rays. In view of (13.10) $\nabla(\bar{y}v-\bar{v}y) = 0$ and thus

$$\nabla g_y = (\bar{y}_+ v_+ - \bar{v}_+ y_+)\nabla f_y = 0 \quad ,$$

which will be the case if and only if

$$\nabla f_y = 0 \quad . \tag{16.2}$$

Moreover, since $\Delta(\lambda_1 y) = 0$,

$$\Delta g_y = \lambda_1 \bar{y}\Delta(f_y v/\lambda_1) - \lambda_1 y\Delta(f_y \bar{v}/\lambda_1) \quad . \tag{16.3}$$

The first term of this is

$$\lambda_1 \bar{y}[(f_y'/N_y')\Delta(N_y v/\lambda_1)+(N_y v/\lambda_1)\Delta(f_y/N_y)] \quad .$$

Since $\Delta(N_y v/\lambda_1)$ is proportional to y (see (13.7)),

$$\Delta g_y = N_y(\bar{y}v-\bar{v}y)\Delta(f_y/N_y) \quad .$$

This will be zero for all rays if and only if $\Delta(f_y/N_y) = 0$, that is, if $f_y = aN_y$ where a is a constant which, without loss of generality, may be taken to be unity. $f_y = N_y$ satisfies (16.2) so that the quantity

$$g_y = N_y(\bar{y}v-\bar{v}y)$$

is in fact a parabasal invariant. In a similar fashion $N_z(\bar{z}w-\bar{w}z)$ can be shown to be a parabasal invariant where $\underset{\sim}{y}$,$\underset{\sim}{v}$ and $\bar{\underset{\sim}{y}}$,$\bar{\underset{\sim}{v}}$ are the canonical variables of two arbitrary parabasal rays Ⓡ$_z$ and $\bar{Ⓡ}_z$ respectively.

Thus, the quantity

$$\underset{\sim}{g} = \underset{\sim}{N}(\bar{\underset{\sim}{y}}\underset{\sim}{v} - \bar{\underset{\sim}{v}}\underset{\sim}{y}) \tag{16.4}$$

is a parabasal invariant. By expressing $\underset{\sim}{y}_j, \underset{\sim}{v}_j$ and $\bar{\underset{\sim}{y}}_j, \bar{\underset{\sim}{v}}_j$ in (16.4) in terms of $\underset{\sim}{s}, \underset{\sim}{t}$ and $\bar{\underset{\sim}{s}}, \bar{\underset{\sim}{t}}$ respectively by (14.2) (or (6.3) at the first surface) the identities (14.3) are readily recovered from the invariance of $\underset{\sim}{g}$.

(b) <u>Note</u>: g_y is independent of the z-components of the canonical variables of Ⓡ$_y$ and Ⓡ̄$_y$ and, moreover, g_z is independent of the y-components of the canonical variables of Ⓡ$_z$ and Ⓡ̄$_z$. Thus, the four rays Ⓡ$_y$, Ⓡ̄$_y$, Ⓡ$_z$ and Ⓡ̄$_z$ may be chosen independently of one another. A very convenient choice (§22,23) is to choose Ⓡ$_y$ and Ⓡ$_z$ as the <u>same</u> ray Ⓡ while Ⓡ̄$_y$ and Ⓡ̄$_z$ are respectively (parabasal) meridional and sagittal rays from the basal point O_B of the object surface. In this case $\underset{\sim}{g}$ is said to be defined for the ray Ⓡ with respect to the basal point O_B. g_y and g_z are called the <u>Lagrange invariants</u> of O_B.

17. <u>The Conrady s- and t-traces</u>

Mere inspection of (13.9,10) or (14.1) suffices to verify that any ray lying in the sagittal plane $y_1 = v_1 = 0$ of the object space of Ⓚ lies in the sagittal planes associated with each medium constituting Ⓚ; in other words, the ray in question is a true sagittal ray (see §3). Let O_B be a point on Ⓡ$_B$ and in the object space of the surface Ⓕ. It was shown by Conrady[1] that in the infinitesimal neighbourhood of the base-ray both a fan of meridional rays and a fan of sagittal rays, from

---

[1] Conrady (1957) §67.

O_B, produce on Ⓡ$_B$ a point image of O_B. These two images are in general distinct. Clearly these fans of rays can be respectively represented by fans of parabasal rays. Let $\underline{\underline{\ell}}$ be the distance of O_B from Ⓟ$_B$ and consider refraction by the single surface Ⓕ. In Ⓒ, the canonical variables of <u>meridional</u> rays from O_B satisfy

$$y = -\underline{\underline{\ell}}\alpha_B v \quad , \tag{17.1}$$

and in the image space of Ⓕ the running coordinates $(\bar{x}',\bar{y}'0)$ on any meridional ray satisfy

$$\bar{y}' = y' + \bar{x}'v' + \bar{x}'V_B' \quad .$$

These rays will intersect Ⓡ$_B$ in a unique point provided

$$y' + \bar{x}'v' = 0 \quad .$$

It follows from (13.9) that the condition that the rays form a point image of O_B at $(\bar{x}',\bar{y}',0)$ is

$$[\lambda_1/\lambda_1'-(c\bar{R}_B\lambda_1/\alpha_B'\cos I_B')\bar{x}']y + \bar{x}'k_y(\lambda_1'/\lambda_1)v = 0 \quad .$$

If y is expressed in terms of v by means of (17.1) this reduces to the requirement that $\underline{\underline{\ell}}$ and $\underline{\underline{\ell}}'$ satisfy

$$-(\lambda_1/\lambda_1'\alpha_B'\underline{\underline{\ell}}') + (k_y\lambda_1'/\lambda_1\alpha_B\underline{\underline{\ell}}) + (c\bar{R}_B\lambda_1/\alpha_B'\cos I_B') = 0 \quad ,$$

where $\underline{\underline{\ell}}'$ is the distance of the image from Ⓟ$_B$. (Note: $\bar{x}' = \underline{\underline{\ell}}'\alpha_B'$). If the detailed expressions for λ_1, k_y and $\bar{R}_B$ in terms of the pseudo-parameters are substituted into this, then

$$N'\cos^2 I_B'/\underline{\underline{\ell}}' - N\cos^2 I_B/\underline{\underline{\ell}} = c(N'\cos I_B'-N\cos I_B) \quad . \tag{17.2}$$

This equation is just Conrady's 67(c). In a similar fashion the sagittal fan from O_B will come to a focus at a point on Ⓡ$_B$ where

$$z' + \underline{\underline{\ell}}'\alpha_B'w' = 0 \quad ,$$

subject to

$$z + \underline{\underline{\ell}}\alpha_B w = 0 \quad .$$

Combined with (14.1), these last two equations yield the condition

$$(k_z\alpha_B'\underline{\underline{\ell}}'-\alpha_B\underline{\underline{\ell}}+c\alpha_B\alpha_B'\bar{\bar{R}}_B\underline{\underline{\ell\ell}}')w = 0 \quad ,$$

for a focus on Ⓡ$_B$ and distant $\underline{\underline{\ell}}'$ from Ⓟ$_B$. Thus

$$N'/\underline{\underline{\ell}}' - N/\underline{\underline{\ell}} = c(N'\cos I_B'-N\cos I_B) \quad , \tag{17.3}$$

in agreement with 67(d) of Conrady. Thus Conrady's s- and t- trace formulae are correctly recovered. It is evident that the two points determined by (17.2,3) are distinct if the angle of incidence I_B is not zero.

Consider a basal object O_B in the object space of a plane symmetric system Ⓚ. After refraction at the first surface, a fan of parabasal meridional rays from O_B will come to a focus at F_{M1}, say, on Ⓡ$_B$ and, furthermore, a fan of sagittal parabasal rays from O_B will come to a distinct focus at F_{S1}, also on Ⓡ$_B$. After subsequent refractions these two fans will respectively remain meridional and sagittal. Consequently, after refraction at the remaining surfaces, the meridional fan will produce foci F_{Mi}, $i = 2,\ldots,k$ and the sagittal fan will produce foci F_{Si}, $i = 2,\ldots k$. However, since F_{M1} and F_{S1} are distinct, any pair F_{Mi} and F_{Si} need not be distinct. The presence of the two foci associated with any surface is characteristic of astigmatism. As will be seen (§21b and 49), parabasal imagery is subject to first order linear astigmatism. The focus F_M is called the <u>meridional focus</u>, and F_S the <u>sagittal focus</u>. Note that (17.2,3) are entirely independent of the choice of coordinates Ⓒ. (See also §18,20c.)

18. Parabasal Imagery and Infinitesimal Canonical Variables

It has been traditional to define the parabasal region as the infinitesimal neighbourhood of the base-ray[1] and to discuss parabasal imagery using hamiltonian coordinates. The hamiltonian coordinate systems $\textcircled{C}_{Hi}$ are defined so that their x-axes lie along the base-ray and their y-axes lie in $\textcircled{M}$ (see Appendix B). There are two such systems associated with each surface. Two results of particular interest obtained under these conditions and appearing in the literature[2] are (i) any line $\textcircled{L}$ normal to $\textcircled{R}_B$ and in $\textcircled{M}$ is imaged by meridional rays into a line normal to $\textcircled{R}_B$, and (ii) corresponding to any basal point O_B there are two focal lines normal to $\textcircled{R}_B$ and to each other, and one of which lies in $\textcircled{M}$. Suppose translated coordinates are used in which $\textcircled{C}_1$ and $\textcircled{C}_{H1}$ are identical. In general, at other surfaces, $\textcircled{C}$ and $\textcircled{C}_H$ will differ by a rotation about the z-axis. Imagery may be determined in $\textcircled{C}$ directly, using the equations of the basal theory developed in this thesis, or by obtaining the image by some method employing hamiltonian coordinates and then referring this to $\textcircled{C}$ by a rotation of coordinates. The images obtained by these two methods must be identical. In fact, in the parabasal region as defined at the beginning of this section, this is not the case.

Consider imagery of the line $\textcircled{L}$ by meridional rays. If SPC are used an arbitrary point O_1 on $\textcircled{L}$ is specified by

$$t_y = \tau h_1 \quad , \quad t_z = 0 \quad ,$$

and meridional rays from O_1 have $s_z = 0$. (Note: in accordance with

---

1 Synge (1937) §9; Luneburg (1964) §33.4 .

2 Synge (1937) §9; Luneburg (1964) §35; Born and Wolf (1964) §4.6 .

§12, $\underset{\sim}{H}_1$ is replaced by $\underset{\sim}{h}_1$ in the parabasal region.) Consequently the equation of any ray from O_1 is of the form

$$y = (y_a+xv_a)s_y + (y_b+xv_b)t_y + xV_B \quad , \quad z = 0 \quad , \qquad (18.1)$$

in the image space, where primes have been omitted for convenience. All (parabasal) rays from O_1 will come to a focus provided the coefficient of s_y vanishes. Thus the image of O_1 is at

$$x = -y_a/v_a \quad , \quad y = -(y|v)\tau h_1/v_a + xV_B \quad , \quad z = 0 \quad . \qquad (18.2)$$

Since x is independent of h_1, all points O_1 on Ⓛ image onto the line $x = -y_a/v_a$ in Ⓜ. This is normal to the x-axis rather than to $Ⓡ_B$ and contradicts (i) above. It can also be shown (see §21) that if the parabasal region is as defined at the beginning of this section, (ii) is contradicted: when determined in the basal theory, the focal lines corresponding to O_B are normal to each other, and to the x-axis, which is not necessarily colinear with $Ⓡ_B$. Moreover, if the image of any point O_1 on Ⓛ is obtained by the two methods it is found that the image determined in one case is distinct from that obtained in the other case, whereas they should be identical. In particular, the coordinates of these two "images" differ by quantities linear in h_1 which certainly are not negligible under the assumptions relating to parabasal imagery in this section.

This strongly suggests that some terms which cannot be neglected have been omitted in the analysis leading to these results. V_B is of the zeroth order and hence terms linear in h_1 and appearing in x will certainly contribute to y. However, x was determined by the requirement that the coefficient of s_y in (18.1) be zero. Since $s_y h_1$ is of the second order of magnitude and hence negligible, those terms of x which are linear in h_1

cannot be determined by this method. Accordingly, consistent results cannot be obtained in parabasal imagery if the canonical variables are assumed to be infinitesimal. In order to be of practical value, the definition of "parabasal" must give consistent results irrespective of the coordinates used. The definition in §12 satisfies this requirement. In order to demonstrate this the dependence of the parabasal and higher order coefficients (in the Taylor series for $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$) on the chosen coordinate systems, must be known. This is found in §19 and parabasal imagery under the definition of §12 is considered in §20. This definition was motivated by the fact that significant terms have apparently been neglected in the traditional analysis. In order to avoid this, the exact equations must be determined to a higher order before any approximations are made.

19. <u>Coordinate Dependence of Paracanonical Coefficients</u>

The <u>paracanonical coefficients</u> of some quantity Q are the coefficients in the Taylor series for Q in terms of the paracanonical coordinates $\underset{\sim}{S}$ and $\underset{\sim}{T}$. In this thesis equations are derived whereby the paracanonical coefficients of quantities such as $\underset{\sim}{H}$, $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ may be computed for the translated coordinates introduced in §5. However, by means of (5.7,8) the canonical variables for arbitrary coordinates may be computed given those for translated coordinates. Alternatively, if the new canonical variables are expanded in terms of the paracanonical coordinates defined with respect to the arbitrary coordinate system, the coefficients in this expansion can be expressed in terms of the paracanonical coefficients of this thesis.

For simplicity it is assumed that the coordinate system Ⓒ of the object space is fixed. In the image space $Ⓒ_H'$ is orientated so that its x-axis lies along the base-ray and its y-axis in the plane of symmetry Ⓜ. Ⓒ' is any coordinate system, for example, the previously defined translated system, which may be obtained from $Ⓒ_H'$ by a rotation about the z-axis. In Ⓒ' $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ will have expansions of the form (14.2) with higher order terms added consistent with the symmetry requirements of §5. Thus

$$Y = y_a S_y + y_b T_y + s_{y1} S_y^2 + s_{y2} S_y T_y + s_{y5} T_y^2 + s_{y8} S_z^2 + s_{y9} S_z T_z + s_{y10} T_z^2 + 0(3) \ ,$$

$$Z = z_a S_z + s_b T_z + s_{z3} S_y S_z + s_{z4} S_y T_z + s_{z6} T_y S_z + s_{z7} T_y T_z + 0(3) \quad , \qquad (19.1)$$

with similar expansions for $\underset{\sim}{V}$ where the higher order coefficients have the subscripts y and z replaced by v and w respectively, for example, $s_{v1}, \ldots, s_{v10}, s_{w3}, \ldots, s_{w7}$.[1] (See also §24a,29.) In $Ⓒ_H'$, $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are distinguished by a subscript "H", thus $\underset{\sim}{Y}_H$ and $\underset{\sim}{V}_H$, and all coefficients pertaining to these also have a subscript H adjacent to the kernel symbol, for example, y_{Ha}, s_{Hz3} . The expansions of $\underset{\sim}{Y}_H$, $\underset{\sim}{V}_H$ are then of the form (19.1). (<u>Note</u>: Since the object space coordinates Ⓒ are fixed, $\underset{\sim}{S}$ and $\underset{\sim}{T}$ are the same for both $Ⓒ_H'$ and Ⓒ'.) From (5.7,8)

$$Y = \frac{1}{\alpha_B} (1+V_B V_H) Y_H + 0(3) \quad , \quad V = \frac{1}{\alpha_B^2} (1+V_B V_H) V_H \quad ,$$

$$Z = Z_H + V_B Y_H W_H + 0(3) \quad , \quad W = \frac{1}{\alpha_B} (1+V_B V_H) W_H \quad . \qquad (19.2)$$

If the series for $\underset{\sim}{Y}_H$ and $\underset{\sim}{V}_H$ corresponding to (19.1) are substituted into (19.2), $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are found in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Comparison of these

---

[1] The coefficients s_{y1} etc., are not to be confused with the parabasal coordinates s_y . The paracanonical coordinates <u>never</u> take suffixes in addition to "y" or "z".

TABLE 19/1 Coordinate Dependence of Paracanonical Coefficients

$$y_a = y_{Ha}/\alpha_B \ , \qquad z_a = z_{Ha} \ ,$$
$$y_b = y_{Hb}/\alpha_B \ , \qquad z_b = z_{Hb} \ ,$$
$$v_a = v_{Ha}/\alpha_B^2 \ , \qquad w_a = w_{Ha}/\alpha_B \ ,$$
$$v_b = v_{Hb}/\alpha_B^2 \ , \qquad w_b = w_{Hb}/\alpha_B \ .$$

$$s_{y1} = (s_{Hy1} + V_B y_{Ha} v_{Ha})/\alpha_B \ , \qquad s_{v1} = (s_{Hv1} + V_B v_{Ha}^2)\alpha_B^2 \ ,$$
$$s_{y2} = [s_{Hy2} + V_B(y_{Ha}v_{Hb} + y_{Hb}v_{Ha})]/\alpha_B \ , \qquad s_{v2} = (s_{Hv2} + 2V_B v_{Ha} v_{Hb})/\alpha_B^2 \ ,$$
$$s_{y5} = (s_{Hy5} + V_B y_{Hb} v_{Hb})/\alpha_B \ , \qquad s_{v5} = (s_{Hv5} + V_B v_{Hb}^2)/\alpha_B^2 \ ,$$
$$s_{yr} = s_{Hyr}/\alpha_B \ , \quad (r = 8,9,10) \qquad s_{vr} = s_{Hvr}/\alpha_B^2 \ . \quad (r = 8,9,10)$$

$$s_{z3} = s_{Hz3} + V_B y_{Ha} w_{Ha} \ , \qquad s_{w3} = (s_{Hw3} + V_B v_{Ha} w_{Ha})/\alpha_B \ ,$$
$$s_{z4} = s_{Hz4} + V_B y_{Ha} w_{Hb} \ , \qquad s_{w4} = (s_{Hw4} + V_B v_{Ha} w_{Hb})/\alpha_B \ ,$$
$$s_{z6} = s_{Hz6} + V_B y_{Hb} w_{Ha} \ , \qquad s_{w6} = (s_{Hw6} + V_B v_{Hb} w_{Ha})/\alpha_B \ ,$$
$$s_{z7} = s_{Hz7} + V_B y_{Hb} w_{Hb} \ , \qquad s_{w7} = (s_{Hw7} + V_B v_{Hb} w_{Hb})/\alpha_B \ .$$

expressions with (19.1) yield the paracanonical coefficients for $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ in terms of those for $\underset{\sim}{Y}_H$ and $\underset{\sim}{V}_H$. For example, Y becomes

$$Y = \frac{1}{\alpha_B}\,[1+V_B(v_{Ha}S_y+v_{Hb}T_y)]\,[y_{Ha}S_y+y_{Hb}T_y+s_{Hy1}S_y^{\;2}+\ldots.]\quad .$$

The coefficient of $S_y^{\;2}$ is thus

$$s_{y1} = \frac{1}{\alpha_B}\,(s_{Hy1}+V_Bv_{Ha}y_{Ha})\quad .$$

In this fashion the equations of Table 19/1 are obtained. These equations show how the parabasal and second order coefficients in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ depend on the orientation of the image space coordinates, the object space coordinates being fixed but arbitrary. However, analagous equations can be similarly derived showing how the coefficients depend on the choice of the object space coordinate system, or both the object and image space coordinate systems simultaneously.

20. Parabasal Imagery

In this and the following section, parabasal imagery will be discussed on the basis of the definition of the term parabasal given in §12.

(a) Let O_1 be a point object of object height $\underset{\sim}{H}_1$. If SPC are used $\underset{\sim}{T} = \tau\underset{\sim}{H}_1$ and rays from O_1 form a two parameter family parameterised by $\underset{\sim}{S}$. In the image space the equation of such a ray is

$$\bar{\underset{\sim}{y}} = \underset{\sim}{Y} + \bar{x}(\underset{\sim}{V}+\underset{\sim}{V}_B)\quad ,$$

where $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are given by (19.1). Thus, in the coordinate system Ⓒ'

$$\bar{y} = \bar{x}V_B+(y_a+\bar{x}v_a)S_y+(y_b+\bar{x}v_b)T_y+(s_{y1}+\bar{x}s_{v1})S_y^{\;2}+\ldots+(s_{y10}+\bar{x}s_{v10})T_z^{\;2}+0(3)\quad ,$$

$$\bar{z} = (z_a+\bar{x}w_a)S_z+(z_b+\bar{x}w_b)T_z+(s_{z3}+\bar{x}s_{w3})S_yS_z+\ldots+(s_{z7}+\bar{x}s_{w7})T_yT_z+0(3)\quad . \quad (20.1)$$

It is in general impossible to find $\bar{x}$ as a function of $\underset{\sim}{T}$ such that $\bar{x},\bar{y}$ and $\bar{z}$ are all independent of $\underset{\sim}{S}$. Consequently, imagery is in general astigmatic. Consider the special case of a meridional object and consider imagery by <u>meridional</u> rays from O_1. Then $S_z = T_z = 0$ and (20.1) becomes

$$\bar{y} = \bar{x}V_B+(y_b+\bar{x}v_b)T_y+(s_{y5}+\bar{x}s_{v5})T_y^2+\sum_{n=2}^{\infty}(\bar{a}_n+\bar{x}\bar{b}_n)T_y^n+[y_a+\bar{x}v_a+(s_{y2}+\bar{x}s_{v2})T_y+$$
$$+\sum_{n=2}^{\infty}(a_n+\bar{x}b_n)T_y^n]S_y+\sum_{n=2}^{\infty}f_n(\bar{x},T_y)S_y^n \quad ,$$

$$\bar{z} = 0 \quad , \tag{20.2}$$

where $\bar{a}_n,\bar{b}_n,a_n,b_n$ are the higher order analogues of s_{y5},s_{v5},s_{y2} and s_{v2} respectively and $f_n(\bar{x},T_y)$, the coefficient of S_y^n, is a linear function of $\bar{x}$ and a series in T_y. All rays from O_1 will intersect in a point provided $\bar{x}$ and $\bar{y}$ are independent of S_y. The coefficient of S_y gives

$$\bar{x} = -(y_a+s_{y2}T_y+\sum_{n=2}^{\infty}a_nT_y^n)/(v_a+s_{v2}T_y+\sum_{n=2}^{\infty}b_nT_y^n) \,. \tag{20.3}$$

The additional requirements

$$f_n(\bar{x},T_y) = 0 \quad , \quad n = 2,3,\ldots$$

will also yield equations of the form (20.3) for x. However, these are in general inconsistent with (20.3) but constitute conditions on the coefficients s_{y1},s_{v1}, etc., which must be satisfied in order to have stigmatic imagery. (20.2,3) thus read

$$\bar{x} = \sum_{n=0}^{\infty}x_nT_y^n \quad , \quad \bar{y} = \sum_{n=0}^{\infty}y_nT_y^n \quad ,$$

where x_0,y_0,x_1,y_1, etc., are combinations of the coefficients of (20.2).

The <u>dominant</u> terms of these series are

$$\bar{x} = -y_a/v_a - (s_{y2}v_a-s_{v2}y_a)t_y/v_a^2 \quad ,$$
$$\bar{y} = -(y_a/v_a)V_B - [(s_{y2}v_a-s_{v2}y_a)V_B+v_a(y|v)]t_y/v_a^2 \quad , \tag{20.4}$$

which give the <u>parabasal image</u> of O_1 formed by meridional rays. In order to refer this image to the coordinate system Ⓒ$_H$' the rotation

$$\bar{x}_H = \alpha_B\bar{x} + \beta_B\bar{y} \quad , \quad \bar{y}_H = -\beta_B\bar{x} + \alpha_B\bar{y} \quad , \quad \bar{z}_H = \bar{z} \quad , \tag{20.5}$$

must be performed. Thus

$$\bar{x}_H = -(y_a/\alpha_B v_a) - [s_{y2}v_a - s_{v2}y_a + \alpha_B{}^2 V_B v_a (y|v)]t_y/\alpha_B v_a{}^2 \quad .$$

If the coefficients y_a, v_a, s_{y2}, s_{v2}, etc., are replaced by their expressions in terms of y_{Ha}, v_{Ha}, etc., (Table 19/1), $\bar{x}_H$ becomes

$$\bar{x}_H = -y_{Ha}/v_{Ha} - (s_{Hy2}v_{Ha} - s_{Hv2}y_{Ha})t_y/v_{Ha}{}^2 \quad . \tag{20.6}$$

In a similar fashion it follows that

$$\bar{y}_H = -(y_H|v_H)t_y/v_H \quad . \tag{20.7}$$

Now suppose the image is determined directly in Ⓒ$_H$'. The basic equations used will be similar to those in the above analysis, with all coefficients bearing an additional subscript "H". The only difference is that in Ⓒ$_H$' $V_B = 0$. The resulting image will therefore be given by (20.4) provided V_B is set equal to zero and the coefficients y_a, v_a, etc., are replaced by the corresponding coefficients with the subscript H. It is then evident that these are identical with (20.6,7). Consequently it is verified that in the case of imagery by meridional rays the definition of <u>parabasal</u> appearing in §12 does give consistent results. If the object O_1 is allowed to run along the line Ⓛ normal to Ⓡ$_B$, it is evident from (20.6,7) that the image (by parabasal meridional rays) of Ⓛ is <u>not</u> necessarily normal to Ⓡ$_B$.

The image height $\underset{\sim}{H}$ is given by $\underset{\sim}{H} = \bar{\underset{\sim}{y}} - \bar{x}\underset{\sim}{V}_B$. Thus, for the meridional object

$$h_y = (y|v)t_y/v_a \quad .$$

Now $T_y = \tau H_{y1}$. Hence, replacing $(y|v)$ by (14.3) and τ by $-v_{a1}g$ according to (6.3), the parabasal image height is

$$h_y = (N_{y1}v_{a1}/N_y v_a)h_{y1} \quad , \quad h_z = 0 \quad .$$

The <u>meridional magnification</u> is defined to be

$$m_y = N_{y1}v_{a1}/N_y v_a \quad ,$$

and depends only on the parabasal coefficients.

(b) Another special case of interest is imagery of a sagittal object by sagittal rays. In this case $S_y = T_y = 0$ and (20.1) reduces to

$$\bar{y} = \bar{x}V_B + (s_{y8}+\bar{x}s_{v8})S_z^2 + (s_{y9}+\bar{x}s_{v9})S_z T_z + (s_{y10}+\bar{x}s_{v10})T_z^2 + 0(3) \quad ,$$

$$\bar{z} = (z_a+\bar{x}w_a)S_z + (z_b+\bar{x}w_b)T_z + 0(3) \quad . \qquad (20.10)$$

Consequently a stigmatic image is obtained with

$$\bar{x} = -z_a/w_a + \sum_{n=1}^{\infty} x_{2n}T_z^{2n} \quad , \quad \bar{y} = \bar{x}V_B + 0(2) \quad ,$$

$$\bar{z} = -(z|w)T_z/w_a + 0(3) \quad ,$$

provided certain conditions on the high order coefficients are satisfied. The dominant terms of this give as the parabasal image

$$\bar{x} = -z_a/w_a \quad , \quad \bar{y} = \bar{x}V_B \quad , \quad \bar{z} = -(z|w)t_z/w_a \quad . \qquad (20.11)$$

The image height is

$$h_y = 0 \quad , \quad h_z = (N_{z1}w_{a1}/N_z w_a)h_{z1} \quad , \qquad (20.12)$$

and the <u>sagittal magnification</u> is

$$m_z = (N_{z1}w_{a1}/N_z w_a) \quad . \qquad (20.13)$$

(20.11) shows that the image (by parabasal sagittal rays) of a line in Ⓢ but normal to Ⓜ is itself normal to Ⓜ.

(c) Finally, if O_1 is a basal object $\underset{\sim}{I} = 0$ and meridional and sagittal rays will respectively come to a focus at F_M and F_S on $\textcircled{R}_B$. According to (20.4) and (20.11) the abscissae of F_M and F_S are

$$\ell_M = -y_a/v_a \quad , \quad \ell_S = -z_a/w_a \quad , \tag{20.14}$$

respectively. Furthermore, F_M and F_S are distant

$$\underline{\underline{\ell}}_M = -y_a/v_a\alpha_B = -y_{Ha}/v_{Ha} \quad , \quad \underline{\underline{\ell}}_S = -z_a/w_a\alpha_B = -z_{Ha}/w_{Ha} \quad , \tag{20.15}$$

along $\textcircled{R}_B$ from the origin (i.e., from $\textcircled{P}_B$) and are the meridional and sagittal foci of §17.

21. Focal Lines and First Order Aberration

(a) With the exceptions of a few authors (see §21d) the first order congruence of rays from a point object has, after traversing a plane symmetric system, been represented in the literature by a family of straight lines intersecting two mutually perpendicular focal lines which are also perpendicular to some member of the congruence.[1] It will be seen below that this representation is too restrictive and leads to inconsistencies. In fact the two focal lines are not necessarily normal to the same member of the congruence. Since rays are the normals to a family of surfaces (i.e., the wave surfaces), the quoted result implies that all the normals to some elementary area of a surface pass through two straight lines. Maxwell[2] considered the form of those surfaces such that <u>all</u> their normals pass through one or two curves, the <u>focal curves</u> of the surface. He showed

---

1 See Synge (1937) §9; Luneburg (1964) §35.

2 Maxwell (1890) p.144.

that if the surface possessed two focal curves then these must be conic sections in a pair of mutually perpendicular planes. Moreover, if one of these is a straight line, the other must be a circle and the corresponding surface is a torus. Accordingly no wave surface can have two straight focal lines of finite or unlimited extent.

(b) As in §18, assume for the moment that the canonical variables are infinitesimal in the parabasal region. In the coordinate system Ⓒ', the aberration (7.7) in the plane $x = \ell$ is

$$\epsilon_y = (y_a + \ell v_a) s_y \quad , \quad \epsilon_z = (z_a + \ell w_a) s_z \quad . \qquad (21.1)$$

In the parabasal region it is usually possible to represent the entrance pupil Ⓔ by an ellipse. Accordingly, introduce polar coordinates (ρ, θ) such that

$$s_y = a\rho \cos \theta \quad , \quad s_z = \rho \sin \theta \quad ,$$

and represent vignetting by $\rho \leq \rho_{max}$ where ρ_{max} is a constant. The factor "a" takes account of the ellipticity of Ⓔ. The aberration (21.1) now becomes

$$\epsilon_y = a\rho(y_a + \ell v_a) \cos \theta \quad , \quad \epsilon_z = \rho(z_a + \ell w_a) \sin \theta \quad . \qquad (21.2)$$

The image is in general an elliptical patch of light which degenerates into a straight line when either $x = -y_a/v_a$ or $x = -z_a/w_a$. In the first case the line is in Ⓢ and passes through F_M and is usually called the meridional focal line; in the second case it is in Ⓜ, through F_S, and is usually called the sagittal focal line. These two focal lines are <u>normal to each other and to the x-axis</u>, but <u>not</u> necessarily to the base-ray. Suppose that the canonical variables are not infinitesimal, but all aberration coefficients of higher order than the first are zero up to some sufficiently high order. Then the equation (21.2) is valid over an extended neighbourhood of the base-ray. The system thus has two straight focal lines of some finite extent which, it has been seen (§21a), is impossible. Hence it is impossible simultaneously to reduce to zero all the high order aberration coefficients of a plane symmetric system. Control of the aberrations must therefore be gained through the balance of one aberration-type against another, or of one order against other orders.

In spite of all this, (21.2) is the parabasal aberration under the definition of §12. However, it does not follow that the "focal lines" derived from this represent the dominant characteristics of the focal curves, provided of course that they exist. (21.2) is the parabasal contribution to the total aberration. The existence of the "focal lines" is characteristic of astigmatism in symmetrical systems. For this reason the parabasal contribution to the aberration is said to be of the astigmatic type. In fact (see §49), it is <u>first order linear astigmatism</u>. Since (21.2) is the dominant term there exists some neighbourhood of the base-ray in which (21.2) does give the aberration of the system. This neighbourhood may be so small that (21.2) never usefully represents the aberration of the system (e.g., §76b). This is naturally the case when the system is already fairly well corrected (in which case F_M and F_S are virtually coincident) and the image plane is sufficiently close to the plane of best focus for parabasal imagery.

(c) If any focal curves exist for Ⓚ it is evident from symmetry that they must lie either in Ⓜ or in some plane normal to Ⓜ, in which case the focal curve must be symmetrical with respect to Ⓜ. A focal curve lying in Ⓜ will be given by the points of intersection of rays with Ⓜ. For the basal object O_B , $\underset{\sim}{T} = 0$ and (20.1) gives

$$\bar{y} = \bar{x}V_B + (y_a+\bar{x}v_a)S_y + (s_{y1}+\bar{x}s_{v1})S_y^2 + (s_{y8}+\bar{x}s_{v8})S_z^2 + 0(3) \quad ,$$

$$\bar{z} = [z_a+\bar{x}w_a+(s_{z3}+\bar{x}s_{w3})S_y+\sum_{n=2}^{\infty}(a_n+\bar{x}b_n)S_y^n]S_z \quad , \tag{21.3}$$

where a_n and b_n are series in S_z^2. Rays will intersect Ⓜ in those points for which $\bar{z} = 0$, i.e.,

$$\bar{x} = -[z_a+s_{z3}S_y+\sum_{n=2}^{\infty}a_nS_y^n]/[w_a+s_{wz}S_y+\sum_{n=2}^{\infty}b_nS_y^n]$$

$$= -z_a/w_a - (s_{z3}w_a-s_{w3}z_a)S_y/w_a^2 + 0(2) \quad . \tag{21.4}$$

Substitution of this into (21.3) yields

$$\bar{y} = -z_aV_B/w_a-[(z_av_a-y_aw_a)w_a+(s_{z3}w_a-s_{w3}z_a)V_B]S_y/w_a^2+\sum_{n=2}^{\infty}\sum_{\nu=0}^{n}f_{n\nu}S_y^{n-\nu}S_z^{\nu} \quad , \tag{21.5}$$

(<u>Note</u>: ν must be even.) where $f_{n\nu}$ is a combination of the coefficients in (21.3). Except when certain terms are zero, that is, when certain conditions between the coefficients are satisfied, (21.4,5) will not be the parametric equations of a curve. However, the dominant terms

$$\bar{x} = -z_a/w_a - (s_{z3}w_a-s_{w3}z_a)S_y/w_a^2 \quad ,$$

$$\bar{y} = -z_aV_B/w_a - [(z_av_a-y_aw_a)w_a+(s_{z3}w_a-s_{w3}z_a)V_B]S_y/w_a^2 \quad , \tag{21.6}$$

represents a straight line, the <u>sagittal focal line</u>, through the sagittal focus but inclined both to the x-axis and the base ray.

Suppose that this analysis is carried out in the coordinates $Ⓒ_H'$. The x-axis being the base-ray, the inclination of the sagittal focal line to the base-ray is θ_H where, from the equations corresponding to (21.6)

$$\tan\theta_H = \frac{w_{Ha}(z_{Ha}v_{Ha}-y_{Ha}w_{Ha})}{(s_{Hz3}w_{Ha}-s_{Hw3}z_{Ha})} \quad , \tag{21.7}$$

since $V_B = 0$ in $Ⓒ_H'$. Consider (21.6). The slope of this line is

$$m = V_B + \frac{(z_a v_a - y_a w_a) w_a}{(s_{z3} w_a - s_{w3} z_a)}$$

and the angle θ it makes with the base-ray is given by

$$\tan\theta = \frac{m - V_B}{1 + mV_B} = \frac{w_a(z_a v_a - y_a w_a)}{(s_{z3} w_a - s_{w3} z_a)/\alpha_B^2 + V_B w_a (z_a v_a - y_a w_a)} \quad .$$

By means of Table 19/1, the coefficients appearing in this may be expressed in terms of the corresponding coefficients in Ⓒ$_H$'. The right hand member of this is then seen to be identical with that of (21.7). Since the sagittal focal line passes through F_S, parabasal imagery as defined in §12, consistently determines the sagittal focal line, in contrast to the definitions based on infinitesimal variables. However, knowledge of certain second order coefficients is required.

If Ⓚ possesses a second focal line it must lie in the plane normal to Ⓜ and through the meridional focus. The plane Ⓗ of §7 is such a plane, provided $\ell = -y_a/v_a$. From (7.4) the points of intersection with Ⓗ of arbitrary rays from O_B are given by

$$H_y = [(s_{y1} v_a - s_{v1} y_a) S_y^2 + (s_{y8} v_a - s_{v8} y_a) S_z^2]/\alpha_H^* v_a (1 + V_H^* V_B) + 0(3) \quad ,$$

$$H_z = \frac{1}{v_a}\,[(z_a v_a - y_a w_a) + (s_{z3} v_a - s_{w3} y_a) S_y] S_z + 0(3) \quad . \qquad (21.8)$$

Irrespective of the inclination of Ⓗ, that is, of the value of V^*, this represents an area of Ⓗ unless certain conditions hold between the various coefficients. The dominant terms however give a focal line, <u>the meridional focal line</u>, through F_M and normal to Ⓜ. Parabasal optics thus gives two mutually perpendicular focal lines, only one of which need be normal to the base-ray. This is not in contradiction to the results of Maxwell since outside the parabasal region there need be no focal curves at all. In order that focal curves exist outside the parabasal region, (21.6,8) must each contain only a single parameter. Evidently this is only possible if certain conditions between the paracanonical coefficients of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are satisfied, in which case, the sagittal and meridional focal lines are the tangents to the focal curves at the points F_S and F_M where these cut the base-ray.

(d) That the sagittal focal line is inclined to the base-ray has not gone unnoticed in the literature. Herzberger,[3] Gullstrand, and Born and Wolf[4] were aware of this although other authors give the erroneous demonstration that it is normal to the base-ray. Born and Wolf give as an example of a focal line which is not normal to the base-ray one of the focal lines associated with the families of rays normal to an element of a torus. In this case the other focal line is of course normal to the base-ray and is the tangent to the focal curve, a circle. Herzberger shows that the foci and focal lines are related to the caustic surfaces of the system. The foci are the points of contact of the base-ray with the caustics, the sagittal caustic has a third order contact with Ⓜ in which it has a tangential line (the sagittal focal line) and the meridional caustic is tangential to the sagittal plane at F_M. The sagittal focal line is not to be confused with Conrady's characteristic focal line of a zonal aperture.[5]

---

3 Herzberger (1958) §23, in particular p.261,2.

4 Born and Wolf (1964) §4.6, p.171.

5 Conrady (1957) §55c.

IV. QUASI-INVARIANTS AND QUASI-LINEAR VARIABLES

22. Augmented Image Heights and Out-of-focus Terms

(a) In §7 the aberration was seen to be of the form

$$\underset{\sim}{\epsilon} = \underset{\sim}{H} - \underset{\sim}{f}(\underset{\sim}{H}_1)$$

where $\underset{\sim}{f}(\underset{\sim}{H}_1) = [f_y(\underset{\sim}{H}_1), f_z(\underset{\sim}{H}_1)]$ and f_y and f_z are distinct functions of $\underset{\sim}{H}_1$. It is convenient to regard $\underset{\sim}{H}$ as the fundamental quantity rather than $\underset{\sim}{\epsilon}$ and to determine the coefficients in the Taylor series for $\underset{\sim}{H}$ in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. The aberration coefficients are readily found from these once $\underset{\sim}{f}(\underset{\sim}{H}_1)$ is given. In the axial theory it was possible to express $\underset{\sim}{H}$ (and $\underset{\sim}{\epsilon}$) directly in terms of the quasi-invariants (M§7). However, in the basal theory, additional terms corresponding to out-of-focus terms are present.

In accordance with §7 the surface Ⓗ will be assumed to be plane and parallel to the x-plane. If Ⓗ has the equation $x = \ell$, the image height of any ray Ⓡ is given by (7.5). Let $Ⓡ_M$ and $Ⓡ_S$ be arbitrary meridional and sagittal parabasal rays respectively from the basal point of the object plane and let Ⓡ be any ray from a point object in the object plane $Ⓗ_1$. If the canonical variables, paracanonical coordinates, etc., of $Ⓡ_M$ and $Ⓡ_S$ are respectively distinguished by subscripts M and S, the abscisse of the foci F_M and F_S are

$$\ell_M = -y_M/v_M \quad , \quad \ell_S = -z_S/w_S \quad . \tag{22.1}$$

In general ℓ is not equal to either of these and it is convenient to define $\underset{\sim}{x}_M = (x_M, x_S)$ by

$$\ell = \ell_M - x_M = \ell_S - x_S \quad . \tag{22.2}$$

Introduce the augmenting factor $\underset{\sim}{\mu} = (\mu_y, \mu_z)$ where

$$\mu_y = -N_y v_M \quad , \quad \mu_z = -N_z w_S \quad ,^{1} \tag{22.3}$$

and define the <u>augmented</u> quantity $\tilde{Q}$ corresponding to Q by

$$\tilde{Q} = \mu Q \quad .$$

Thus the <u>augmented image height</u> is $\tilde{H}$:

$$\tilde{H}_y = -\mu_y x_M V + N_y(y_M V - v_M Y) \quad , \quad \tilde{H}_z = -\mu_z x_S W + N_z(z_S W - w_S Z) \quad ,$$

where use has been made of (22.1,2). If

$$F_y = -x_M V \quad , \quad G_y = N_y(y_M V - v_M Y) \quad ,$$

$$F_z = -x_S W \quad , \quad G_z = N_z(z_S W - w_S Z) \quad , \tag{22.4}$$

the augmented image height is given by

$$\tilde{H} = \tilde{F} + G \quad . \tag{22.5}$$

Suppose the image plane is the plane $Ⓗ_M$ through F_M with the equation $x = \ell_M$. Then $x_M = 0$ and $F_y = 0$. Similarly, if Ⓗ is the plane $Ⓗ_S$ with the equation $x = \ell_S$, $x_S = 0$ and $F_z = 0$. Thus image planes can be found so that each component of F vanishes. For this reason F is called the <u>out-of-focus term</u>. If first order astigmatism is absent, $x_M = x_S$ and consequently an image plane can be found so that $F = 0$. In general however this is not possible.

(b) If any quantity Q is linear in the canonical variables of any parabasal ray it is automatically linear in the corresponding paracanonical coordinates. This is formally expressed by writing (cf. (14.1))

---

[1] Note the minus sign. This ensures that for systems of positive focal length, $\tilde{Q}$ and Q have the same sign.

$$\underset{\sim}{Q} = Q_a \underset{\sim}{s} + Q_b \underset{\sim}{t} \quad . \tag{22.6}$$

Q_a and Q_b are respectively the <u>a- and b-components</u> of Q. Evidently both $\underset{\sim}{\mu}$ and $\underset{\sim}{G}$ may be resolved into a- and b-components:

$$\mu_y = -N_y(v_a s_M + v_b t_M) \quad , \quad G_y = G_{ya} s_M + G_{yb} t_M \quad ,$$

$$\mu_z = -N_z(w_a s_S + w_b t_S) \quad , \quad G_z = G_{za} s_S + G_{zb} t_S \quad .$$

Suppose SPC are used. $t_M = t_S = 0$ and thus

$$\mu_y = -N_y v_a s_M \quad , \quad G_y = G_{ya} s_M \quad ,$$

with similar expressions for the z-components of $\underset{\sim}{\mu}$ and $\underset{\sim}{G}$. If these are substituted into (22.5) it is evident that

$$N_y v_a H_y = N_y v_a F_y + G_{ya} \quad , \quad N_z w_a H_z = N_z w_a F_z + G_{za} \quad .$$

Hence, when SPC are used it is sufficient to define the augmenting factor as

$$\underset{\sim}{\mu} = -(N_y v_a \, , \, N_z w_a) \quad . \tag{22.7}$$

Then

$$\tilde{\underset{\sim}{H}} = \tilde{\underset{\sim}{F}} + \underset{\sim}{G}_a \quad , \tag{22.8}$$

which may be rewritten so as to give $\underset{\sim}{H}$ directly:

$$\underset{\sim}{H} = \underset{\sim}{F} + \underset{\sim}{G}_a/\underset{\sim}{\mu} \quad . \tag{22.9}$$

The b-component of $\underset{\sim}{G}$ is apparently not required. However, in order to determine $\underset{\sim}{G}_a$ correct to the nth order, say, $\underset{\sim}{G}_b$ must be known correct to the (n-1)th order. This becomes apparent in §37 where iteration is discussed. Further, $\underset{\sim}{G}_b$ is required to determine $\underset{\sim}{V}$ and hence $\underset{\sim}{F}$ (see §25). Consequently, in order to determine $\underset{\sim}{H}$ correct to the nth order, $\underset{\sim}{G}_a$ and $\underset{\sim}{G}_b$ are both required correct to the nth order. This is a result of the presence of first order astigmatism in the parabasal region. (cf. M§13e.)

23. Quasi-invariants and their Geometrical Interpretation

In the parabasal region $\underset{\sim}{G}$, as defined by (22.4), reduces to $\underset{\sim}{g}$:

$$g_y = N_y(y_M v - v_M y) \quad , \quad g_z = N_z(z_S w - w_S z) \quad .$$

From the definitions of y_M, v_M, z_S and w_S it is evident that $\underset{\sim}{g}$ is the Lagrange invariant defined in §16b. $\underset{\sim}{G}$ is thus a quasi-invariant and is referred to as <u>the quasi-invariant</u>. In general, then $\Delta\underset{\sim}{G} = 0(2)$. However, $\nabla\underset{\sim}{G} = 0$. Suppose the image plane is the plane Ⓗ$_M$. Then, for any ray Ⓡ, $\tilde{H}_y = G_y$ and G_y is the y-component of the augmented image height in Ⓗ$_M$. Similarly, if the image plane is Ⓗ$_S$, $\tilde{H}_z = G_z$ and G_z is the z-component of the augmented image height in Ⓗ$_S$.

24. Paracanonical Coefficients. Surface Increments to $\underset{\sim}{H}$.

(a) Let $\underset{\sim}{Q}$ be any proper two-vector which is a function of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Formally, $\underset{\sim}{Q}$ may be expanded as a Taylor series in $\underset{\sim}{S}$ and $\underset{\sim}{T}$:

$$\underset{\sim}{Q} = \underset{\sim}{Q}_0 + \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{\sim}{q}_{\mu\nu\tau}{}^{(n)} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} \quad , \tag{24.1}$$

where the summation $\sum_{\mu\nu\tau}^{n} \equiv \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \sum_{\tau=0}^{\nu}$. The expansion (24.1) is said to be the <u>paracanonical expansion</u> of $\underset{\sim}{Q}$ and the coefficients $\underset{\sim}{q}_{\mu\nu\tau}{}^{(n)} = (q_{y\mu\nu\tau}{}^{(n)}, q_{z\mu\nu\tau}{}^{(n)})$ are collectively the <u>nth order (paracanonical) $\underset{\sim}{q}$-coefficients</u>. In particular, $q_{y\mu\nu\tau}{}^{(n)}$ and $q_{z\mu\nu\tau}{}^{(n)}$ are the q_y- and q_z-coefficients respectively. In view of the symmetry properties of $\underset{\sim}{Q}$, $Q_{z0} = 0$ and ν must be even in Q_y and odd in Q_z. Let $q^{(n)}$ be the sum of all <u>n</u>th order terms in (24.1). In the parabasal region

$$\underset{\sim}{q} = \underset{\sim}{Q}_0 + \underset{\sim}{q}^{(1)} \quad .$$

Then

$$\underset{\sim}{Q} = \underset{\sim}{q} + \sum_{n=2}^{\infty} \underset{\sim}{q}^{(n)} = \underset{\sim}{q} + \underset{\sim}{Q}^{\dagger} \quad , \tag{24.2}$$

where $\underset{\sim}{Q}^{\dagger}$ are the second and higher order terms of $\underset{\sim}{Q}$. The notation needs to be broadened when Q_j naturally appears in the theory in the form of a sum over surfaces:

$$\underset{\sim}{Q}_j = \sum_{i=1}^{j-1} \Delta \underset{\sim}{Q}_i \quad , \quad \underset{\sim}{Q}_j{}' = \sum_{i=1}^{j} \Delta \underset{\sim}{Q}_i \quad .$$

In this case the coefficients of the expansion of $\Delta \underset{\sim}{Q}_i$ will be denoted by lower case letters, e.g., $\underset{\sim}{q}_{\mu\nu\tau i}{}^{(n)}$, with the corresponding upper case letters reserved for $\sum_i \Delta \underset{\sim}{Q}_i$. Thus

$$\underset{\sim}{Q}_{\mu\nu\tau j}{}^{(n)} = \sum_{i=1}^{j-1} \underset{\sim}{q}_{\mu\nu\tau i}{}^{(n)} \quad , \quad \underset{\sim}{Q}_{\mu\nu\tau j}{}^{(n)\prime} = \sum_{i=1}^{j} \underset{\sim}{q}_{\mu\nu\tau i}{}^{(n)} \quad . \tag{24.3}$$

The $\underset{\sim}{q}_{\mu\nu\tau i}{}^{(n)}$ are the <u>contributions by the ith surface</u> to the <u>intermediate coefficients</u> $\underset{\sim}{Q}_{\mu\nu\tau j}{}^{(n)}$. In the image space of Ⓚ, the $\underset{\sim}{Q}_{\mu\nu\tau k}{}^{(n)\prime}$ are the <u>final</u> $\underset{\sim}{Q}$-<u>coefficients</u>. Moreover, if $\underset{\sim}{Q}$ or $\Delta\underset{\sim}{Q}$ permit a decomposition into $\underline{a}$- and $\underline{b}$-components, so do the $\underset{\sim}{q}$- or $\underset{\sim}{Q}$-coefficients. For example, for ΔG_y,

$$g_{y\mu\nu\tau}{}^{(n)} = g_{y\mu\nu\tau a}{}^{(n)} s_{yM} + g_{y\mu\nu\tau b}{}^{(n)} t_{yM} \quad . \tag{24.4}$$

The number of coefficients of any order appearing in (24.1) can be readily evaluated. From the definition of the coefficients it is clear that $n \geqq \mu \geqq \nu \geqq \tau$. Taking into account that in Q_y and Q_z ν is even and odd respectively, it follows that the numbers $\underline{\underline{N}}_{yn}$ of $\underline{n}$th order q_y-coefficients, and $\underline{\underline{N}}_{zn}$ of $\underline{n}$th order q_z-coefficients are given by

$$\underline{\underline{N}}_{yn} = \frac{1}{12}(n+2)(n^2+4n+6) \quad , \quad \underline{\underline{N}}_{zn} = \frac{1}{12} n(n+2)(n+4) \qquad \text{(n even)}$$

$$\underline{\underline{N}}_{yn} = \underline{\underline{N}}_{zn} = \frac{1}{12}(n+1)(n+2)(n+3) \quad . \qquad \text{(n odd)} \tag{24.5}$$

(b) The image height is by definition the difference between the coordinates of two points and as such is invariant under translations.

Thus $\nabla \underset{\sim}{H} = 0$. Moreover, $\nabla \underset{\sim}{\mu} = 0$ since $\nabla N = 0$, $\nabla \underset{\sim}{V} = 0$. Thus

$$\nabla \tilde{\underset{\sim}{H}} = 0 \quad . \tag{24.6}$$

However, in general the refraction increments are not zero. In fact

$$\Delta \tilde{\underset{\sim}{H}} = \Delta \tilde{\underset{\sim}{F}} + \Delta \underset{\sim}{G} \quad ,$$

where $\Delta \tilde{\underset{\sim}{F}}$ is of the first order, and $\Delta \underset{\sim}{G}$ of the second. At the first surface $\underset{\sim}{F}_1 = 0$. Write $\underset{\sim}{F} = \underset{\sim}{F}^{\dagger} + \underset{\sim}{f}$ and sum $\Delta \tilde{\underset{\sim}{H}}$ over the surfaces. Thus

$$\tilde{\underset{\sim}{H}}_j = \underset{\sim}{G}_1 + \tilde{\underset{\sim}{f}}_j + \sum_{i=1}^{j-1} (\Delta \tilde{\underset{\sim}{F}}_i{}^{\dagger} + \underset{\sim}{G}_i)$$

in view of (24.6). Now $\underset{\sim}{G}_1 = \underset{\sim}{g}_1 = \underset{\sim}{g}_j$. Thus

$$\tilde{\underset{\sim}{H}}_j = \tilde{\underset{\sim}{h}}_j + \sum_{i=1}^{j-1} (\Delta \tilde{\underset{\sim}{F}}_i{}^{\dagger} + \underset{\sim}{G}_i) \quad , \tag{24.7}$$

expresses the augmented image height as a parabasal term plus second order contributions from the surfaces. $\tilde{\underset{\sim}{H}}_j{}'$ is obtained by extending the summation to $\underline{j}$.

For a symmetric system pairs of conjugate points and planes can be readily found and in general object and image planes are taken to be conjugate. When the system is not symmetric, pairs of conjugate points for any two rays may be defined in a general manner.[1] However, the point conjugate to a given fixed point depends on the rays chosen and in general conjugate planes or surfaces cannot be defined. It is evident, therefore, that given any pair of object and image planes for Ⓚ as a whole, the choice of image planes in other media of Ⓚ is arbitrary. Consequently there is little point in considering contributions of the form $\Delta \tilde{\underset{\sim}{F}} = \tilde{F}' - \tilde{F}$ to $\underset{\sim}{F}$. In the absence of first order astigmatism Ⓗ will be initially

---

[1] Smith (1930).

chosen such that parabasal imagery is stigmatic. In this case $\underset{\sim}{F} = 0$ and the contribution by the $\underline{i}$th surface to $\tilde{\underset{\sim}{H}}^{\dagger}$ (i.e., the contribution to the augmented aberration) is simply $\Delta\underset{\sim}{G}_i$. In the presence of first order astigmatism the term $\tilde{\underset{\sim}{F}}$ must be taken into account. Under some circumstances, for example, when $\underset{\sim}{x}_M$ is small, the effect of the $\underline{i}$th surface may still be taken as the contribution $\Delta\underset{\sim}{G}_i$. This does not mean, however, that the out-of-focus term may be neglected in the final aberration since this is to be exact to the $\underline{n}$th order, that is, all terms of order less than or equal to n <u>must</u> be included. In general the effect of $\tilde{\underset{\sim}{F}}^{\dagger}$ will need to be taken into account when analyzing the surface contribution to $\underset{\sim}{H}$.

$\underset{\sim}{x}_M$ is chosen by the designer so as to give the best overall image. Consequently the only part of $\tilde{\underset{\sim}{F}}$ which depends on rays traversing Ⓚ is V. Since the surfaces affect $\tilde{\underset{\sim}{F}}$ only through $\underset{\sim}{V}$ it seems appropriate to consider the contribution by the $\underline{i}$th surface to $\tilde{\underset{\sim}{F}}_j{}^{\dagger}$ to be of the form

$$\Delta\tilde{\underset{\sim}{F}}_i{}^{\dagger} = -\mu_j \underset{\sim}{x}_{Mj} \Delta\underset{\sim}{V}_i{}^{\dagger} \quad .$$

$\underset{\sim}{H}_j$ may now be written

$$\underset{\sim}{H}_j = \underset{\sim}{h}_j - \underset{\sim}{x}_{Mj} \sum_{i=1}^{j-1} \Delta\underset{\sim}{V}_i{}^{\dagger} + \frac{1}{\mu_j} \sum_{i=1}^{j-1} \Delta\underset{\sim}{G}_{ai} \quad , \tag{24.8}$$

and the contribution by the $\underline{i}$th surface to $\underset{\sim}{H}_j{}^{\dagger}$ is <u>exactly</u>

$$\Delta\underset{\sim}{H}_j{}^{\dagger} = -\underset{\sim}{x}_{Mj} \Delta\underset{\sim}{V}_i{}^{\dagger} + \frac{1}{\mu_j} \Delta\underset{\sim}{G}_{ai} \quad . \tag{24.9}$$

The first order terms of $\underset{\sim}{H}_j$ are

$$\underset{\sim}{h}_j = (\underset{\sim}{y}_a + \ell \underset{\sim}{v}_a)\underset{\sim}{s} + (\underset{\sim}{y}_b + \ell \underset{\sim}{v}_b)\underset{\sim}{t} = -\underset{\sim}{x}_M \underset{\sim}{v}_a \underset{\sim}{s} + \underset{\sim}{m} H_1 \quad , \tag{24.10}$$

where the magnification $\underset{\sim}{m} = \tau(\underset{\sim}{y}_b + \ell \underset{\sim}{v}_b)$.

(c) $\underset{\sim}{H}$, $\underset{\sim}{V}$ and $\Delta\underset{\sim}{G}$ are all proper two-vectors and have expansions of the form (24.1). If these are substituted into (24.8) it is seen that the $\underset{\sim}{h}$-, or image height coefficients, are

$$\underset{\sim}{h}_{\mu\nu\tau k}^{(n)'} = \frac{1}{\underset{\sim}{\mu}_k'}\underset{\sim}{G}_{\mu\nu\tau ak}^{(n)'} - \underset{\sim}{x}_{Mk}\underset{\sim}{v}_{\mu\nu\tau k}^{(n)'} \quad . \tag{24.11}$$

From (24.9) it is evident that the contributions by the $\underline{i}$th surface to these are

$$\Delta\underset{\sim}{h}_{\mu\nu\tau i}^{(n)} = \frac{1}{\mu_k'}\underset{\sim}{G}_{\mu\nu\tau ai}^{(n)} - \underset{\sim}{x}_{Mk}\Delta\underset{\sim}{v}_{\mu\nu\tau i}^{(n)} \quad , \tag{24.12}$$

(Note that the corresponding intermediate coefficients at the $\underline{j}$th surface do not represent image height coefficients at the $\underline{j}$th surface since $\underline{\mu}$ and $\underset{\sim}{x}_M$ are relevant to the image space of the last surface.) If the general definition of $\underset{\sim}{\epsilon}$ is used, the function $\underset{\sim}{f}(\underset{\sim}{H}_1)$ may be expressed as a series in $\underset{\sim}{T}$ and subtracted from the series for $\underset{\sim}{H}$. The aberration coefficients $\underset{\sim}{E}_{\mu\nu\tau k}^{(n)'}$ are then identical with (24.11) except for those of distortion - that is, the coefficients $\underset{\sim}{E}_{n\mu\mu k}^{(n)'}$ which multiply $T_y^{\,n-\mu}T_z^{\,\mu}$.

25. Quasi-linear Variables

Since $\nabla\underset{\sim}{G} = 0$, $\underset{\sim}{G}_j$ is given by

$$\underset{\sim}{G}_j = \underset{\sim}{G}_1 + \sum_{i=1}^{j-1}\Delta\underset{\sim}{G}_i \quad . \tag{25.1}$$

Consider $\underset{\sim}{G}_1$. The canonical coordinates in (22.4) may be replaced by the paracanonical coordinates (not necessarily SPC) according to (6.3). Then

$$G_{y1} = (N_{y1}/g)(s_{yM}T_y - t_{yM}S_y) \quad , \quad G_{z1} = (N_{z1}/g)(s_{zS}T_z - t_{zS}S_z) \quad . \tag{25.2}$$

Moreover,

$$G_y = G_{ya}s_{yM} + G_{yb}t_{yM} \quad , \quad G_z = G_{za}s_{zS} + G_{zb}t_{zS} \quad ,$$

where

$$G_{ya} = N_y(y_aV - v_aY) \quad , \quad G_{za} = N_z(z_aW - w_aZ) \quad ,$$

$$G_{yb} = N_y(y_bV - v_bY) \quad , \quad G_{zb} = N_z(z_bW - w_bZ) \quad . \tag{25.3}$$

Since $\circledR_M$ and $\circledR_S$ are arbitrary, (25.1-3) must in particular be valid for the a- and b-rays. Consequently, by taking $\circledR_M$ first as the meridional a-ray from O_B and then as the meridional b-ray and substituting (25.2,3) into (25.1), the following pair of equations is obtained:

$$N_{yj}(y_{aj}V_j - v_{aj}Y_j) = (N_{y1}/g_y)T_y + \sum_{i=1}^{j-1}\Delta G_{yai} \quad ,$$

$$N_{yj}(y_{bj}V_j - v_{bj}Y_j) = -(N_{y1}/g_y)S_y + \sum_{i=1}^{j-1}\Delta G_{ybi} \quad , \qquad (25.4)$$

These may be solved for Y_j and V_j and, using the identity (14.3),

$$Y_j = y_{aj}(S_y+\delta_{yj}) + y_{bj}(T_y+\delta_{vj}) \quad , \quad V_j = v_{aj}(S_y+\delta_y) + v_{bj}(T_y+\delta_v) , (25.5)$$

where

$$\delta_{yj} = -(g/N_{y1})\sum_{i=1}^{j-1}\Delta G_{yb} \quad , \quad \delta_{vj} = (g/N_{y1})\sum_{i=1}^{j-1}\Delta G_{ya} \quad .$$

(cf. M§9) An analagous pair of equations which may be solved for Z and W are obtained from the z-component of (25.1). The solutions of these, along with (25.5), are conveniently written as

$$\underset{\sim}{Y}_j = \underset{\sim}{y}_{aj}(\underset{\sim}{S}+\underset{\sim}{\delta}_{yj}) + \underset{\sim}{y}_{bj}(\underset{\sim}{T}+\underset{\sim}{\delta}_{vj}) \quad , \quad \underset{\sim}{V}_j = \underset{\sim}{v}_{aj}(\underset{\sim}{S}+\underset{\sim}{\delta}_{yj}) + \underset{\sim}{v}_{bj}(\underset{\sim}{T}+\underset{\sim}{\delta}_{vj}) \ , (25.6)$$

where

$$\underset{\sim}{\delta}_{yj} = (\delta_{yj},\delta_{zj}) = -\sum_{i=1}^{j-1}{}^{\backprime}\Delta\underset{\sim}{G}_b \quad , \quad \underset{\sim}{\delta}_{vj} = (\delta_{vj},\delta_{wj}) = \sum_{i=1}^{j-1}{}^{\backprime}\Delta\underset{\sim}{G}_a \quad . \ (25.7)$$

(Note: If the $\underset{\sim}{\delta}$'s are primed, the summation extends to j and if j = 1, the $\underset{\sim}{\delta}$'s are zero.) The convention applying to the use of the ante-prime "ˋ" is as follows:

$${}^{\backprime}Q_y = (g/N_{y1})Q_y \quad , \quad {}^{\backprime}Q_z = (g/N_{z1})Q_z$$

for any quantity $\underset{\sim}{Q}$ or coefficient of $\underset{\sim}{Q}$. (cf. M§12b.) (25.6,7) exhibit $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ as quasi-linear variables in the sense that in the parabasal region, $\underset{\sim}{\delta}_y = \underset{\sim}{\delta}_v = 0$ and (14.2) is recovered. Since $\underset{\sim}{V}$ occurs in $\underset{\sim}{F}$, knowledge of $\Delta\underset{\sim}{G}$ at each surface completely determines $\underset{\sim}{H}$, $\underset{\sim}{Y}$, $\underset{\sim}{V}$ etc. In other words the behaviour of Ⓚ is uniquely specified by the $\underset{\sim}{g}$-coefficients. At each surface there are

$$\underline{\underline{N}}_{gn} = 2(\underline{\underline{N}}_{yn}+\underline{\underline{N}}_{zn}) = \tfrac{1}{3}(n+1)(n+2)(n+3) \tag{25.8}$$

$\underline{n}$th order $\underset{\sim}{g}$- (or $\underset{\sim}{G}$-) coefficients, including $\underline{a}$-, $\underline{b}$-, $\underline{y}$- and $\underline{z}$-components.

26. $\Delta\underset{\sim}{G}$ as a Function of $\underset{\sim}{Y}$, $\underset{\sim}{V}$

(a) The $\underset{\sim}{g}$-coefficients are the coefficients in the Taylor series for $\Delta\underset{\sim}{G}$ in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. In order to determine these, $\Delta\underset{\sim}{G}$ must first be expressed in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. This is accomplished by expressing $\Delta\underset{\sim}{G}$ in terms of $\Delta\underset{\sim}{Y}$ and $\Delta\underset{\sim}{V}$ and hence in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ through (10.12). Consider ΔG_y:

$$\begin{aligned}\Delta G_y &= \Delta N_y(y_M V - v_M Y)\\ &= N_y'(y_M'\Delta V - v_M'\Delta Y) + V\Delta N_y y_M - Y\Delta N_y v_M\\ &= -N_y'\bar{R}_B[v_M'P + y_M'(R-1)](I_y+I_{By}) - (N_y'c\bar{R}_B y_M' + \Delta N_y v_M)Y + (\Delta N_y y_M - N_y' n_{Bx}\bar{R}_B y_M')V\\ &\equiv -N_y'\bar{R}_B[v_M'P + y_M'(R-1)](I_y+I_{By}) + a_y Y + b_y V \quad ,\end{aligned} \tag{26.1}$$

where the identity defines a_y and b_y. The corresponding expression for ΔG_z can be obtained from the above by replacing y- and v- components of two-vectors by the corresponding z- and w-components, and v_M', y_M' by w_S', z_S' . Thus

$$a_z = -\Delta N_z w_S - N_z'c\bar{R}_B z_S' = N_z'c\bar{R}_B z_S' - N_z'c\bar{R}_B z_S' = 0 \quad ,$$

and in a similar fashion $b_z = 0$. Consequently:

$$\Delta G_z = -N_z'\bar{R}_B[w_S'P + z_S'(R-1)]I_z \quad . \tag{26.2}$$

Since $\underset{\sim}{G}$ is a quasi-invariant, $\Delta\underset{\sim}{G}$ should be of the second order. X, and hence P, is of the first order, and R = 1 for the base-ray. Thus ΔG_z as given by (26.2) is certainly of the second order. However, in (26.1) ΔG_y is not explicitly of the second order. It is important to rewrite (26.1) as a second order expression.

(b) If $\underset{\sim}{Q}$ is any proper two-vector function of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, it may be formally expanded as a power series in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$:

$$\underset{\sim}{Q} = \underset{\sim}{Q}_0 + \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \hat{\underset{\sim}{q}}_{\mu\nu\tau}{}^{(n)} Y^{(n-\mu)} V^{(\mu-\nu)} Z^{(\nu-\tau)} W^{\tau} \quad , \tag{26.3}$$

(cf. (24.1)) where $\hat{\underset{\sim}{q}}_{\mu\nu\tau}{}^{(n)} = (\hat{q}_{y\mu\nu\tau}{}^{(n)}, \hat{q}_{z\mu\nu\tau}{}^{(n)})$. Just as in (24.1) ν is even in Q_y and odd in Q_z and $Q_{z0} = 0$. The expansion (26.3) is called the <u>surface expansion</u> of Q. $\hat{\underset{\sim}{q}}_{\mu\nu\tau}{}^{(n)}$ are the $\hat{\underset{\sim}{q}}$- or <u>surface coefficients</u> of $\underset{\sim}{Q}$ and are equal in number to the corresponding $\underset{\sim}{q}$-coefficients. $\hat{q}^{(n)}$ denotes the sum of all terms of the <u>n</u>th order in (26.3). Thus, in the parabasal region,

$$\underset{\sim}{q} = \underset{\sim}{Q}_0 + \hat{\underset{\sim}{q}}^{(1)} \quad .$$

Define

$$\underset{\sim}{Q}^* = \underset{\sim}{Q} - \underset{\sim}{q} = \sum_{n=2}^{\infty} \underset{\sim}{q}^{(n)} \quad . \tag{26.4}$$

In the above $\underset{\sim}{q}$ is a function of Y and V whereas in (24.2) it was a function of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. $\underset{\sim}{q}$ always represents the dominant or parabasal terms of $\underset{\sim}{Q}$ irrespective of the variables used. If $\underset{\sim}{q}$ in (26.4) is expressed as a function of $\underset{\sim}{S}$ and $\underset{\sim}{T}$, terms of the second and higher orders will arise due to the $\underset{\sim}{\delta}$'s in (25.6). <u>Thus $\underset{\sim}{Q}^{\dagger}$ is the sum of $\underset{\sim}{Q}^*$ and these terms</u>. Similarly, although the $\hat{\underset{\sim}{q}}$- and $\hat{q}^*$-coefficients (i.e., the coefficients in the surface expansion of $\underset{\sim}{Q}^*$) are identical, the q- and q*-coefficients (i.e., the coefficients in a paracanonical expansion of $\underset{\sim}{Q}^*$) are not.

These notations are carried over to scalar quantities Q, that is, quantities which are invariant under reflection in Ⓜ, for example P and R. Finally it is convenient to define a two-vector $\underset{\sim}{Q}$ corresponding to the scalar Q by

$$\underset{\sim}{Q} = (Q-Q_0)\underset{\sim}{I} + Q^*\underset{\sim}{I}_B \quad , \tag{26.5}$$

which is of the second order in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$.

(c) Accordingly, write

$$P = p + P^* \quad , \quad R = r + R^* \quad , \tag{26.6}$$

and substitute these into (26.1):

$$\Delta G_y = -N_y'\bar{\bar{R}}_B\{[v_M'P+y_M'(R-1)]I_y+[v_M'P^*+y_M'R^*]I_{By}\} - N_y'\bar{\bar{R}}_B[v_M'p+y_M'(r-1)]I_{By} +$$
$$+ a_yY + b_yV \quad .$$

The term in the braces is obviously of the second order and the remainder of the first order. Thus, since G_y is a quasi-invariant

$$N_y'\bar{\bar{R}}_B[v_M'p+y_M'(r-1)]I_{By} - a_yY - b_yV \equiv 0 \quad , \tag{26.7}$$

and this is verified in §34. Thus

$$\Delta G_y = -N_y'\bar{\bar{R}}_B(v_M'P_y+y_M'R_y) \quad , \quad \Delta G_z = -N_z'\bar{\bar{R}}_B(w_S'P_z+z_S'R_z) \quad ,$$

where the last equation follows immediately from (26.2) and P_y, P_z, R_y and R_z are defined by (26.5). Define

$$\bar{\underset{\sim}{y}} = -\bar{\bar{R}}_B(N_y'y_M', N_z'z_S') \quad , \quad \bar{\underset{\sim}{v}} = -\bar{\bar{R}}_B(N_y'v_M', N_z'w_S') \quad , \tag{26.8}$$

in which case the pair of equations following (26.7) may be written

$$\Delta\underset{\sim}{G} = \bar{\underset{\sim}{v}}\underset{\sim}{P} + \bar{\underset{\sim}{y}}\underset{\sim}{R} \quad . \tag{26.9}$$

The $\underline{a}$- and $\underline{b}$-components of $\Delta\underset{\sim}{G}$ may be read off directly from (26.9). For example

$$\Delta G_{ya} = \bar{v}_{ya} P_y + \bar{y}_{ya} R_y \quad , \tag{26.10}$$

where

$$\bar{v}_{ya} = -N_y{}' v_a{}' \bar{R}_B \quad , \quad \bar{y}_{ya} = -N_y{}' y_a{}' \bar{R}_B \quad . \tag{26.11}$$

27. Quadratic Equations for R and X

(a) The point of intersection Ⓟ of any ray with Ⓕ has the coordinates $\vec{y} = (X,y,z)$ where $\vec{y}$ satisfies (10.1). Now $\underset{\sim}{y} = \underset{\sim}{Y} + X(\underset{\sim}{V}+\underset{\sim}{V}_B)$ and substitution of this into (10.1) yields

$$(cX-n_{Bx})^2 + [c\underset{\sim}{Y}+cX(\underset{\sim}{V}+\underset{\sim}{V}_B)-\underset{\sim}{n}_B]^2 = 1 \quad .$$

Thus, X is given in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ by the quadratic equation

$$\rho_2 X^2 + 2\rho_1 X + \rho_0 = 0 \quad , \tag{27.1}$$

where

$$\rho_2 = c^2(1/\alpha_B{}^2+2\underset{\sim}{V}_B\cdot\underset{\sim}{V}+\underset{\sim}{V}\cdot\underset{\sim}{V}) \quad , \quad \rho_1 = c[\lambda_1{}^{-1}+\underset{\sim}{n}_B\cdot V-c\underset{\sim}{Y}\cdot(\underset{\sim}{V}+\underset{\sim}{V}_B)] \quad ,$$

$$\rho_0 = c^2\underset{\sim}{Y}\cdot\underset{\sim}{Y} - 2c\underset{\sim}{n}_B\cdot\underset{\sim}{Y} \quad . \tag{27.2}$$

(b) If $\underset{\sim}{I}_A = \underset{\sim}{I} + \underset{\sim}{I}_B$ and $\underset{\sim}{V}_A = \underset{\sim}{V} + \underset{\sim}{V}_B$ are substituted into the equation (9.7,8) for $\bar{R}_A$, the corresponding equation for R is found to be

$$\tau R^2 + 2\sigma R + \mu = 0 \tag{27.3}$$

where

$$\mu = (1-1/k^2)(1/\alpha_B{}^2+2\underset{\sim}{V}_B\cdot\underset{\sim}{V}+\underset{\sim}{V}\cdot\underset{\sim}{V}) \quad ,$$

$$\sigma = -\bar{R}_B[\underset{\sim}{I}_B\cdot\underset{\sim}{V}_B+\underset{\sim}{I}_B\cdot\underset{\sim}{V}+\underset{\sim}{I}\cdot\underset{\sim}{V}_B+\underset{\sim}{I}\cdot\underset{\sim}{V}+n_x\mu/(1-k^2)] \quad ,$$

$$\tau = \bar{R}_B{}^2[\underset{\sim}{I}_B\cdot\underset{\sim}{I}_B+2\underset{\sim}{I}_B\cdot\underset{\sim}{I}+\underset{\sim}{I}\cdot\underset{\sim}{I}+n_x{}^2\mu/(1-k^2)] \quad . \tag{27.4}$$

Here

$$n_x = n_{Bx} - cX \quad . \tag{27.5}$$

(Note: (27.3,4) are in fact valid for any surface, provided the appropriate expression for n_x is used in place of (27.5).)

V. SURFACE EXPANSIONS

28. On the Determination of the g-coefficients

(a) Equations (27.1-4) may be solved systematically for X and R as series in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ (see §30). Since P = XR, both P and R are, in principle, known as surface expansions and by substitution of these into (26.9) so is $\Delta\underset{\sim}{G}$. However, $\Delta\underset{\sim}{G}$ is ultimately required as a series in the paracanonical coordinates. $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are expressed in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ through (25.6) and if these are substituted into the surface expansion for $\Delta\underset{\sim}{G}$, the required series in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ is obtained. $\underset{\sim}{Y}_j$ and $\underset{\sim}{V}_j$ themselves involve the $\Delta\underset{\sim}{G}_i$ as a sum over surfaces. However, the summation is from $\underline{i}$ = 1 to $\underline{j}$-1 and consequently, $\Delta\underset{\sim}{G}_j$ is itself not required. Thus if $\Delta\underset{\sim}{G}_i$ is known at each of the first ($\underline{j}$-1) surfaces of Ⓚ as a series in $\underset{\sim}{S}$ and $\underset{\sim}{T}$, $\Delta\underset{\sim}{G}_j$ can be obtained as a similar series. Moreover, at the first surface $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$ are linear in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ (see (6.3)) and it is a simple matter to express $\Delta\underset{\sim}{G}_1$ in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Thus at all surfaces $\Delta\underset{\sim}{G}$ may, in principle, be determined as a Taylor series in $\underset{\sim}{S}$ and $\underset{\sim}{T}$.

$\Delta\underset{\sim}{G}$ is of the second order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ and when the series (25.6) is substituted in place of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, terms of the latter which are of the $\underline{n}$th order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ will not contribute to the $\underline{n}$th order terms of $\Delta\underset{\sim}{G}$. Consequently, if the terms of $\Delta\underset{\sim}{G}$ of order less than $\underline{n}$ are known, those of order $\underline{n}$ can be found. Since the parabasal coefficients are known it is possible to find the second order terms of $\Delta\underset{\sim}{G}$ at all surfaces. It is evident from the above discussion that the g-coefficients can be found by either of two equivalent procedures: on the one hand all the g-coefficients up to the required order may be computed at the first surface. From these the corresponding coefficients at the second surface may be determined, which in turn are used to determine the coefficients at the third surface, and so on. On the other hand, given the parabasal coefficients at all surfaces, the second order $\underset{\sim}{g}$-coefficients may be determined at all surfaces. These in turn render it possible to compute the third order coefficients, and so on. In the actual computation of the coefficients the first of these methods was employed.

(b) In order to determine both the $\underline{a}$- and $\underline{b}$-components of $\Delta\underset{\sim}{G}$, four surface expansions, that is, two for each of $\Delta\underset{\sim}{G}_a$ and $\Delta\underset{\sim}{G}_b$, must be converted into the corresponding paracanonical expansions. It is more convenient for the purposes of developing the theory, and the subsequent numerical calculations are somewhat shorter, if the paracanonical expansions of P and R are determined in lieu of the surface expansions of $\Delta\underset{\sim}{G}$. This may also be accomplished by an iterative process and has the advantage that only two surface expansions are converted into the corresponding paracanonical expansions. Moreover, both P and R are scalars and (as will be seen in §36,37) the same generic set of equations for performing pseudo-expansions and iteration may be applied to both. This is to be compared with the situation when $\Delta\underset{\sim}{G}$ is first expressed as a surface expansion. ΔG_y and ΔG_z behave differently under reflection in Ⓜ and two generically different sets of iteration equations etc., are required (§38).

In spite of the length of the numerical computations involved, the whole process is highly systematic and as such is readily extended to high orders. The basic steps are discussed in (i) to (v) below:

(i) X and R are expressed in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ by solving in series the equations (27.1,3). The method employed is that of M§75 and is discussed in §30 and applied in §31,32. The surface expansion of P = XR is then simply determined.

(ii) The surface expansions for P and R are converted into the corresponding paracanonical expansions by first, as a purely formal device, setting $\underset{\sim}{\delta}_y = \underset{\sim}{\delta}_v = 0$ so that $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are linear in $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Substitution of this linear relationship into the surface expansions yields a new type of expansion, the so-called pseudo-expansion, in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. (See M§10a and §35.36 of this thesis.)

(iii) Since $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ were neglected the pseudo-expansion is not an exact expansion. However, by replacing $\underset{\sim}{S}$ by $\underset{\sim}{S} + \underset{\sim}{\delta}_y$ and $\underset{\sim}{T}$ by $\underset{\sim}{T} + \underset{\sim}{\delta}_v$, the pseudo-expansion becomes identically equal to the paracanonical expansion (see (35.3)). By expressing $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ up to the ($\underline{n}$-1)th order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ use can be made of this identity to express the $\underline{n}$th order p*- and r*-coefficients in terms of the pseudo-coefficients (i.e., the coefficients occurring in the pseudo-expansion) and the $\underset{\sim}{G}$-coefficients of orders less than $\underline{n}$ (see §35,37 and 38 and M§10,11,81 and 84). At this stage the p*- and r*-coefficients have been determined to the $\underline{n}$th order by assuming that the $\underset{\sim}{g}$-coefficients of orders less than $\underline{n}$ are known.

(iv) In view of this assumption $\hat{p}^{(1)}$ and $\hat{r}^{(1)}$, known from step (i), may be expanded to the ($\underline{n}$-1)th order (see §39). The $\underset{\sim}{i}$-coefficients of orders less than $\underline{n}$ follow directly from (25.6) according to (10.4).

(v) Consequently P, R and $\underset{\sim}{I}$ are known to the ($\underline{n}$-1)th order, and P* and R* to the $\underline{n}$th order. $\underset{\sim}{P}$ and $\underset{\sim}{R}$ and the corresponding $\underset{\sim}{p}$- and $\underset{\sim}{r}$-coefficients can be calculated to the $\underline{n}$th order (see (26.5) and §40) and the $\underset{\sim}{g}$-coefficients follow immediately from (26.9). Summation over the surfaces gives the $\underset{\sim}{G}$-coefficients.

29. Coefficients and Notation

The paracanonical coefficients of Q are denoted by $q_{\mu\nu\tau}^{(n)}$. This is the basic form of all coefficients and is extremely useful when writing down general relations such as (24.11,12) and when first deriving various equations, especially when products of series are involved. The notation for all coefficients is derived from $q_{\mu\nu\tau}^{(n)}$ by the addition of various affixes. In particular the type of the expansion is indicated by a modification of the kernel symbol q of the coefficient. For pseudo-coefficients (see §36) q becomes $\underline{\underline{q}}$ and for surface coefficients, $\hat{q}$. All possible affixes are characterized by their <u>rank</u> such that if two or more affixes are present, the one with the lower rank appears nearer to the kernel symbol. Affixes of positive rank are to the right of the kernel symbol, those with negative rank, to the left. The ^ and = are assigned zero rank. The possible affixes (to basal coefficients) along with their rank and interpretation are given in Table 29/1. Appendix H may also be consulted in this context. If the kernel symbol is a capital letter, the coefficient in question arises in an expansion of $\sum_i \Delta Q_i$. A comprehensive example is $`G_{y211aj}^{(3)\prime}$ denoting:

$$(g/N_{y1})(\text{coefficient of } S_y T_y T_z \text{ in the } \underline{a}\text{-component of } \sum_{i=1}^{j} \Delta G_{yi}) \,.$$

The † never appears on a coefficient since $Q^{\dagger}$ is the sum of the terms of the second and higher orders in the <u>paracanonical</u> expansion of Q, the coefficients of which are exactly the corresponding q-coefficients. For a similar reason ^ and * never appear on the same coefficient.

An alternative and more readable notation which will be used interchangeably with the above is as follows. The general form of the new coefficients corresponding to $q_{\mu\nu\tau}^{(n)}$ is $k_{q\alpha}$ where k is a new kernel symbol determined uniquely by (n), q designates the quantity to which the coefficient pertains and α is uniquely determined by μ, ν and τ. As before, other coefficients are distinguished by the addition of affixes to the basic form. The affixes of rank -1,0,+1,4,5 and 7 of Table 29/1 apply unchanged. The new affixes of rank 2 and 3 are given in Table 29/2. For n = 1,2 or 3, k is p,s or t respectively. A capital P, S or T as a kernel symbol denotes a coefficient of $\sum_i \Delta Q_i$. Surface and pseudo-coefficients are denoted by $\hat{k}$ and $\underline{\underline{k}}$ respectively. α replaces μ, ν and τ and is given by

$$\alpha = 1 + \tau + \tfrac{1}{2}\nu(\nu+1) + \tfrac{1}{6}\mu(\mu+1)(\mu+2) \quad , \qquad (29.1)$$

which is equivalent to placing all the terms of the expansion into one-to-one correspondence with the integers α = 1,2,3,... in the following manner: for any two terms the one with the higher power of S_y (or Y) corresponds to the lower value of α. If the two powers of S_y are the same, then the one with the higher power of T_y (or V) has the lower value of α. If still no distinction has been drawn, the powers of S_z (or Z) and finally of T_z (or W) are similarly taken into account. As an example, '$G^{(3)}_{y211aj}$' becomes 'T_{gy7aj}'. If Q has some particular symmetry property, the coefficients corresponding to certain values of α may be identically zero. For example, since $\underset{\sim}{G}$ is a proper two-vector, 'T_{gy7aj}' is zero. However, these coefficients are allowed for in the numbering by α, ensuring that the notation carries over unchanged when the base-ray is no longer meridional.

Table 29/3 lists all possible terms of the first three orders occurring in proper two-vectors and both forms of the corresponding coefficients. The column headed "dual" is relevant to Appendix E. Essentially, if any valid relationship between coefficients is transformed

TABLE 29/1 Interpretation of Affixes

Affix	Rank	Type[1] and Meaning
`	-1	Sp. The ante-prime, see §25.
^ or =	0	Surface coefficients or pseudo-coefficients respectively.
*	+1	Sp. Indicates a coefficient of Q*.
y or z	2	Sb. Imply y- or z-components of coefficients of a two-vector.
μ ν τ	3	Sb. Together with (n), these indicate which term the coefficient multiplies.
a or b	4	Sb. a- or b-components of a coefficient.
i,j or k	5	Sb. Surface indicator, k implies last surface.
(n)	6	Sp. Order of coefficient.
'	7	Sp. Prime - implies after refraction.

TABLE 29/2 Affixes Characteristic of $k_{q\alpha}$ Coefficients

Affix	Rank	Type and Meaning
q,qy or qz	2	Sb. Indicates a coefficient of Q, Q_y or Q_z respectively.
α	3	Sb. Used in place of μ ν τ .

---

[1] Sp. is an abbreviation for superscript, and Sb. for subscript.

TABLE 29/3 Coefficients and Their Duals

Term	Coefficients		Dual	Term	Coefficients		Dual
S_y	$q_{y000}^{(1)}$	p_{qy1}	p_{qy2}	S_z	$q_{z110}^{(1)}$	p_{qz3}	p_{qz4}
T_y	$q_{y100}^{(1)}$	p_{qy2}	p_{qy1}	T_z	$q_{z111}^{(1)}$	p_{qz4}	p_{qz3}
S_y^2	$q_{y000}^{(2)}$	s_{qy1}	s_{qy5}	S_yS_z	$q_{z110}^{(2)}$	s_{qz3}	s_{qz7}
S_yT_y	$q_{y100}^{(2)}$	s_{qy2}	s_{qy2}	S_yT_z	$q_{z111}^{(2)}$	s_{qz4}	s_{qz6}
T_y^2	$q_{y200}^{(2)}$	s_{qy5}	s_{qy1}	T_yS_z	$q_{z210}^{(2)}$	s_{qz6}	s_{qz4}
S_z^2	$q_{y220}^{(2)}$	s_{qy8}	s_{qy10}	T_yT_z	$q_{z211}^{(2)}$	s_{qz7}	s_{qz3}
S_zT_z	$q_{y221}^{(2)}$	s_{qy9}	s_{qy9}				
T_z^2	$q_{y222}^{(2)}$	s_{qy10}	s_{qy8}				
S_y^3	$q_{y000}^{(3)}$	t_{qy1}	t_{qy11}	$S_y^2S_z$	$q_{z110}^{(3)}$	t_{qz3}	t_{qz13}
$S_y^2T_y$	$q_{y100}^{(3)}$	t_{qy2}	t_{qy5}	$S_y^2T_z$	$q_{z111}^{(3)}$	t_{qz4}	t_{qz12}
$S_yT_y^2$	$q_{y200}^{(3)}$	t_{qy5}	t_{qy2}	$S_yT_yS_z$	$q_{z210}^{(3)}$	t_{qz6}	t_{qz7}
$S_yS_z^2$	$q_{y220}^{(3)}$	t_{qy8}	t_{qy16}	$S_yT_yT_z$	$q_{z211}^{(3)}$	t_{qz7}	t_{qz6}
$S_yS_zT_z$	$q_{y221}^{(3)}$	t_{qy9}	t_{qy15}	$T_y^2S_z$	$q_{z310}^{(3)}$	t_{qz12}	t_{qz4}
$S_yT_z^2$	$q_{y222}^{(3)}$	t_{qy10}	t_{qy14}	$T_y^2T_z$	$q_{z311}^{(3)}$	t_{qz13}	t_{qz3}
T_y^3	$q_{y300}^{(3)}$	t_{qy11}	t_{qy1}	S_z^3	$q_{z330}^{(3)}$	t_{qz17}	t_{qz20}
$T_yS_z^2$	$q_{y320}^{(3)}$	t_{qy14}	t_{qy10}	$S_z^2T_z$	$q_{z331}^{(3)}$	t_{qz18}	t_{qz19}
$T_yS_zT_z$	$q_{y321}^{(3)}$	t_{qy15}	t_{qy9}	$S_zT_z^2$	$q_{z332}^{(3)}$	t_{qz19}	t_{qz18}
$T_yT_z^3$	$q_{y322}^{(3)}$	t_{qy16}	t_{qy8}	T_z^3	$q_{z333}^{(3)}$	t_{qz20}	t_{qz17}

by replacing all coefficients by their duals (obtained from Table 29/3), the relationship remains valid.

30. Series Solution of Quadratic Equations

Let Q be a scalar function of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ given by the quadratic equation

$$\tau Q^2 + 2\sigma Q + \mu = 0 \tag{30.1}$$

where τ, σ and μ are functions of the canonical variables. In the notation of §26b write

$$\tau = \sum_{n=0}^{\infty}\hat{\tau}^{(n)} \quad , \quad \sigma = \sum_{n=0}^{\infty}\hat{\sigma}^{(n)} \quad , \quad \mu = \sum_{n=0}^{\infty}\hat{\mu}^{(n)} \quad , \quad Q = \sum_{n=0}^{\infty}\hat{q}^{(n)} \quad .$$

Substitution of these into (30.1) and selection of the nth order terms yields:

$$\sum_{m=0}^{n}\left[\sum_{r=0}^{m}\hat{\tau}^{(n-m)}\hat{q}^{(m-r)}\hat{q}^{(r)}+2\hat{\sigma}^{(n-m)}\hat{q}^{(m)}\right] + \hat{\mu}^{(n)} = 0 \quad . \tag{30.2}$$

For n = 0 this reduces to

$$\tau_0 Q_0{}^2 + 2\sigma_0 Q_0 + \mu_0 = 0 \quad .$$

The relevant solution for Q_0 is of course the value of Q for the base-ray. (30.2) can be rearranged as a recurrence relation for the $\hat{q}^{(n)}$:

$$\hat{q}^{(n)} = -\left\{ Q_0{}^2\hat{\tau}^{(n)}+2Q_0\hat{\sigma}^{(n)}+\hat{\mu}^{(n)}+\sum_{m=1}^{n-1}\left[\sum_{r=1}^{m-1}\hat{\tau}^{(n-m)}\hat{q}^{(m-r)}\hat{q}^{(r)}+\hat{q}^{(n-m)}(\tau_0\hat{q}^{(m)}+\right.\right.$$
$$\left.\left. +2\hat{\sigma}^{(m)}+2Q_0\hat{\tau}^{(m)})\right]\right\} / 2(\tau_0 Q_0+\sigma_0) \quad .$$

Without loss of generality[1] take $2(\tau_0 Q_0+\sigma_0) = -1$ and define

---

[1] This may always be arranged simply by multiplying (30.1) through by a suitable constant factor.

$$\hat{\Phi}^{(n)} = Q_0{}^2\hat{\tau}^{(n)} + 2Q_0\hat{\sigma}^{(n)} + \hat{\mu}^{(n)} \;,\; \hat{\Psi}^{(n)} = \tau_0\hat{q}^{(n)} + 2\hat{\sigma}^{(n)} + 2Q_0\hat{\tau}^{(n)} \;. \qquad (30.3)$$

Then

$$\hat{q}^{(n)} = \hat{\Phi}^{(n)} + \sum_{m=1}^{n-1}\left[\hat{q}^{(n-m)}\hat{\Psi}^{(m)} + \hat{\tau}^{(n-m)}\sum_{r=1}^{m-1}\hat{q}^{(m-r)}\hat{q}^{(r)}\right] \;. \qquad (30.4)$$

In particular, for the first three orders,

$$\hat{q}^{(1)} = \hat{\Phi}^{(1)} \;,$$

$$\hat{q}^{(2)} = \hat{\Phi}^{(2)} + \hat{q}^{(1)}\hat{\Psi}^{(1)} \;,$$

$$\hat{q}^{(3)} = \hat{\Phi}^{(3)} + \hat{q}^{(1)}(\hat{\Psi}^{(2)} + \hat{\tau}^{(1)}\hat{q}^{(1)}) + \hat{q}^{(2)}\hat{\Psi}^{(1)} \;. \qquad (30.5)$$

From §26b, $\hat{q}^{(n)}$ has the expansion

$$q^{(n)} = \sum_{\mu=0}^{n}\sum_{\nu=0}^{\mu}\sum_{\tau=0}^{\nu}\hat{q}_{\mu\nu\tau}{}^{(n)}Y^{(n-\mu)}V^{(\mu-\nu)}Z^{(\nu-\tau)}W^{\tau} \;.$$

If the corresponding expansions of $\hat{\mu}^{(n)}$, $\hat{\sigma}^{(n)}$ and $\hat{\tau}^{(n)}$ are substituted into (30.4), the $\hat{q}$-coefficients of a given order are readily expressed in terms of the $\hat{\tau}$-, $\hat{\sigma}$- and $\hat{\mu}$-coefficients and the lower order $\hat{q}$-coefficients. In general the $\hat{\Phi}$- and $\hat{\Psi}$-coefficients are given by

$$\hat{\Phi}_{\mu\nu\tau}{}^{(n)} = Q_0{}^2\hat{\tau}_{\mu\nu\tau}{}^{(n)} + 2Q_0\hat{\sigma}_{\mu\nu\tau}{}^{(n)} + \hat{\mu}_{\mu\nu\tau}{}^{(n)} \;,$$

$$\hat{\Psi}_{\mu\nu\tau}{}^{(n)} = \tau_0\hat{q}_{\mu\nu\tau}{}^{(n)} + 2\hat{\sigma}_{\mu\nu\tau}{}^{(n)} + 2Q_0\hat{\tau}_{\mu\nu\tau}{}^{(n)} \;. \qquad (30.6)$$

In terms of these, the $\hat{q}$-coefficients of the first three orders are given in Table 30/1. The equations of this table will in particular apply to both X and R. Whereas it proves possible to express the $\hat{x}$-coefficients simply and explicitly in terms of the parameters of the system (§31), with the exception of those of the first order, this is in general not possible for the $\hat{r}$-coefficients which are accordingly left in terms of the $\hat{\mu}$-, $\hat{\sigma}$- and $\hat{\tau}$-coefficients (§32).

TABLE 30/1 The $\hat{q}$-coefficients in terms of the $\hat{\tau}$-, $\hat{\sigma}$- and $\hat{\mu}$-coefficients

$$\hat{p}_{q1} = \hat{p}_{\phi 1}$$

$$\hat{p}_{q2} = \hat{p}_{\phi 2}$$

$$\hat{s}_{q1} = \hat{s}_{\phi 1} + \hat{p}_{q1}\hat{p}_{\psi 1}$$

$$\hat{s}_{q2} = \hat{s}_{\phi 2} + \hat{p}_{q1}\hat{p}_{\psi 2} + \hat{p}_{q2}\hat{p}_{\psi 1}$$

$$\hat{s}_{q5} = \hat{s}_{\phi 5} + \hat{p}_{q2}\hat{p}_{\psi 2}$$

$$\hat{s}_{q8} = \hat{s}_{\phi 8}$$

$$\hat{s}_{q9} = \hat{s}_{\phi 9}$$

$$\hat{s}_{q10} = \hat{s}_{\phi 10}$$

$$\hat{t}_{q1} = \hat{t}_{\phi 1} + \hat{p}_{q1}(\hat{s}_{\psi 1}+\hat{p}_{\tau 1}\hat{p}_{q1}) + \hat{s}_{s1}\hat{p}_{\psi 1}$$

$$\hat{t}_{q2} = \hat{t}_{\phi 2} + \hat{p}_{q1}(\hat{s}_{\psi 2}+\hat{p}_{\tau 1}\hat{p}_{q2}+\hat{p}_{\tau 2}\hat{p}_{q1}) + \hat{p}_{q2}(\hat{s}_{\psi 1}+\hat{p}_{\tau 1}\hat{p}_{q1}) + \hat{p}_{\psi 1}\hat{s}_{q2} + \hat{p}_{\psi 2}\hat{s}_{q1}$$

$$\hat{t}_{q5} = \hat{t}_{\phi 5} + \hat{p}_{q1}(\hat{s}_{\psi 5}+\hat{p}_{\tau 2}\hat{p}_{q2}) + \hat{p}_{q2}(\hat{s}_{\psi 2}+\hat{p}_{\tau 1}\hat{p}_{q2}+\hat{p}_{\tau 2}\hat{p}_{q1}) + \hat{p}_{\psi 1}\hat{s}_{q5} + \hat{p}_{\psi 2}\hat{s}_{q2}$$

$$\hat{t}_{q8} = \hat{t}_{\phi 8} + \hat{p}_{q1}\hat{s}_{\psi 8} + \hat{s}_{q8}\hat{p}_{\psi 1}$$

$$\hat{t}_{q9} = \hat{t}_{\phi 9} + \hat{p}_{q1}\hat{s}_{\psi 9} + \hat{s}_{q9}\hat{p}_{\psi 1}$$

$$\hat{t}_{q10} = \hat{t}_{\phi 10} + \hat{p}_{q1}\hat{s}_{\psi 10} + \hat{s}_{q10}\hat{p}_{\psi 1}$$

$$\hat{t}_{q11} = \hat{t}_{\phi 11} + \hat{p}_{q2}(\hat{s}_{\psi 5}+\hat{p}_{\tau 2}\hat{p}_{q2}) + \hat{s}_{q5}\hat{p}_{\psi 2}$$

$$\hat{t}_{q14} = \hat{t}_{\phi 14} + \hat{p}_{q2}\hat{s}_{\psi 8} + \hat{s}_{q8}\hat{p}_{\psi 2}$$

$$\hat{t}_{q15} = \hat{t}_{\phi 15} + \hat{p}_{q2}\hat{s}_{\psi 9} + \hat{s}_{q9}\hat{p}_{\psi 2}$$

$$\hat{t}_{q16} = \hat{t}_{\phi 16} + \hat{p}_{q2}\hat{s}_{\psi 10} + \hat{s}_{q10}\hat{p}_{\psi 2}$$

31. The $\hat{x}$-coefficients

After being renormalised so that $2(\rho_{20}X_0+\rho_{10}) = -1$, equations (27.1,2) for X read

$$\rho_2 X^2 + 2\rho_1 X + \rho_0 = 0 \quad ,$$

where

$$\rho_2 = \tfrac{1}{2}c\lambda_1(1/\alpha_B{}^2+2V_BV+\underset{\sim}{V}.\underset{\sim}{V}) \quad , \quad \rho_1 = -\tfrac{1}{2}[1+\lambda_1 n_{By}V-c\lambda_1(V_BY+\underset{\sim}{V}.\underset{\sim}{Y})] \quad ,$$

$$\rho_0 = \lambda_1(\tfrac{1}{2}c\underset{\sim}{Y}.\underset{\sim}{Y}-n_{By}Y) \quad . \tag{31.1}$$

The series for X is

$$X = \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \hat{x}_{\mu\nu\tau}{}^{(n)} Y^{(n-\mu)} V^{(\mu-\nu)} Z^{(\nu-\tau)} W^{\tau} \quad , \tag{31.2}$$

where $X_0 = 0$ since the base-ray passes through the origin of Ⓒ. Thus, in the notation of §30,

$$\hat{\Phi}^{(n)} = \hat{\rho}_0{}^{(n)} \quad , \quad \hat{\Psi}^{(n)} = \rho_{20}\hat{x}^{(n)} + 2\hat{\rho}_1{}^{(n)} \quad . \tag{31.3}$$

It follows that

$$\hat{x}^{(1)} = \hat{\rho}_0{}^{(1)} = -\lambda_1 n_{By} Y \quad ,$$

and, accordingly,

$$\hat{p}_{x1} = -\lambda_1 n_{By} \quad , \quad \hat{p}_{x2} = 0 \quad .$$

The higher order $\hat{x}$-coefficients are determined readily from Table 30/1. For instance, considering $\hat{s}_{x1}$, the coefficients $\hat{s}_{\phi 1}$ and $\hat{p}_{\psi 1}$ are required. These are obtained from (31.3,1):

$$\hat{s}_{\phi 1} = \tfrac{1}{2}c\lambda_1 \quad , \quad \hat{p}_{\psi 1} = c\lambda_1(V_B-\tfrac{1}{2}\lambda_1 n_{By}/\alpha_B{}^2) \quad .$$

Thus

$$\hat{s}_{x1} = \tfrac{1}{2}\lambda_1 c + \tfrac{1}{2}\lambda_1{}^2 n_{By} c(\lambda_1 n_{By}/\alpha_B{}^2-2V_B)$$

$$= \tfrac{1}{2}(\lambda_1 c/\cos^2 I_B)[\cos^2 I_B+n_{By}{}^2-2\beta_B n_{By}\cos I_B]$$

$$= \tfrac{1}{2}(\lambda_1 c/\cos^2 I_B)[(\cos I_B-n_{By}\beta_B)^2+n_{By}{}^2\alpha_B{}^2] = \tfrac{1}{2}c\lambda_1{}^3 \quad ,$$

where (11.2) has been used. In this fashion the second and third order $\hat{x}$-coefficients are determined and are given in Table 31/1.

TABLE 31/1 The $\hat{x}$-coefficients

$\hat{p}_{x1} = -n_{By}\lambda_1 \qquad \hat{p}_{x2} = 0 \quad ,$

$\hat{s}_{x1} = \tfrac{1}{2}c\lambda_1^{\;3} \qquad \hat{s}_{x2} = \hat{p}_{x1}^{\;2} \qquad \hat{s}_{x5} = 0$

$\hat{s}_{x8} = \tfrac{1}{2}c\lambda_1 \qquad \hat{s}_{x9} = 0 \qquad \hat{s}_{x10} = 0 \quad ,$

$\hat{t}_{x1} = cI_{By}\lambda_1^{\;2}\hat{s}_{x1} \qquad \hat{t}_{x2} = 3\hat{p}_{x1}\hat{s}_{x1} \qquad \hat{t}_{x5} = \hat{p}_{x1}^{\;3}$

$\hat{t}_{x8} = cI_{By}\lambda_1^{\;2}\hat{s}_{x8} \qquad \hat{t}_{x9} = 2\hat{p}_{x1}\hat{s}_{x8} \qquad \hat{t}_{x10} = 0$

$\hat{t}_{x11} = 0 \qquad \hat{t}_{x11} = \hat{p}_{x1}\hat{s}_{x8} \qquad \hat{t}_{x15} = 0$

$\hat{t}_{x16} = 0 \quad .$

32. The $\hat{r}$-coefficients

(a) It has already been mentioned that except for those of first order, it is not possible to express the $\hat{r}$-coefficients in a simple manner directly in terms of the system parameters. It is of value therefore to find as simple forms as possible for the $\hat{\mu}$-, $\hat{\sigma}$- and $\hat{\tau}$-coefficients. After renormalising so that $2(\tau_0 R_0 + \sigma_0) = -1$, equations (27.3,4) for R read:

$$\tau R^2 + 2\sigma R + \mu = 0 \quad , \tag{32.1}$$

where

$$\begin{aligned} \mu &= \mu_0\bar{\mu} = \mu_0(1+2\alpha_B\beta_B V + \alpha_B^{\;2}\underset{\sim}{V}\cdot\underset{\sim}{V}) \quad , \\ \sigma &= -\lambda_2 n_x\bar{\mu} + \lambda_3[I_{By}V_B + (I_{By}+n_{Bx}V_B)V + cV_B Y + \underset{\sim}{I}\cdot\underset{\sim}{V}] \quad , \\ \tau &= \bar{R}_B\lambda_2 n_x^{\;2}\bar{\mu} - \bar{R}_B\lambda_3[I_{By}^{\;2} + 2I_{By}(cY + n_{Bx}V) + \underset{\sim}{I}\cdot\underset{\sim}{I}] \quad . \end{aligned} \tag{32.2}$$

For convenience the following combinations of parameters have been defined in addition to λ_1 (see (13.2)):

$$\lambda_2 = (2k\alpha_B \cos I_B')^{-1} \quad , \quad \lambda_3 = k^2\alpha_B{}^2\lambda_2 \quad , \quad \lambda_4 = -2\bar{R}_B n_{Bx} \quad ,$$

$$\mu_0 = (1-k^2)\lambda_2/\bar{R}_B \quad . \tag{32.3}$$

(λ_4 is for use later in this section.) For the base-ray R = 1 and thus the relevant solution of the equation following (30.2) is $R_0 = 1$. That this is indeed a solution is readily checked: from (32.2) it is found that

$$\tau_0 + 2\sigma_0 + \mu_0 = \bar{R}_B\lambda_2 n_{Bx}{}^2 - \bar{R}_B\lambda_3 I_{By}{}^2 - 2\lambda_2 n_{Bx} + 2\lambda_3 I_{By}V_B + (1-k^2)\lambda_2/\bar{R}_B$$

$$= -\lambda_2 n_{Bx}(1+k\alpha_B/\alpha_B') + \lambda_3 I_{By}(V_B'+V_B) + (1-k^2)\lambda_2/\bar{R}_B$$

$$= (\lambda_2/\bar{R}_B)[k^2(\alpha_B/\alpha_B')^2 - 1 + k^2\alpha_B{}^2(V_B{}^2 - V_B'{}^2) + 1 - k^2]$$

$$= (\lambda_2/\alpha_B'{}^2\bar{R}_B)[k^2\alpha_B{}^2 - k^2\alpha_B'{}^2 + k^2\alpha_B'{}^2\beta_B{}^2 - k^2\alpha_B{}^2\beta_B'{}^2] = 0 \quad ,$$

as required. Moreover

$$2(\tau_0 R_0 + \sigma_0) = 2\lambda_2 n_{Bx}(\bar{R}_B n_{Bx} - 1) - 2\lambda_3 I_{By}(\bar{R}_B I_{By} + V_B)$$

$$= 2(\lambda_2 k\alpha_B/\alpha_B')(\alpha_B' I_{By}'\beta_B' - n_{Bx})$$

$$= -2\lambda_2 k\alpha_B \cos I_B' = -1 \quad ,$$

where in both cases extensive use has been made of (11.1-7). From these two equations it follows that

$$\sigma_0 = \tfrac{1}{2} - \mu_0 \quad , \quad \tau_0 = \mu_0 - 1 \quad . \tag{32.4}$$

(b) The $\hat{\mu}$-, $\hat{\sigma}$- and $\hat{\tau}$-coefficients of all orders can be read off (32.2). Apart from the use of λ_1, λ_2, λ_3 and λ_4, no attempt has been made to replace all repeated combinations of parameters with a single symbol. When constructing a computing scheme, either for an electronic computer, or for a desk calculating machine, it is of advantage to do this. However, a certain amount of simplification has been carried out. For example, from (32.2) the coefficients of V in σ and τ are found to be

$$\hat{p}_{\tau 2} = 2\bar{R}_B n_{Bx}(\lambda_2 n_{Bx}\alpha_B\beta_B - \lambda_3 I_{By}) \quad ,$$

$$\hat{p}_{\sigma 2} = \lambda_3(I_{By} + n_{Bx}V_B) - 2\lambda_2 n_{Bx}\alpha_B\beta_B \quad .$$

The last of these may be used to eliminate $2\lambda_2 n_{Bx}\alpha_B\beta_B$ from $\hat{p}_{\tau 2}$:

$$\begin{aligned}\hat{p}_{\tau 2} &= -\bar{R}_B n_{Bx}[(I_{By} - n_{Bx}V_B)\lambda_3 + \hat{p}_{\sigma 2}] \\ &= \tfrac{1}{2}\lambda_4(\hat{p}_{\sigma 2} - n_{By}\lambda_3) \quad .\end{aligned}$$

The third order coefficients warrant special mention. All third order $\hat{\mu}$-coefficients are zero and from (32.2)

$$\hat{\sigma}^{(3)} = c\lambda_2(\hat{x}^{(3)} + 2\alpha_B\beta_B\hat{x}^{(3)}V + \alpha_B{}^2\hat{x}^{(1)}\underline{V}.\underline{V}) \quad ,$$

$$\hat{\tau}^{(3)} = \lambda_4\hat{\sigma}^{(3)} + 2c^2\bar{R}_B\lambda_2\hat{x}^{(1)}(\hat{x}^{(2)} + \alpha_B\beta_B\hat{x}^{(1)}V) \quad ,$$

from which the third order coefficients are easily obtained. The <u>only</u> non-zero $\hat{\mu}$-coefficients are

$$\mu_0 = (1-k^2)\lambda_2/\bar{R}_B \quad , \quad \hat{p}_{\mu 2} = 2\alpha_B\beta_B\mu_0 \quad , \quad \hat{s}_{\mu 5} = \hat{s}_{\mu 10} = \alpha_B{}^2\mu_0 \quad . \qquad (32.5)$$

The $\hat{\sigma}$- and $\hat{\tau}$-coefficients of the first three orders are given in Table 32/1 and used in conjunction with Table 30/1 to calculate $\hat{r}$-coefficients,

TABLE 32/1 The $\hat{\sigma}$- and $\hat{\tau}$-coefficients

$\hat{p}_{\sigma 1} = c(\lambda_3 V_B + \lambda_2 \hat{p}_{x1})$	$\hat{p}_{\tau 1} = \lambda_4 \hat{p}_{\sigma 1} + 2c\bar{R}_B \lambda_3 n_{By}$
$\hat{p}_{\sigma 2} = \lambda_3(I_{By} + n_{Bx} V_B) - 2\lambda_2 n_{Bx} \alpha_B \beta_B$	$\hat{p}_{\tau 2} = \frac{1}{2}\lambda_4(\hat{p}_{\sigma 2} - n_{By}\lambda_3)$
$\hat{s}_{\sigma 1} = c\lambda_2 \hat{s}_{x1}$	$\hat{s}_{\tau 1} = \lambda_4 \hat{s}_{\sigma 1} + c^2 \bar{R}_B(\lambda_2 \hat{s}_{x2} - \lambda_3)$
$\hat{s}_{\sigma 2} = c\lambda_3 + c\lambda_2(2\alpha_B \beta_B \hat{p}_{x1} + \hat{s}_{x2})$	$\hat{s}_{\tau 2} = \lambda_4 \hat{s}_{\sigma 2}$
$\hat{s}_{\sigma 5} = n_{Bx}(\lambda_3 - \alpha_B{}^2 \lambda_2)$	$\hat{s}_{\tau 5} = \frac{1}{2}\lambda_4 \hat{s}_{\sigma 5}$
$\hat{s}_{\sigma 8} = c\lambda_2 \hat{s}_{x8}$	$\hat{s}_{\tau 8} = \lambda_4 \hat{s}_{\sigma 8} - \bar{r}_B \lambda_3 c^2$
$\hat{s}_{\sigma 9} = c\lambda_3$	$\hat{s}_{\tau 9} = \lambda_4 \hat{s}_{\sigma 9}$
$\hat{s}_{\sigma 10} = \hat{s}_{\sigma 5}$	$\hat{s}_{\tau 10} = \hat{s}_{\tau 5}$
$\hat{t}_{\sigma 1} = c\lambda_2 \hat{t}_{x1}$	$\hat{t}_{\tau 1} = \lambda_4 \hat{t}_{\sigma 1} + 2c^2 \bar{R}_B \lambda_2 \hat{s}_{x1} \hat{p}_{x1}$
$\hat{t}_{\sigma 2} = c\lambda_2(\hat{t}_{x2} + 2\alpha_B \beta_B \hat{s}_{x1})$	$\hat{t}_{\tau 2} = \lambda_4 \hat{t}_{\sigma 2} + 2c^2 \bar{R}_B \lambda_2 \hat{s}_{x2}(\alpha_B \beta_B + \hat{p}_{x1})$
$\hat{t}_{\sigma 5} = c\lambda_2(\hat{t}_{x5} + 2\alpha_B \beta_B \hat{s}_{x2} + \alpha_B{}^2 \hat{p}_{x1})$	$\hat{t}_{\tau 5} = \lambda_4 \hat{t}_{\sigma 5}$
$\hat{t}_{\sigma 8} = c\lambda_2 \hat{t}_{x8}$	$\hat{t}_{\tau 8} = \lambda_4 \hat{t}_{\sigma 8} + c\bar{S}_B \hat{t}_{\sigma 9}$
$\hat{t}_{\sigma 9} = c\lambda_2 \hat{t}_{x9}$	$\hat{t}_{\tau 9} = \lambda_4 \hat{t}_{\sigma 9}$
$\hat{t}_{\sigma 10} = c\lambda_2 \alpha_B{}^2 \hat{p}_{x1}$	$\hat{t}_{\tau 10} = \lambda_4 \hat{t}_{\sigma 10}$
$\hat{t}_{\sigma 11} = 0$	$\hat{t}_{\tau 11} = 0$
$\hat{t}_{\sigma 14} = c\lambda_2(\hat{t}_{x14} + 2\alpha_B \beta_B \hat{s}_{x8})$	$\hat{t}_{\tau 14} = \lambda_4 \hat{t}_{\sigma 14}$
$\hat{t}_{\sigma 15} = 0$	$\hat{t}_{\tau 15} = 0$
$\hat{t}_{\sigma 16} = 0$	$\hat{t}_{\tau 16} = 0$

33. The $\hat{p}$-coefficients

By definition P = XR. Thus, if the nth order terms of P are $\hat{p}^{(n)}$,

$$\hat{p}^{(n)} = \sum_{m=0}^{n} \hat{x}^{(n-m)} \hat{r}^{(m)} \quad .$$

Since X vanishes for the base-ray $P_0 = 0$. From the equation above:

$$\hat{p}^{(1)} = \hat{x}^{(1)}$$

$$\hat{p}^{(2)} = \hat{x}^{(2)} + \hat{x}^{(1)}\hat{r}^{(1)}$$

$$\hat{p}^{(3)} = \hat{x}^{(3)} + \hat{x}^{(2)}\hat{r}^{(1)} + \hat{x}^{(1)}\hat{r}^{(2)} \quad . \tag{33.1}$$

By expressing $\hat{x}^{(n)}$ and $\hat{r}^{(n)}$ as the appropriate series in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, the $\hat{p}$-coefficients may be written down and are given in Table 33/1.

TABLE 33/1 The $\hat{p}$-coefficients

$\hat{p}_{p1} = \hat{p}_{x1}$ $\qquad$ $\hat{p}_{p2} = 0$,

$\hat{s}_{p1} = \hat{s}_{x1} + \hat{p}_{x1}\hat{p}_{r1}$ $\qquad$ $\hat{s}_{p2} = \hat{s}_{x2} + \hat{p}_{x1}\hat{p}_{r2}$

$\hat{s}_{p5} = 0$ $\qquad$ $\hat{s}_{p8} = \hat{s}_{x8}$

$\hat{s}_{p9} = 0$ $\qquad$ $\hat{s}_{p10} = 0$,

$\hat{t}_{p1} = \hat{t}_{x1} + \hat{s}_{x1}\hat{p}_{r1} + \hat{p}_{x1}\hat{s}_{r1}$ $\qquad$ $\hat{t}_{p2} = \hat{t}_{x2} + \hat{s}_{x2}\hat{p}_{r1} + \hat{s}_{x1}\hat{p}_{r2} + \hat{p}_{x1}\hat{s}_{r2}$

$\hat{t}_{p5} = \hat{t}_{x5} + \hat{s}_{x2}\hat{p}_{r2} + \hat{p}_{x1}\hat{s}_{r5}$ $\qquad$ $\hat{t}_{p8} = \hat{t}_{x8} + \hat{s}_{x8}\hat{p}_{r1} + \hat{p}_{x1}\hat{s}_{r8}$

$\hat{t}_{p9} = \hat{t}_{x9} + \hat{p}_{x1}\hat{s}_{r9}$ $\qquad$ $\hat{t}_{p10} = \hat{p}_{x1}\hat{s}_{r10}$

$\hat{t}_{p11} = 0$ $\qquad$ $\hat{t}_{p14} = \hat{t}_{x14} + \hat{s}_{x8}\hat{p}_{r2}$

$\hat{t}_{p15} = 0$ $\qquad$ $\hat{t}_{p16} = 0$.

34. The Second Order Nature of ΔG_y

In order to verify (26.7), in other words that ΔG_y as given by (26.1) is of the second order, it is necessary to obtain the first order $\hat{r}$-coefficients explicitly in terms of the systems parameters. For $\hat{p}_{r1}$, Table 30/1 yields

$$\hat{p}_{r1} = \hat{p}_{\phi 1} = \hat{p}_{\tau 1} + 2\hat{p}_{\sigma 1} \quad ,$$

and by Table 32/1

$$\begin{aligned}
\hat{p}_{r1} &= (\lambda_4+2)\hat{p}_{\sigma 1} + 2c\bar{R}_B\lambda_3 n_{By} \\
&= (2c\lambda_3/\alpha_B')(k\beta_B+\alpha_B'\bar{R}_B n_{By}) - 2k\alpha_B{}^2 n_{By}\lambda_2 c/\alpha_B'\cos I_B \\
&= (c\alpha_B/\cos I_B\cos I_B'\alpha_B')(k\cos I_B\beta_B'-n_{By}) \quad .
\end{aligned}$$

Similarly,

$$\hat{p}_{r2} = \hat{p}_{\tau 2} + 2\hat{p}_{\sigma 2} + \hat{p}_{\mu 2} = k(\alpha_B{}^2/\alpha_B'\cos I_B'\bar{R}_B)(\beta_B-k\beta_B')$$

where in both cases extensive use of §11 was made. The coefficient of Y in (26.7) is

$$-N_y'\bar{R}_B(v_M'\hat{p}_{p1}+y_M'\hat{p}_{r1})I_{By} - N_y'c\bar{R}_B y_M' - \Delta N_y v_M \quad .$$

Expressed in terms of y_M' and v_M',

$$\Delta N_y v_M = N_y'(\lambda_1'-\lambda_1)v_M'/\lambda_1' - (N_y'c\bar{R}_B\lambda_1/\alpha_B'\cos I_B')y_M' \quad .$$

Consequently, the coefficient of $v_M'Y$ in (26.7) is

$$(N_y'/\lambda_1')(\bar{R}_B I_{By} n_{By}\lambda_1\lambda_1'-\lambda_1'+\lambda_1) \propto \alpha_B\cos I_B' - \alpha_B'\cos I_B + n_{By}\alpha_B\alpha_B'(V_B-V_B') = 0.$$

Furthermore, the coefficient of $y_M'Y$ in (26.7) is proportional to

$$\begin{aligned}
&k\alpha_B I_{By}\cos I_B\beta_B' - \alpha_B(I_{By}n_{By}+1) + \alpha_B'\cos I_B'\cos I_B = \\
&\qquad = \cos I_B(\alpha_B'\beta_B'I_{By}'-n_{Bx}+\alpha_B'\cos I_B') = 0 \; .
\end{aligned}$$

Thus the coefficient of Y vanishes. In a similar fashion it can be shown that the coefficient of V in (26.7) also vanishes. Consequently ΔG_y, as given by (26.1), is indeed of the second order.

VI. PSEUDO-EXPANSIONS, ITERATION AND G-COEFFICIENTS

35. The Principles of Pseudo-expansions and Iteration

P and R are known as surface expansions. Since $\underset{\sim}{I}$ is linear in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, it is a simple matter to determine the surface expansion of $\Delta\underset{\sim}{G}$ from (26.5,9). This must be converted into the corresponding paracanonical expansion. As has already been noted (§28), it is of advantage to convert the surface expansions of P and R into the corresponding paracanonical expansions. (The reasons for this will become apparent in the next few sections.) Inspection of (26.5) shows that if $\underset{\sim}{Q}$ is required to the $\underline{n}$th order, Q* must be found to the $\underline{n}$th order and Q to the ($\underline{n}$-1)th order. If Q* is determined to the $\underline{n}$th order, it is possible to obtain Q to the ($\underline{n}$-1)th from Q* (See §39). Thus, let Q be a scalar whose surface expansion is known and for which the paracanonical expansion of Q* must be found. In particular, Q may be P or R.

The paracanonical expansions of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are obtained from (25.6) and may be substituted into the surface expansion of Q*. This gives the paracanonical expansion of Q* and the q*-coefficients are expressed in terms of the $\hat{q}$-coefficients and the $\underset{\sim}{y}$- and $\underset{\sim}{v}$-coefficients. These expressions are quite lengthy for the third and higher orders but, if the $\underset{\sim}{y}$- and $\underset{\sim}{v}$-coefficients are expressed in terms of the $\underset{\sim}{G}$-coefficients (§40a), considerable simplifications can be made. However, it is considerably more systematic to carry out the conversion from surface to paracanonical expansions in the following manner. Suppose for the moment that parabasal optics completely describes imagery by the first ($\underline{j}$-1) surfaces of the system. Then, before refraction at the $\underline{j}$th surface (14.2) is exact and may be used to express Q* as a paracanonical expansion. Since (14.2) is linear in the paracanonical coordinates, $q^{*(n)}$ arises solely from $\hat{q}^{(n)}$ and the

$q^{*}_{\mu\nu\tau}{}^{(n)}$ are expressed in terms of the $\hat{q}_{\mu\nu\tau}{}^{(n)}$. In general, however, parabasal optics will not describe the imagery produced by the first (j-1) surfaces. However, as a purely formal step the increments δ_y and δ_v in (25.6) are set equal to zero and (14.2) is recovered. If (14.2) is now used to express Q in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$, a new kind of expansion, the pseudo-expansion of Q, is obtained. Since (14.2) is linear, the pseudo-expansion of Q* consists of the terms of the second and higher orders in the pseudo-expansion of Q and is formally written:

$$Q^{*} = \sum_{n=2}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{=}{q}_{\mu\nu\tau}{}^{(n)} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} \quad . \tag{35.1}$$

From the manner in which the pseudo-expansion is constructed, the pseudo-coefficients $\underset{=}{q}_{\mu\nu\tau}{}^{(n)}$ (of both Q and Q*) are obtained from the identity:

$$\sum_{\mu\nu\tau}^{n} [\underset{=}{q}_{\mu\nu\tau}{}^{(n)} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} - \hat{q}_{\mu\nu\tau}{}^{(n)} (y_a S_y + y_b T_y)^{n-\mu} (v_a S_y + v_b T_y)^{\mu-\nu} \times$$
$$\times (z_a S_z + z_b T_z)^{\nu-\tau} (w_a S_z + w_b T_z)^{\tau}] \equiv 0 \quad . \tag{35.2}$$

Since the increments $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ in (25.6) have been neglected, the pseudo-expansion is not an exact expansion. If $\underset{\sim}{S}$ and $\underset{\sim}{T}$ in the pseudo-expansion are replaced by $\underset{\sim}{S} + \underset{\sim}{\delta}_y$ and $\underset{\sim}{T} + \underset{\sim}{\delta}_v$ respectively, (35.1) becomes exact and identical with the exact paracanonical expansion of Q*. This follows from the fact that the same substitution converts (14.2) into the corresponding exact equations (25.6). Accordingly, the exact paracanonical coefficients $q^{*}_{\mu\nu\tau}{}^{(n)}$ of Q* are obtained from the identity:

$$\sum_{n=2}^{\infty} \sum_{\mu\nu\tau}^{n} [q^{*}_{\mu\nu\tau}{}^{(n)} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} - \underset{=}{q}_{\mu\nu\tau}{}^{(n)} (S_y+\delta_y)^{n-\mu} (T_y+\delta_v)^{\mu-\nu} (S_z+\delta_z)^{\nu-\tau} \times$$
$$\times (T_z+\delta_w)^{\tau}] \equiv 0 \quad , \tag{35.3}$$

where $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ are expressed as their paracanonical expansions, the coefficients of which are, apart from numerical factors, the $\underset{\sim}{G}_b$- and $\underset{\sim}{G}_a$-

coefficients respectively (see (25.7) and (37.3)). If the right hand terms of this identity are expanded, equations are obtained which express the $\underline{n}$th order q*-coefficients in terms of the $\underline{q}$-coefficients of orders up to and including the $\underline{n}$th, and the $\underset{\sim}{G}$-coefficients of orders up to the ($\underline{n}$-1)th. These equations are the so-called iteration equations and constitute the key equations for obtaining the $\underline{n}$th order $\underset{\sim}{G}$-coefficients once those of the ($\underline{n}$-1)th order are known. They correctly take into account the aberrations of preceding surfaces. The process whereby a set of iteration equations are applied to some quantity Q is called simply iteration with Q.

At the first surface $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ are identically zero and consequently there is no difference between the paracanonical coefficients and the corresponding pseudo-coefficients. Thus, since the incident rays are aberration free, the pseudo-coefficients completely characterize the aberrations introduced by the first surface. Consider for the moment the iteration equations (M11.3) of the axial theory. If the first ($\underline{j}$-1) surfaces of a symmetric system produce a stigmatic image of the object plane, the $\underline{a}$-components of the axial $\underset{\sim}{G}$-coefficients are zero at the $\underline{j}$th surface. It does not follow, however, that $g_{\mu\nu}^{(n)} = \underline{g}_{\mu\nu}^{(n)}$, $\bar{g}_{\mu\nu}^{(n)} = \bar{\underline{g}}_{\mu\nu}^{(n)}$ (in the notation of M). This is because the $\underline{b}$-components of the $\underset{\sim}{G}$-coefficients are not zero, in other words that imagery between any pair of conjugate planes is not stigmatic. If and only if the first ($\underline{j}$-1) surfaces of the system constitute a perfect optical system will the pseudo-coefficients (or intrinsic coefficients as they are sometimes called) represent the aberrations introduced by the $\underline{j}$th surface. This is also the case in the basal theory. Suppose all the (basal) $\underset{\sim}{G}$-coefficients are zero at the $\underline{j}$th surface. It follows that $q_{\mu\nu\tau}^{(n)} = \underline{\underline{q}}_{\mu\nu\tau}^{(n)}$. However, unless parabasal imagery is stigmatic (and hence imagery as a whole is perfect) there exist two mutually perpendicular straight focal lines (see §21a), which is impossible. Thus imagery must be perfect in order that $q_{\mu\nu\tau}^{(n)} = \underline{\underline{q}}_{\mu\nu\tau}^{(n)}$.

36. Pseudo-coefficients

(a) The pseudo-coefficients of any scalar quantity Q are obtained from the identity (35.2). Since Q is invariant under reflection in Ⓜ, those terms for which ν is odd are absent from (36.2). Consider the second order terms of this identity:

$$\underline{\underline{q}}_{000}^{(2)} S_y^2 + \underline{\underline{q}}_{100}^{(2)} S_y T_y + \underline{\underline{q}}_{200}^{(2)} T_y^2 + \ldots - \hat{q}_{000}^{(2)} (y_a S_y + y_b T_y)^2 -$$

$$- \hat{q}_{100}^{(2)} (y_a S_y + y_b T_y)(v_a S_y + v_b T_y) - \ldots - \hat{q}_{222}^{(2)} (w_a S_z + w_b T_z)^2 \equiv 0 \quad .$$

Hence, by performing the expansions and selecting the coefficients of like terms:

$$\underline{\underline{q}}_{000}^{(2)} = y_a^2 \hat{q}_{000}^{(2)} + y_a v_a \hat{q}_{100}^{(2)} + v_a^2 \hat{q}_{200}^{(2)} \quad ,$$

$$\underline{\underline{q}}_{100}^{(2)} = 2y_a y_b \hat{q}_{000}^{(2)} + (y_a v_b + v_a y_b) \hat{q}_{100}^{(2)} + 2v_a v_b \hat{q}_{200}^{(2)}, \quad (36.1)$$

and so on. In this manner the auxiliary quantities given in Tables 36/1,2 were obtained. The pseudo-coefficients may be read from these tables as follows: multiply the surface coefficients at the head of each column by the corresponding entry in the row headed by the pseudo-coefficient in question and sum the term obtained for each column. Thus, the pseudo-coefficient $\underline{\underline{s}}_{q2}$ is given by

$$\underline{\underline{s}}_{q2} = 2y_a y_b \hat{s}_{q1} + (y_a v_b + v_a y_b) \hat{s}_{q2} + 2v_a v_b \hat{s}_{q5} \quad ,$$

(cf. (36.1)). The first order pseudo-coefficients are given by

$$\underline{\underline{p}}_{q1} = y_a\hat{p}_{q1} + v_a\hat{p}_{q2} \quad , \quad \underline{\underline{p}}_{q2} = y_b\hat{p}_{q1} + v_b\hat{p}_{q2} \quad . \tag{36.2}$$

Needless to say, Tables 36/1,2 apply to both P and R.

(b) Although probably of academic interest only, especially for low orders, a closed formula can be found for the pseudo-coefficients. By performing the expansions in the right hand term of (35.2) the following identity is obtained.

$$\sum_{\mu\nu\tau}^{n} \underline{\underline{q}}_{\mu\nu\tau}^{(n)} S_y^{\,n-\mu} T_y^{\,\mu-\nu} S_z^{\,\nu-\tau} T_z^{\,\tau} \equiv \sum_{\bar{\mu}\bar{\nu}\bar{\tau}}^{n} \sum_{p=0}^{n-\bar{\mu}} \sum_{r=0}^{\bar{\mu}-\bar{\nu}} \sum_{s=0}^{\bar{\nu}-\bar{\tau}} \sum_{t=0}^{\bar{\tau}} \hat{q}_{\bar{\mu}\bar{\nu}\bar{\tau}}^{(n)} \binom{n-\bar{\mu}}{p}\binom{\bar{\mu}-\bar{\nu}}{r} \times$$
$$\times \binom{\bar{\mu}-\bar{\tau}}{s}\binom{\bar{\tau}}{t} y_a^{\,n-\bar{\mu}-p} y_b^{\,p} v_a^{\,\bar{\mu}-\bar{\nu}-r} v_b^{\,r} z_a^{\,\bar{\nu}-\bar{\tau}-s} z_b^{\,s} w_a^{\,\bar{\tau}-t} w_b^{\,t} S_y^{\,n-\bar{\mu}-p-r} T_y^{\,p+r} S_z^{\,\bar{\nu}-s-t} T_z^{\,s+t} . \tag{36.3}$$

The coefficient of $S_y^{\,n-\mu} T_y^{\,\mu-\nu} S_z^{\,\nu-\tau} T_z^{\,\tau}$ in the right hand member of this is obtained by selecting only those terms for which

$$s + t = \tau \quad , \quad \bar{\nu} = \nu \quad , \quad p + r = \mu - \nu \quad .$$

Thus

$$\underline{\underline{q}}_{\mu\nu\tau}^{(n)} = \sum_{\bar{\mu}=0}^{n} \sum_{\bar{\tau}=0}^{\nu} \left[\sum_{p=0}^{n-\bar{\mu}} \binom{n-\bar{\mu}}{p}\binom{\bar{\mu}-\nu}{\mu-\nu-p} y_a^{\,n-\bar{\mu}-p} y_b^{\,p} v_a^{\,\bar{\mu}-\mu+p} v_b^{\,\mu-\nu-p} \right] \times$$
$$\times \left[\sum_{s=0}^{\nu-\bar{\tau}} \binom{\nu-\bar{\tau}}{s}\binom{\bar{\tau}}{\tau-s} z_a^{\,\nu-\bar{\tau}-s} z_b^{\,s} w_a^{\,\bar{\tau}-\tau+s} w_b^{\,\tau-s} \right] \hat{q}_{\bar{\mu}\nu\bar{\tau}}^{(n)} \quad , \tag{36.4}$$

where the summation over p and s is such that none of the exponents of the parabasal coefficients in (36.4) are negative. The coefficient $\hat{q}_{\alpha\beta\gamma}^{(n)}$ in (36.4) is taken to be zero if the inequality $n \geqq \alpha \geqq \beta \geqq \gamma$ is <u>not</u> satisfied. As mentioned in the previous section, the pseudo-coefficients are linear combinations of the surface coefficients of the same order. Moreover, the value of ν for the pseudo-coefficients is equal to that of the surface coefficients of which it is a combination. Thus

TABLE 36/1 Auxiliary Quantities for the Second Order Pseudo-coefficients

	$\hat{s}_{q1}$	$\hat{s}_{q2}$	$\hat{s}_{q5}$
$\underline{\underline{s}}_{q1}$	y_a^2	$y_a v_a$	v_a^2
$\underline{\underline{s}}_{q2}$	$2y_a y_b$	$y_a v_b + y_b v_a$	$2v_a v_b$
$\underline{\underline{s}}_{q5}$	y_b^2	$y_b v_b$	v_b^2

	$\hat{s}_{q8}$	$\hat{s}_{q9}$	$\hat{s}_{q10}$
$\underline{\underline{s}}_{q8}$	z_a^2	$z_a w_a$	w_a^2
$\underline{\underline{s}}_{q9}$	$2z_a z_b$	$z_a w_b + z_b w_a$	$2w_a w_b$
$\underline{\underline{s}}_{q10}$	z_b^2	$z_b w_b$	w_b^2

TABLE 36/2 Auxiliary Quantities for the Third Order Pseudo-coefficients

	$\hat{t}_{q1}$	$\hat{t}_{q2}$	$\hat{t}_{q5}$	$\hat{t}_{q11}$
$\underline{\underline{t}}_{q1}$	y_a^3	$y_a^2 v_a$	$y_a v_a^2$	v_a^3
$\underline{\underline{t}}_{q2}$	$3y_a^2 y_b$	$y_a(y_a v_b + 2v_a y_b)$	$v_a(2y_a v_b + v_a y_b)$	$3v_a^2 v_b$
$\underline{\underline{t}}_{q5}$	$3y_a y_b^2$	$y_b(2y_a v_b + v_a y_b)$	$v_b(y_a v_b + 2v_a y_b)$	$3v_a v_b^2$
$\underline{\underline{t}}_{q11}$	y_b^3	$y_b^2 v_b$	$y_b v_b^2$	v_b^3

	$\hat{t}_{q8}$	$\hat{t}_{q9}$	$\hat{t}_{q10}$	$\hat{t}_{q14}$	$\hat{t}_{q15}$	$\hat{t}_{q16}$
$\underline{\underline{t}}_{q8}$	$y_a z_a^2$	$y_a z_a w_a$	$y_a w_a^2$	$v_a z_a^2$	$v_a z_a w_a$	$v_a w_a^2$
$\underline{\underline{t}}_{q9}$	$2y_a z_a z_b$	$y_a(z_a w_b + z_b w_a)$	$2y_a w_a w_b$	$2v_a z_a z_b$	$v_a(z_a w_b + z_b w_a)$	$2v_a w_a w_b$
$\underline{\underline{t}}_{q10}$	$y_a z_b^2$	$y_a z_b w_b$	$y_a w_b^2$	$v_a z_b^2$	$v_a z_b w_b$	$v_a w_b^2$
$\underline{\underline{t}}_{q14}$	$y_b z_a^2$	$y_b z_a w_a$	$y_b w_a^2$	$v_b z_a^2$	$v_b z_a w_a$	$v_b w_a^2$
$\underline{\underline{t}}_{q15}$	$2y_b z_a z_b$	$y_b(z_a w_b + z_b w_a)$	$2y_b w_a w_b$	$2v_b z_a z_b$	$v_b(z_a w_b + z_b w_a)$	$2v_b w_a w_b$
$\underline{\underline{t}}_{q16}$	$y_b z_b^2$	$y_b z_b w_b$	$y_b w_b^2$	$v_b z_b^2$	$v_b z_b w_b$	$v_b w_b^2$

the equations giving the pseudo-coefficients of any order are split into distinct groups characterized by ν. For example, the second and third order equations split into two groups since ν can take the values 0 and 2, whereas for the fourth and fifth orders, ν can take the additional value 4 and there are three groups.

(c) Suppose that the surface expansions of $\Delta \underline{G}$ are obtained. The pseudo-expansion of ΔG_y may be obtained using Tables 36/1,2. However, ΔG_z requires special consideration: the identity (35.2) still holds but in this case the terms for which ν is <u>even</u> are zero. Tables corresponding to 36/1,2 can be constructed, in which case there is a single group of second order equations, entailing 16 auxiliary quantities, and groups of four and six third order equations, entailing 52 auxiliary quantities. Thus the number of auxiliary quantities required are doubled, with a corresponding increase in subsequent numerical work.

37. The Second and Third Order Iteration Formulae

Express $\underline{\delta}_y$ and $\underline{\delta}_v$ as sums of terms of the <u>n</u>th order,

$$\underline{\delta}_y = \sum_{n=2}^{\infty} \underline{\delta}_y^{(n)} \quad , \quad \underline{\delta}_v = \sum_{n=2}^{\infty} \underline{\delta}_v^{(n)} \quad ,$$

and substitute these into (35.3). If the terms $(S_y+\delta_y)^{n-\mu}$, $(T_y+\delta_v)^{\mu-\nu}$, etc., are expanded, the <u>n</u>th order terms of (35.3) are obtained. In particular, for the second and third orders

$$\sum_{\mu\nu\tau}^{2} (q^*{}_{\mu\nu\tau}^{(2)} - \underline{\underline{q}}_{\mu\nu\tau}^{(2)}) S_y^{2-\mu} T_y^{\mu-\nu} S_z^{\nu-\tau} T_z^{\tau} \equiv 0 \quad ,$$

$$\sum_{\mu\nu\tau}^{3} (q^*{}_{\mu\nu\tau}^{(3)} - \underline{\underline{q}}_{\mu\nu\tau}^{(3)}) S_y^{3-\mu} T_y^{\mu-\nu} S_z^{\nu-\tau} T_z^{\tau} \equiv 2\underline{\underline{s}}_{q1}\delta_y^{(2)} S_y + \underline{\underline{s}}_{q2}(\delta_y^{(2)} T_y + \delta_v^{(2)} S_y) +$$
$$+ 2\underline{\underline{s}}_{q5}\delta_v^{(2)} T_y + 2\underline{\underline{s}}_{q8}\delta_z^{(2)} S_z + \underline{\underline{s}}_{q9}(\delta_z^{(2)} T_z + \delta_w^{(2)} S_z) + 2\underline{\underline{s}}_{q10}\delta_w^{(2)} T_z \quad . \quad (37.1)$$

From the first of these it follows immediately that the second order pseudo- and paracanonical coefficients of a second order quantity are identical

$$s^*_{q\alpha} = \underline{\underline{s}}_{q\alpha} \quad . \tag{37.2}$$

From (25.7), $\underset{\sim}{\delta}_y$ and $\underset{\sim}{\delta}_v$ have the second order expansions

$$\delta_y^{(2)} = -({}^{\backprime}S_{gy1b}S_y^{\,2} + {}^{\backprime}S_{gy2b}S_yT_y + {}^{\backprime}S_{gy5b}T_y^{\,2} + {}^{\backprime}S_{gy8b}S_z^{\,2} + {}^{\backprime}S_{gy9b}S_zT_z + {}^{\backprime}S_{gy10b}T_z^{\,2}) \, ,$$

$$\delta_v^{(2)} = {}^{\backprime}S_{gy1a}S_y^{\,2} + {}^{\backprime}S_{gy2a}S_yT_y + {}^{\backprime}S_{gy5a}T_y^{\,2} + {}^{\backprime}S_{gy8a}S_z^{\,2} + {}^{\backprime}S_{gy9a}S_zT_z + {}^{\backprime}S_{gy10a}T_z^{\,2} \, ,$$

$$\delta_z^{(2)} = -({}^{\backprime}S_{gz3b}S_yS_z + {}^{\backprime}S_{gz4b}S_yT_z + {}^{\backprime}S_{gz6b}T_yS_z + {}^{\backprime}S_{gz7b}T_yT_z) \, ,$$

$$\delta_w^{(2)} = {}^{\backprime}S_{gz3a}S_yS_z + {}^{\backprime}S_{gz4a}S_yT_z + {}^{\backprime}S_{gz6a}T_yS_z + {}^{\backprime}S_{gz7a}T_yT_z \quad . \tag{37.3}$$

After substituting (37.3) into (37.1), the required third order iteration equations are read off. Thus, the coefficient of $S_y^{\,3}$ gives

$$t^*_{q1} = \underline{\underline{t}}_{q1} - 2\,{}^{\backprime}S_{gy1b}\underline{\underline{s}}_{q1} + {}^{\backprime}S_{gy1a}\underline{\underline{s}}_{q2} \quad . \tag{37.4}$$

Proceeding in this manner the third order iteration equations are easily and systematically obtained. Those relevant to higher orders could, if desired, be obtained in exactly the same way, starting with the higher order terms of (35.3). The third order equations are given in Table 37/1 and are valid for any scalar quantity Q* of the second order in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. Note that both the $\underline{a}$- and $\underline{b}$-components of the second order $\underset{\sim}{G}$-coefficients turn up in the third order iteration equations.

TABLE 37/1 Third Order Iteration Formulae

$$t^*_{q1} = \underline{\underline{t}}_{q1} - 2\,{}^{\backprime}S_{gy1b}\underline{\underline{s}}_{q1} + {}^{\backprime}S_{gy1a}\underline{\underline{s}}_{q2}$$

$$t^*_{q2} = \underline{\underline{t}}_{q2} - 2\,{}^{\backprime}S_{gy2b}\underline{\underline{s}}_{q1} + ({}^{\backprime}S_{gy2a} - {}^{\backprime}S_{gy1b})\underline{\underline{s}}_{q2} + 2\,{}^{\backprime}S_{gy1a}\underline{\underline{s}}_{q5}$$

$$t^*_{q5} = \underline{\underline{t}}_{q5} - 2\,{}^{\backprime}S_{gy5b}\underline{\underline{s}}_{q1} + ({}^{\backprime}S_{gy5a} - {}^{\backprime}S_{gy2b})\underline{\underline{s}}_{q2} + 2\,{}^{\backprime}S_{gy2a}\underline{\underline{s}}_{q5}$$

$$t^*_{q8} = \underline{\underline{t}}_{q8} - 2\,{}^{\backprime}S_{gy8b}\underline{\underline{s}}_{q1} + {}^{\backprime}S_{gy8a}\underline{\underline{s}}_{q2} - 2\,{}^{\backprime}S_{gz3b}\underline{\underline{s}}_{q8} + {}^{\backprime}S_{gz3a}\underline{\underline{s}}_{q9}$$

$$t^*_{q9} = \underline{\underline{t}}_{q9} - 2\,{}^{\backprime}S_{gy9b}\underline{\underline{s}}_{q1} + {}^{\backprime}S_{gy9a}\underline{\underline{s}}_{q2} - 2\,{}^{\backprime}S_{gz4b}\underline{\underline{s}}_{q8} + ({}^{\backprime}S_{gz4a} - {}^{\backprime}S_{gz3b})\underline{\underline{s}}_{q9} + $$
$$+ 2\,{}^{\backprime}S_{gz3a}\underline{\underline{s}}_{q10}$$

$$t^*_{q10} = \underline{\underline{t}}_{q10} - 2\,{}^{\backprime}S_{gy10b}\underline{\underline{s}}_{q1} + {}^{\backprime}S_{gy10a}\underline{\underline{s}}_{q2} - {}^{\backprime}S_{gz4b}\underline{\underline{s}}_{q9} + 2\,{}^{\backprime}S_{gz4a}\underline{\underline{s}}_{q10}$$

$$t^*_{q11} = \underline{\underline{t}}_{q11} - {}^{\backprime}S_{gy5b}\underline{\underline{s}}_{q2} + 2\,{}^{\backprime}S_{gy5a}\underline{\underline{s}}_{q5}$$

$$t^*_{q14} = \underline{\underline{t}}_{q14} - {}^{\backprime}S_{gy8b}\underline{\underline{s}}_{q2} + 2\,{}^{\backprime}S_{gy8a}\underline{\underline{s}}_{q5} - 2\,{}^{\backprime}S_{gz6b}\underline{\underline{s}}_{q8} + {}^{\backprime}S_{gz6a}\underline{\underline{s}}_{q9}$$

$$t^*_{q15} = \underline{\underline{t}}_{q15} - {}^{\backprime}S_{gy9b}\underline{\underline{s}}_{q2} + 2\,{}^{\backprime}S_{gy9a}\underline{\underline{s}}_{q5} - 2\,{}^{\backprime}S_{gz7b}\underline{\underline{s}}_{q8} + ({}^{\backprime}S_{gz7a} - {}^{\backprime}S_{gz6b})\underline{\underline{s}}_{q9} + $$
$$+ 2\,{}^{\backprime}S_{gz6a}\underline{\underline{s}}_{q10}$$

$$t^*_{q16} = \underline{\underline{t}}_{q16} - {}^{\backprime}S_{gy10b}\underline{\underline{s}}_{q2} + 2\,{}^{\backprime}S_{gy10a}\underline{\underline{s}}_{q5} - {}^{\backprime}S_{gz7b}\underline{\underline{s}}_{q9} + 2\,{}^{\backprime}S_{gz7a}\underline{\underline{s}}_{q10}$$

38. Iteration for Quantities of the First Order and for Two-Vectors

(a) In the previous section the iteration equations relevant to a second order scalar, Q*, were derived. Although these are the ones which will be used in computing the coefficients, it is of interest to consider the iteration equations for Q, a scalar which is of the first order in $\underline{Y}$ and $\underline{V}$. In this case the summation over $\underline{n}$ in (35.3) is from $\underline{n} = 1$ to ∞. Corresponding to (37.1), the first and second order terms of (35.3) are

$$\sum_{\mu\nu\tau}^{1}(q_{\mu\nu\tau}{}^{(1)}-\underline{\underline{q}}_{\mu\nu\tau}{}^{(1)})S_y{}^{1-\mu}T_y{}^{\mu-\nu}S_z{}^{\nu-\tau}T_z{}^{\tau} \equiv 0 \quad ,$$

$$\sum_{\mu\nu\tau}^{2}(q_{\mu\nu\tau}{}^{(2)}-\underline{\underline{q}}_{\mu\nu\tau}{}^{(2)})S_y{}^{2-\mu}T_y{}^{\mu-\nu}S_z{}^{\nu-\tau}T_z{}^{\tau} \equiv \underline{\underline{p}}_{q1}\delta_y{}^{(2)} + \underline{\underline{p}}_{q2}\delta_v{}^{(2)} \quad . \qquad (38.1)$$

Thus, substituting (37.3) and selecting coefficients of like terms, the relevant iteration equations are

$$p_{q\alpha} = \underline{\underline{p}}_{q\alpha} \qquad (\alpha = 1,2)$$

$$s_{q\alpha} = \underline{\underline{s}}_{q\alpha} - {}^{`}S_{gy\alpha b}\underline{\underline{p}}_{q1} + {}^{`}S_{gy\alpha a}\underline{\underline{p}}_{q2} \qquad (\alpha = 1,2,5,8,9,10) \quad . \qquad (38.2)$$

Similarly, the first member of the third order group is

$$t_{q1} = \underline{\underline{t}}_{q1} - 2{}^{`}S_{gy1b}\underline{\underline{s}}_{q1} + {}^{`}S_{gy1a}\underline{\underline{s}}_{q2} - {}^{`}T_{gy1b}\underline{\underline{p}}_{q1} + {}^{`}T_{gy1a}\underline{\underline{p}}_{q2} \quad . \qquad (38.3)$$

Note that the first three terms on the right of (38.3) are identical with the right hand term of (37.4). In general, it is found that the paracanonical coefficient $q_{\mu\nu\tau j}{}^{(n)}$ is given by

$$q_{\mu\nu\tau j}{}^{(n)} = \underline{\underline{q}}_{\mu\nu\tau j}{}^{(n)} + \text{(function of all the } \underset{\sim}{G}_j\text{-coefficients of order} \leq \underline{n} \text{ and of the } \underline{\underline{q}}_j\text{-coefficients of orders} < \underline{n}) \quad . \qquad (38.4)$$

It is evident from (25.7) that the $\underset{\sim}{G}_j$-coefficients in (38.4) do not involve the $\underset{\sim}{g}_j$-coefficients. $\underset{\sim}{Y}_j$ and $\underset{\sim}{V}_j$ involve $\underset{\sim}{\delta}_{yj}$ and $\underset{\sim}{\delta}_{vj}$, for which the summation over $\underline{i}$ is from $\underline{i} = 1$ to $\underline{j} - 1$. If Q had been of the $\underline{m}$th order in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, only the $\underset{\sim}{G}$-coefficients of orders less than or equal to $\underline{n}-\underline{m}+1$ would appear in (38.4).

(b) The iteration formulae of Table 37/1 and §38a were obtained under the assumption that Q was invariant under reflection in Ⓜ. Now suppose the iteration formulae for Q_z, the z-component of $\underset{\sim}{Q}$, are required. The identity (35.3) will still hold, but with ν odd. Then, the first member of (38.1) is unchanged and the second becomes

$$\sum_{\mu\nu\tau}^{2}\left(q_{z\mu\nu\tau}^{(2)} - \underline{\underline{q}}_{z\mu\nu\tau}^{(2)}\right) S_y^{\,2-\mu} T_y^{\,\mu-\nu} S_z^{\,\nu-\tau} T_z^{\,\tau} \equiv \underline{\underline{p}}_{qz3}\delta_z^{(2)} + \underline{\underline{p}}_{qz4}\delta_w^{(2)} \ .$$

Thus

$$p_{qz\alpha} = \underline{\underline{p}}_{qz\alpha} \qquad (\alpha = 3,4) \ ,$$

$$s_{qz\alpha} = \underline{\underline{s}}_{qz\alpha} - \grave{S}_{gz\alpha b}\underline{\underline{p}}_{qz3} + \grave{S}_{gz\alpha a}\underline{\underline{p}}_{qz4} \qquad (\alpha = 3,4,6,7) \ . \qquad (38.5)$$

The second member of this is generically distinct from the second member of (38.2) - the $S_{gy\alpha}$ are replaced by $S_{gz\alpha}$ and there are only four equations in (38.5) but six in (38.2). Similarly, the higher order equations differ. This has its origin in the different symmetry properties of Q, in (38.2), and Q_z, in (38.5). It is evident, therefore, that if the surface expansion of $\Delta\underset{\sim}{G}$ is converted into the corresponding paracanonical expansion, two generically distinct sets of iteration equations are required, one for ΔG_y and the other for ΔG_z. Moreover, those relevant to ΔG_y are given in Table 37/1 and are applied twice - once for ΔG_{ya} and once for ΔG_{yb}. Thus, the use of those relevant to ΔG_z involves numerical computation over and above that required when the iteration is with P and R. Together with the discussion of §36c, this demonstrates that less computation is required if the paracanonical expansions of P and R are determined.

39. The q-coefficients

The surface expansions of P and R are known and it has been seen (§37) how the paracanonical expansions of P* and R* can be determined to any order (in principle). The next step (§28iv) is to determine the corresponding coefficients for P and R. As before let Q be any scalar function (of the canonical variables). Assume that the q*-coefficients up to the nth order have been found as in §36, 37. It is required to find the paracanonical coefficients of order (n-1) of Q = Q* + q. This may be accomplished in one of two ways. On the one hand the iteration equations relevant to Q (§38a) could be used and on the other hand $\hat{q}^{(1)}$ could be expressed as a function of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Since the q*-coefficients are known, it is most convenient to take the second alternative and start with the surface expansion of $\hat{q}^{(1)}$,

$$\hat{q}^{(1)} = \hat{p}_{q1} Y + \hat{p}_{q2} V \quad ,$$

and express Y and V in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ by means of (25.6). Accordingly, it is found that

$$\hat{q}^{(1)} = \underline{\underline{p}}_{q1} S_y + \underline{\underline{p}}_{q2} T_y + \sum_{n=2}^{\infty} \sum_{\mu\nu\tau}^{n} (\underline{\underline{p}}_{q2}\, {}^{\backprime}G_{y\mu\nu\tau a}{}^{(n)} - \underline{\underline{p}}_{q1}\, {}^{\backprime}G_{y\mu\nu\tau b}{}^{(n)}) S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} . \tag{39.1}$$

Thus, from (26.4) and (24.1), the q-coefficients are given by

$$p_{q1} = \underline{\underline{p}}_{q1} \quad , \quad p_{q2} = \underline{\underline{p}}_{q2}$$

$$q_{\mu\nu\tau}{}^{(n)} = q^*_{\mu\nu\tau}{}^{(n)} + \underline{\underline{p}}_{q2}\, {}^{\backprime}G_{y\mu\nu\tau a}{}^{(n)} - \underline{\underline{p}}_{q1}\, {}^{\backprime}G_{y\mu\nu\tau b}{}^{(n)} \quad . \tag{39.2}$$

If the q*-coefficients from (37.2) and Table 37/1 are substituted into (39.2), equations (38.2,3) are recovered and, as would be expected, iteration with Q* followed by use of (39.2) is equivalent to iteration with Q. Note: The $\underset{\sim}{G}$-coefficients of the ($\underline{n}$-1)th order are required in order to determine the $\underline{n}$th order q*-coefficients (§37) and knowledge of these $\underset{\sim}{G}$-coefficients enables the ($\underline{n}$-1)th order q-coefficients to be determined from (39.2).

40. The $\underset{\sim}{i}$-coefficients. The Vectors $\underset{\sim}{P}$ and $\underset{\sim}{R}$.

(a) P, R, P* and R* have all been determined to the required order in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Before $\underset{\sim}{P}$ and $\underset{\sim}{R}$ can be determined, one further quantity, at present known only in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, must be found to the ($\underline{n}$-1)th order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$. This is $\underset{\sim}{I}$, given by (10.4). From (25.6) the $\underline{n}$th order terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ are

$$\underset{\sim}{y}^{(n)} = \underset{\sim}{y}_a \underset{\sim}{\delta}_y{}^{(n)} + \underset{\sim}{y}_b \underset{\sim}{\delta}_v{}^{(n)} \quad , \quad \underset{\sim}{v}^{(n)} = \underset{\sim}{v}_a \underset{\sim}{\delta}_y{}^{(n)} + \underset{\sim}{v}_b \underset{\sim}{\delta}_v{}^{(n)} \quad , \tag{40.1}$$

and, from (10.4), the corresponding terms of $\underset{\sim}{I}$ are

$$\underset{\sim}{i}^{(n)} = c\underset{\sim}{y}^{(n)} + n_{Bx}\underset{\sim}{v}^{(n)} = \underset{\sim}{i}_a \underset{\sim}{\delta}_y{}^{(n)} + \underset{\sim}{i}_b \underset{\sim}{\delta}_v{}^{(n)} \quad , \tag{40.2}$$

where the <u>parabasal $\underset{\sim}{i}$-coefficients</u> are

$$\underset{\sim}{i}_a = c\underset{\sim}{y}_a + n_{Bx}\underset{\sim}{v}_a \quad , \quad \underset{\sim}{i}_b = c\underset{\sim}{y}_b + n_{Bx}\underset{\sim}{v}_b \quad . \tag{40.3}$$

It follows from (40.1,2) and (25.7) that

$$\underset{\sim}{y}^{(n)} = \sum_i (\Delta \grave{G}_i{}^{(n)} | \underset{\sim}{y}) \quad , \quad \underset{\sim}{v}^{(n)} = \sum_i (\Delta \grave{G}_i{}^{(n)} | \underset{\sim}{v}) \quad , \quad \underset{\sim}{i}^{(n)} = \sum_i (\Delta \grave{G}_i{}^{(n)} | \underset{\sim}{i}) \quad . \tag{40.4}$$

Thus, substituting the paracanonical expansions of $\Delta\underset{\sim}{G}$,

$$\underset{\sim}{y}_{\mu\nu\tau}{}^{(n)} = (\grave{\underset{\sim}{G}}_{\mu\nu\tau}{}^{(n)} | \underset{\sim}{y}) \quad , \quad \underset{\sim}{v}_{\mu\nu\tau}{}^{(n)} = (\grave{\underset{\sim}{G}}_{\mu\nu\tau}{}^{(n)} | \underset{\sim}{v}) \quad ,$$

$$\underset{\sim}{i}_{\mu\nu\tau}{}^{(n)} = c\underset{\sim}{y}_{\mu\nu\tau}{}^{(n)} + n_{Bx}\underset{\sim}{v}_{\mu\nu\tau}{}^{(n)} = (\grave{\underset{\sim}{G}}_{\mu\nu\tau}{}^{(n)} | \underset{\sim}{i}) \quad , \tag{40.5}$$

and the $\underset{\sim}{y}$-, $\underset{\sim}{v}$- and $\underset{\sim}{i}$-coefficients are known to the same order as the $\underset{\sim}{G}$-coefficients.

(b) The stage has now been reached at which the paracanonical expansions of $\underset{\sim}{P}$ and $\underset{\sim}{R}$ can be determined. As usual let Q be a scalar, in particular, P or R. $\underset{\sim}{Q}$ is defined in terms of Q and $\underset{\sim}{I}$ by (26.5), from which it follows that

$$\underset{\sim}{q}^{(n)} = \sum_{r=1}^{n} q^{(n-r)} \underset{\sim}{i}^{(r)} + q^{*(n)} \underset{\sim}{I}_{By} \quad , \tag{40.6}$$

and in particular,

$$\underset{\sim}{q}^{(2)} = q^{(1)} \underset{\sim}{i}^{(1)} + q^{*(2)} \underset{\sim}{I}_{By} \quad ,$$

$$\underset{\sim}{q}^{(3)} = q^{(2)} \underset{\sim}{i}^{(1)} + q^{(1)} \underset{\sim}{i}^{(2)} + q^{*(3)} \underset{\sim}{I}_{By} \quad . \tag{40.7}$$

TABLE 40/1 Second Order $\underset{\sim}{q}$-coefficients

$s_{qy1} = s^*_{q1} I_{By} + i_{ya} p_{q1}$

$s_{qy2} = s^*_{q2} I_{By} + i_{ya} p_{q2} + i_{yb} p_{q1}$

$s_{qy5} = s^*_{q5} I_{By} + i_{yb} p_{q2}$

$s_{qy8} = s^*_{q8} I_{By}$

$s_{qy9} = s^*_{q9} I_{By}$

$s_{qy10} = s^*_{q10} I_{By}$

$s_{qz3} = p_{q1} i_{za}$

$s_{qz4} = p_{q1} i_{zb}$

$s_{qz6} = p_{q2} i_{za}$

$s_{qz7} = p_{q2} i_{zb}$

TABLE 40/2 Third Order $\underset{\sim}{q}$-coefficients

$t_{qy1} = t^*_{q1} I_{By} + s_{q1} i_{ya} + p_{q1} s_{iy1}$

$t_{qy2} = t^*_{q2} I_{By} + s_{q2} i_{ya} + s_{q1} i_{yb} + p_{q1} s_{iy2} + p_{q2} s_{iy1}$

$t_{qy5} = t^*_{q5} I_{By} + s_{q5} i_{ya} + s_{q2} i_{yb} + p_{q1} s_{iy5} + p_{q2} s_{iy2}$

$t_{qy8} = t^*_{q8} I_{By} + s_{q8} i_{ya} + p_{q1} s_{iy8}$

$t_{qy9} = t^*_{q9} I_{By} + s_{q9} i_{ya} + p_{q1} s_{iy9}$

$t_{qy10} = t^*_{q10} I_{By} + s_{q10} i_{ya} + p_{q1} s_{iy10}$

$t_{qy11} = t^*_{q11} I_{By} + s_{q5} i_{yb} + p_{q2} s_{iy5}$

$t_{qy14} = t^*_{q14} I_{By} + s_{q8} i_{yb} + p_{q2} s_{iy8}$

$t_{qy15} = t^*_{q15} I_{By} + s_{q9} i_{yb} + p_{q2} s_{iy9}$

$t_{qy16} = t^*_{q16} I_{By} + s_{q10} i_{yb} + p_{q2} s_{iy10}$

$t_{qz3} = s_{q1} i_{za} + p_{q1} s_{iz3}$

$t_{qz4} = s_{q1} i_{zb} + p_{q1} s_{iz4}$

$t_{qz6} = s_{q2} i_{za} + p_{q1} s_{iz6} + p_{q2} s_{iz3}$

$t_{qz7} = s_{q2} i_{zb} + p_{q1} s_{iz7} + p_{q2} s_{iz4}$

$t_{qz12} = s_{q5} i_{za} + p_{q2} s_{iz6}$

$t_{qz13} = s_{q5} i_{zb} + p_{q2} s_{iz7}$

$t_{qz17} = s_{q8} i_{za}$

$t_{qz18} = s_{q9} i_{za} + s_{q8} i_{zb}$

$t_{qz19} = s_{q10} i_{za} + s_{q9} i_{zb}$

$t_{qz20} = s_{q10} i_{zb}$

By substituting the series for $q^{*(n)}$, $q^{(n)}$ and $\underset{\sim}{i}^{(n)}$ and selecting the coefficients of like terms, the second and third order $\underset{\sim}{q}$-coefficients may be read off (40.7) and are given in Tables 40/1,2. Note that although Q is a scalar, $\underset{\sim}{Q}$ is a two-vector. Hence the sets of equations for the q_y- and q_z-coefficients are generically distinct.

41. The $\underset{\sim}{g}$-coefficients. Aberration Coefficients.

(a) It is now a simple matter to obtain the $\underset{\sim}{g}$-coefficients. (26.9) gives immediately

$$\underset{\sim}{g}_{\mu\nu\tau}{}^{(n)} = \bar{\underset{\sim}{v}}\underset{\sim}{p}_{\mu\nu\tau}{}^{(n)} + \bar{\underset{\sim}{y}}\underset{\sim}{r}_{\mu\nu\tau}{}^{(n)} \quad . \tag{41.1}$$

To obtain the g_{ya}-, g_{yb}-, g_{za}- or g_{zb}-coefficients, it is sufficient to replace $\bar{\underset{\sim}{v}}$ and $\bar{\underset{\sim}{y}}$ by $\bar{v}_{ya}$ and $\bar{y}_{ya}$, $\bar{v}_{yb}$ and $\bar{y}_{yb}$, $\bar{v}_{za}$ and $\bar{y}_{za}$ or $\bar{v}_{zb}$ and $\bar{y}_{zb}$ respectively, where these may be obtained from (26.8). For example

$$g_{y\mu\nu\tau a}{}^{(n)} = -N_y{}'\bar{R}_B(v_a{}'p_{y\mu\nu\tau}{}^{(n)} + y_a{}'r_{y\mu\nu\tau}{}^{(n)}) \quad .$$

In this manner the $\underline{a}$- and $\underline{b}$-components of the g_y- and g_z-coefficients are determined for all required orders. The corresponding intermediate or $\underset{\sim}{G}$-coefficients follow from summation over the surfaces:

$$\underset{\sim}{G}_{\mu\nu\tau aj}{}^{(n)} = \sum_{i=1}^{j-1}\underset{\sim}{g}_{\mu\nu\tau ai}{}^{(n)} \quad , \quad \underset{\sim}{G}_{\mu\nu\tau bj}{}^{(n)} = \sum_{i=1}^{j-1}\underset{\sim}{g}_{\mu\nu\tau bi}{}^{(n)} \quad . \tag{41.2}$$

If the $\underset{\sim}{G}_{\mu\nu\tau j}{}^{(n)}$' are required, the summation is taken from $\underline{i} = 1$ to $\underline{j}$. Once the $\underset{\sim}{G}$-coefficients are known the $\underline{n}$th order $\underset{\sim}{y}$- and $\underset{\sim}{v}$-coefficients follow from (40.5).

The surface and pseudo-coefficients of any quantity can be determined to an arbitrary order without the knowledge of any paracanonical coefficients. However, in deriving the $\underline{n}$th order iteration equations, it was assumed that the $\underset{\sim}{G}$-coefficients were known up to the ($\underline{n}$-1)th order.

Under this assumption, the p- and r-coefficients up to the ($\underline{n}$-1)th order are obtainable from the p*- and r*-coefficients. Moreover, the $\underset{\sim}{i}$-coefficients of the first ($\underline{n}$-1) orders are expressed in terms of the known $\underset{\sim}{G}$-coefficients and hence the $\underline{n}$th order $\underset{\sim}{p}$- and $\underset{\sim}{r}$-coefficients may be determined. From these follow the $\underset{\sim}{g}$- and $\underset{\sim}{G}$-coefficients of the $\underline{n}$th order. Thus, the results of this chapter verify in detail that given the ($\underline{n}$-1)th order $\underset{\sim}{G}$-coefficients, those of the $\underline{n}$th order may be found. Finally, the $\underset{\sim}{G}$-coefficients of the first order at all surfaces, and of all orders at the first surface, are trivially zero. Thus the $\underset{\sim}{G}$-coefficients may indeed be determined by the iterative process sketched in §28.

(b) Now that the $\underset{\sim}{G}$-coefficients are known the image height or $\underset{\sim}{h}$-coefficients may be determined for any position of the image plane from (24.11), and the corresponding surface contributions by (24.12). Suppose, for the sake of illustration, that imagery is to be distortion free and the aberration is defined by

$$\underset{\sim}{\epsilon}_k{}' = \underset{\sim}{H}_k{}' - \underset{\sim}{h}_{bk}{}'\underset{\sim}{T} \quad . \tag{41.3}$$

Then

$$\underset{\sim}{\epsilon}_k{}' = \underset{\sim}{h}_{ak}{}'\underset{\sim}{S} + \sum_{n=2}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{\sim}{h}_{\mu\nu\tau k}{}^{(n)\prime} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} \quad ,$$

and the <u>aberration coefficients</u> are, in this case, defined to be

$$\underset{\sim}{E}_{ak}{}' = \underset{\sim}{h}_{ak}{}' \quad , \quad \underset{\sim}{E}_{\mu\nu\tau k}{}^{(n)\prime} = \underset{\sim}{h}_{\mu\nu\tau k}{}^{(n)\prime} \quad . \tag{41.4}$$

The corresponding surface increments are

$$\underset{\sim}{e}_{\mu\nu\tau i}{}^{(n)\prime} = \Delta\underset{\sim}{E}_{\mu\nu\tau i}{}^{(n)} = \frac{1}{\underset{\sim}{\mu}_k{}'} \underset{\sim}{g}_{\mu\nu\tau ai}{}^{(n)} - \underset{\sim}{x}_{Mk}{}'\Delta\underset{\sim}{v}_{\mu\nu\tau i}{}^{(n)} \quad . \tag{41.5}$$

Surface contributions to $\underset{\sim}{E}_{ak}{}'$ are <u>not</u> considered, since $\underset{\sim}{h}_{ak}{}' = -\underset{\sim}{x}_{Mk}{}'\underset{\sim}{v}_{ak}{}'$ and the condition for stigmatic parabasal imagery is $y_a/v_a = z_a/w_a$, <u>not</u> $\underset{\sim}{v}_{ak}{}' = 0$.

42. Summary of a Method for Computing Basal Coefficients

Appendix F consists of a listing of a working computer programme (programme TIRIKI) for computing the basal coefficients of the first three orders referred to a meridional base-ray in a symmetric system of spherical surfaces. The programme incorporates its own brief explanatory notes and these are supplemented by subsections (a) to (g) below. Programme TIRIKI differs only slightly from the scheme outlined here. After each subsection the names of the relevant subroutines are given in parenthesis. The use of SPC is presupposed.

(a) The Trace of the Base-Ray: For a symmetric system of spherical surfaces Ford's basic ray trace scheme[1] is admirably suited to the task of tracing base-rays. In some instances the computation of pseudo-parameters is accomplished automatically in the course of the trace, for instance, $\overleftarrow{R}_B$, n_{Bx}, $\underset{\sim}{V}_B$ and $\underset{\sim}{I}_B$. It is a simple matter to determine $\vec{\beta}_B$ from $\underset{\sim}{V}_B$. Further simple calculations yield $\underset{\sim}{n}_B$, and hence cos I_B and cos I_B' , and d' . (BASRAY)

(b) Determination of the Parabasal Coefficients: This has already been discussed in §15. The parabasal $\underset{\sim}{i}$-coefficients, the augmenting factors $\underset{\sim}{\mu}_k'$ and the displacement $\underset{\sim}{x}_{Mk}'$ of the meridional and sagittal foci from some suitably chosen image plane[2] may well be computed here. This is accomplished by equations (40.3), (22.7) and

---

1 P. W. Ford (1960). If the system is decentred, the scheme of Appendix C may be used.

2 For a symmetric system, a suitable initial choice of receiving plane is the paraxial ideal image plane.

$$\ell_M = -y_a/v_a \quad , \quad \ell_S = -z_a/w_a \quad , \tag{42.1}$$

respectively. (PARBAS)

(c) Determination of the $\hat{r}$- and $\hat{p}$-coefficients: First the quantities $\lambda_1,\ldots,\lambda_4$ are computed from (13.2) and (32.3) and the $\hat{x}$-coefficients from Table 31/1. μ_0 and τ_0 follow from (32.3,4) and the few non-zero $\hat{\mu}$-coefficients from (32.5). The $\hat{\sigma}$- and $\hat{\tau}$-coefficients are calculated from Table 32/1 and provided the $\hat{\psi}$- and $\hat{\Phi}$-coefficients are first computed from (30.6), an application of Table 30/1 yields the $\hat{r}$-coefficients. Finally, the $\hat{p}$-coefficients follow from the $\hat{x}$- and $\hat{r}$-coefficients by means of Table 33/1. (XST123)

(d) Computation of the Pseudo-Coefficients: The auxiliary quantities of Table 36/1 are computed and used to determine the $\underline{\underline{p}}$- and $\underline{\underline{r}}$-coefficients according to the examples of §36a. (PSEUDO)

(e) Calculation of the Second Order $\underset{\sim}{g}$-coefficients: So that the $\underset{\sim}{g}$-coefficients can be calculated by forming linear combinations of the $\underset{\sim}{r}$- and $\underset{\sim}{p}$-coefficients, the $\bar{\underset{\sim}{v}}$- and $\bar{\underset{\sim}{y}}$-coefficients are determined from (26.8,11). The second order p*- and r*-coefficients are identically equal to the corresponding pseudo-coefficients (37.2) and the first order p- and r-coefficients are identical with the first order $\underline{\underline{p}}$- and $\underline{\underline{r}}$ coefficients (39.2). The second order $\underset{\sim}{r}$- and $\underset{\sim}{p}$-coefficients then follow from Table 40/1 and the second order $\underset{\sim}{g}$-coefficients are computed by forming the appropriate linear combinations of the corresponding $\underset{\sim}{r}$- and $\underset{\sim}{p}$- coefficients according to (41.1). Summation gives the $\underset{\sim}{G}$-coefficients (41.2). (GCOEFS)

(f) Determination of the Third Order g-coefficients: The iteration equations of Table 37/1 must be applied to the third order p- and r-coefficients in order to compute the corresponding p*- and r*-coefficients. Next, (40.5) yields the second order i-coefficients and the second order r- and p-coefficients follow from (39.2). Table 40/2 is then applied in order to compute the third order p- and r-coefficients and the third order g-coefficients follow from (41.1). The G-coefficients are again found by summation. (GCOEFS, ITRATE and ST2VEC)

(g) Determination of the Image Height Coefficients: Finally, the h-coefficients in the presupposed image plane are determined by first computing the v-coefficients according to (40.5) and then the h-coefficients from (24.11). If desired, the surface contributions to the latter can be found from (24.12). (HDCOEF)

VII. THE EXTREMAL IDENTITIES

43. The Existence of Identities

In so much as they determine $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, the $\underset{\sim}{G}$-coefficients completely determine the behaviour of Ⓚ. The aberration coefficients determine only the point of intersection of a ray with the chosen image plane. If the $\underset{\sim}{v}$-coefficients are also known, the point of intersection of the ray with any plane may be computed and the ray is uniquely determined (cf. the case of SPC where rays were specified by their points of intersection with two planes). Thus the aberration and $\underset{\sim}{v}$-coefficients together determine the behaviour of Ⓚ. However, since these coefficients are expressed in terms of the $\underset{\sim}{G}$-coefficients which arise from quantities of fundamental importance in the theory, that is, the quasi-invariants, the $\underset{\sim}{G}$-coefficients are regarded as the basic set of coefficients completely determining the behaviour of Ⓚ. It has been seen (§25) that there are

$$\underline{\underline{N}}_{gn} = \tfrac{1}{3}(n+1)(n+2)(n+3) \qquad (43.1)$$

$\underline{n}$th order[1] $\underset{\sim}{G}$-coefficients, including both $\underline{a}$- and $\underline{b}$-components.

The $\underline{n}$th order imagery of any system is determined by the coefficients of the terms of the ($\underline{n}$+1)th order appearing in a characteristic function of the system. If the system is plane symmetric, and the base-ray lies in the plane of symmetry, the number of such coefficients is given by $\underline{\underline{N}}_{y(n+1)}$. If this is compared with (43.1) it is seen that $\underline{\underline{N}}_{gn} > \underline{\underline{N}}_{y(n+1)}$ for all $n \geq 1$. Thus the $\underline{\underline{N}}_{gn}$ $\underset{\sim}{G}$-coefficients of the $\underline{n}$th order are not all mutually independent and there exist $\underline{\underline{N}}_{In}$ identities between them where

---

1 When n = 1, imagery is characterized by the eight parabasal coefficients, a number correctly given by (43.1).

$$\underset{=}{N}_{In} = \underset{=}{N}_{gn} - \underset{=}{N}_{y(n+1)} = \tfrac{1}{4}(n+3)(n^2+2n-1) \qquad \text{when n is odd} \quad ,$$

$$= \tfrac{1}{4}n(n+2)(n+3) \qquad \text{when n is even} \quad . \qquad (43.2)$$

As in the axial theory, these identities, here called the extremal identities, may be determined from six integrability conditions obtained from an extremal condition.[2] The extremal condition is in effect Fermat's principle. These identities play no direct role in the development of the theory, or in the computation of coefficients. Rather, their value lies in the fact that they are derived from a principle completely outside the theory and hence can be used first as a very valuable numerical check on the algebra of the theory and later as an aid to the "debugging" of the programme for computing the coefficients.

44. The Integrability Conditions

With the x-planes of the first and jth surfaces as base planes, let the point characteristic of the plane symmetric system Ⓚ be $\underset{=}{V}$. If the base points are taken to be the points of intersection of the base-ray with the first and jth surfaces,

$$d\underset{=}{V} = \Delta_j(N\underset{\sim}{\beta}.d\underset{\sim}{Y}) \quad , \qquad (44.1)$$

is a total differential irrespective of the independent variables where, for any quantity Q, $\Delta_j Q$ is defined to be $Q_j - Q_1$.[1] The set of independent variables will be taken to be $\underset{\sim}{S}$ and $\underset{\sim}{T}$. If the partial derivatives of Q with respect to S_y, S_z, T_y and T_z are respectively denoted by $Q_{,y}$, $Q_{,z}$,

---

2 See, for instance, M§22b and Buchdahl (1965) §9.

1 Also, write $\Delta_j{}'Q = Q_j{}' - Q_1$. Note that $\Delta_j Q = \sum_{i=1}^{j-1}\Delta Q_i$.

$Q_{,v}$ and $Q_{,w}$, the differential dQ is given by

$$dQ = Q_{,y}dS_y + Q_{,v}dT_y + Q_{,z}dS_z + Q_{,w}dT_z \quad .$$

Applying this to the total differential $d\underline{\underline{V}}$, the four equations

$$\underline{\underline{V}}_{,y} = \Delta_j N\underset{\sim}{\beta}\cdot\underset{\sim}{Y}_{,y} = N_j(\beta_j Y_{j,y}+\gamma_j Z_{j,y}) - N_1(\beta_1 Y_{1,y}+\gamma_1 Z_{1,y}) \quad ,$$

$$\underline{\underline{V}}_{,v} = \Delta_j(N\underset{\sim}{\beta}\cdot\underset{\sim}{Y}_{,v}) \quad , \quad \underline{\underline{V}}_{,z} = \Delta_j(N\underset{\sim}{\beta}\cdot\underset{\sim}{Y}_{,z}) \quad , \quad \underline{\underline{V}}_{,w} = \Delta_j(N\underset{\sim}{\beta}\cdot\underset{\sim}{Y}_{,w}) \quad , \tag{44.2}$$

are obtained. Now, for any pair of distinct indices r and s taken from y, v, z and w, $\underline{\underline{V}}_{,rs} = \underline{\underline{V}}_{,sr}$. There are six such equations and these are just the six identities from which the extremal identities are found. In detail

$$\Delta_j N(\underset{\sim}{\beta}_{,y}\cdot\underset{\sim}{Y}_{,v}-\underset{\sim}{\beta}_{,v}\underset{\sim}{Y}_{,y}) = 0 \quad , \quad \Delta_j N(\underset{\sim}{\beta}_{,y}\cdot\underset{\sim}{Y}_{,z}-\underset{\sim}{\beta}_{,z}\cdot\underset{\sim}{Y}_{,y}) = 0 \quad ,$$

$$\Delta_j N(\underset{\sim}{\beta}_{,y}\cdot\underset{\sim}{Y}_{,w}-\underset{\sim}{\beta}_{,w}\underset{\sim}{Y}_{,y}) = 0 \quad , \quad \Delta_j N(\underset{\sim}{\beta}_{,v}\cdot\underset{\sim}{Y}_{,z}-\underset{\sim}{\beta}_{,z}\cdot\underset{\sim}{Y}_{,v}) = 0 \quad ,$$

$$\Delta_j N(\underset{\sim}{\beta}_{,v}\cdot\underset{\sim}{Y}_{,w}-\underset{\sim}{\beta}_{,w}\underset{\sim}{Y}_{,v}) = 0 \quad , \quad \Delta_j N(\underset{\sim}{\beta}_{,z}\cdot\underset{\sim}{Y}_{,w}-\underset{\sim}{\beta}_{,w}\cdot\underset{\sim}{Y}_{,z}) = 0 \quad . \tag{44.3}$$

The direction cosines may be expressed as functions of the basal variables V and W, and thus both $\underset{\sim}{\beta}$ and $\underset{\sim}{Y}$ are in principle known as para-canonical expansions. Substitution of these into (44.3) yields the required identities between the $\underset{\sim}{G}$-coefficients. Note: of (44.3), the first and last are even in Z and W, whereas the rest are odd. Thus $\frac{1}{2}n(n+1)(n+2)$ identities between the $\underline{n}$th order $\underset{\sim}{G}$-coefficients are obtained from (44.3) when $\underline{n}$ is even, or $\frac{1}{2}(n+1)(n^2+2n-1)$ when $\underline{n}$ is odd. Clearly this is in excess of the number given in §43 and the identities obtained from (44.3) are not mutually independent. It can be checked by inspection of the identities themselves that only the required number are indeed independent.

The algebra required to derive the extremal identities is tedious but systematic and it is sufficient to give the details of the derivation of a single identity for each of the required orders (§46), the remainder being derived in the same manner. All of the identities of the first three orders are given in §47.

45. Expansions of α, β and γ

The direction cosines of an arbitrary ray must be expressed in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$. This is done in the usual manner by first expanding in terms of the canonical variables and then using (25.6). From the definition of $\underset{\sim}{V}$ it follows that

$$\alpha = \alpha_B(1+2\alpha_B\beta_B V+\alpha_B{}^2\underset{\sim}{V}.\underset{\sim}{V})^{-\frac{1}{2}} \quad , \quad \beta = \alpha(V_B+V) \quad , \quad \gamma = \alpha W \quad .$$

For the first three orders the non-vanishing surface coefficients of α are

$$\hat{p}_{\alpha 2} = -\alpha_B{}^2\beta_B \quad , \qquad \hat{s}_{\alpha 5} = \tfrac{1}{2}\alpha_B{}^3(3\beta_B{}^2-1) \quad , \quad \hat{s}_{\alpha 10} = -\tfrac{1}{2}\alpha_B{}^3 \quad ,$$
$$\hat{t}_{\alpha 11} = \tfrac{1}{2}\alpha_B\beta_B(3-5\beta_B{}^2) \quad , \quad \hat{t}_{\alpha 16} = \tfrac{3}{2}\alpha_B{}^4\beta_B \quad . \tag{45.1}$$

Thus, the non-vanishing surface coefficients of β and γ are

$$\hat{p}_{\beta 2} = \alpha_B{}^3 \quad , \qquad \hat{p}_{\gamma 4} = \alpha_B \quad ,$$
$$\hat{s}_{\beta 5} = -\tfrac{3}{2}\alpha_B{}^4\beta_B \quad , \qquad \hat{s}_{\gamma 7} = -\alpha_B{}^2\beta_B \quad ,$$
$$\hat{s}_{\beta 10} = -\tfrac{1}{2}\alpha_B{}^2\beta_B \quad , \qquad \hat{t}_{\gamma 13} = \tfrac{1}{2}\alpha_B{}^3(3\beta_B{}^2-1) \quad ,$$
$$\hat{t}_{\beta 11} = \tfrac{1}{2}\alpha_B{}^5(5\beta_B{}^2-1) \quad , \quad \hat{t}_{\gamma 20} = -\tfrac{1}{2}\alpha_B{}^3 \quad ,$$
$$\hat{t}_{\beta 16} = \tfrac{1}{2}\alpha_B{}^3(3\beta_B{}^2-1) \quad . \tag{45.2}$$

It is convenient to write

$$\beta = \beta_B + \alpha_B{}^3 V + \beta^* \quad , \quad \gamma = \alpha_B W + \gamma^* \quad , \tag{45.3}$$

and determine the paracanonical expansions of β^* and γ^* rather than β and

TABLE 45/1 The β*-coefficients

$$t^*_{\beta 1} = \hat{t}_{\beta 11} v_a^3 + 2\hat{s}_{\beta 5} v_a s_{v1}$$

$$t^*_{\beta 2} = 3\hat{t}_{\beta 11} v_a^2 v_b + 2\hat{s}_{\beta 5}(v_a s_{v2} + v_b s_{v1})$$

$$t^*_{\beta 5} = 3\hat{t}_{\beta 11} v_a v_b^2 + 2\hat{s}_{\beta 5}(v_a s_{v5} + v_b s_{v2})$$

$$t^*_{\beta 8} = \hat{t}_{\beta 16} v_a w_a^2 + 2\hat{s}_{\beta 5} v_a s_{v8} + 2\hat{s}_{\beta 10} w_a s_{w3}$$

$$t^*_{\beta 9} = 2\hat{t}_{\beta 16} v_a w_a w_b + 2\hat{s}_{\beta 5} v_a s_{v9} + 2\hat{s}_{\beta 10}(w_a s_{w4} + w_b s_{w3})$$

$$t^*_{\beta 10} = \hat{t}_{\beta 16} v_a w_b^2 + 2\hat{s}_{\beta 5} v_a s_{v10} + 2\hat{s}_{\beta 10} w_b s_{w4}$$

$$t^*_{\beta 11} = \hat{t}_{\beta 11} v_b^3 + 2\hat{s}_{\beta 5} v_b s_{v5}$$

$$t^*_{\beta 14} = \hat{t}_{\beta 16} v_b w_a^2 + 2\hat{s}_{\beta 5} v_b s_{v8} + 2\hat{s}_{\beta 10} w_a s_{w6}$$

$$t^*_{\beta 15} = 2\hat{t}_{\beta 16} v_b w_a w_b + 2\hat{s}_{\beta 5} v_b s_{v9} + 2\hat{s}_{\beta 10}(w_a s_{w7} + w_b s_{w6})$$

$$t^*_{\beta 16} = \hat{t}_{\beta 16} v_b w_b^2 + 2\hat{s}_{\beta 5} v_b s_{v10} + 2\hat{s}_{\beta 10} w_b s_{w7}$$

$$s^*_{\beta 1} = \hat{s}_{\beta 5} v_a^2$$

$$s^*_{\beta 2} = 2\hat{s}_{\beta 5} v_a v_b$$

$$s^*_{\beta 5} = \hat{s}_{\beta 5} v_b^2$$

$$s^*_{\beta 8} = \hat{s}_{\beta 10} w_a^2$$

$$s^*_{\beta 9} = 2\hat{s}_{\beta 10} w_a w_b$$

$$s^*_{\beta 10} = \hat{s}_{\beta 10} w_b^2$$

$$t^*_{\gamma 3} = \hat{t}_{\gamma 13} v_a^2 w_a + \hat{s}_{\gamma 7}(v_a s_{w3} + w_a s_{v1})$$

$$t^*_{\gamma 4} = \hat{t}_{\gamma 13} v_a^2 w_b + \hat{s}_{\gamma 7}(v_a s_{w4} + w_b s_{v1})$$

$$t^*_{\gamma 6} = 2\hat{t}_{\gamma 13} v_a v_b w_a + \hat{s}_{\gamma 7}(v_a s_{w6} + v_b s_{w3} + w_a s_{v2})$$

$$t^*_{\gamma 7} = 2\hat{t}_{\gamma 13} v_a v_b w_b + \hat{s}_{\gamma 7}(v_a s_{w7} + v_b s_{w4} + w_b s_{v2})$$

$$t^*_{\gamma 12} = \hat{t}_{\gamma 13} v_b^2 w_a + \hat{s}_{\gamma 7}(w_a s_{v5} + v_b s_{w6})$$

$$t^*_{\gamma 13} = \hat{t}_{\gamma 13} v_b^2 w_b + \hat{s}_{\gamma 7}(w_b s_{v5} + v_b s_{w7})$$

$$t^*_{\gamma 17} = \hat{t}_{\gamma 20} w_a^3 + \hat{s}_{\gamma 7} w_a s_{v8}$$

$$t^*_{\gamma 18} = 3\hat{t}_{\gamma 20} w_a^2 w_b + \hat{s}_{\gamma 7}(w_b s_{v8} + w_a s_{v9})$$

$$t^*_{\gamma 19} = 3\hat{t}_{\gamma 20} w_a w_b^2 + \hat{s}_{\gamma 7}(w_b s_{v9} + w_a s_{v10})$$

$$t^*_{\gamma 20} = \hat{t}_{\gamma 20} w_b^3 + \hat{s}_{\gamma 7} w_b s_{v10}$$

$$s^*_{\gamma 3} = \hat{s}_{\gamma 7} v_a w_a$$

$$s^*_{\gamma 4} = \hat{s}_{\gamma 7} v_a w_b$$

$$s^*_{\gamma 6} = \hat{s}_{\gamma 7} v_b w_a$$

$$s^*_{\gamma 7} = \hat{s}_{\gamma 7} v_b w_b$$

γ. This is accomplished by substituting for $\underset{\sim}{V}$ in $\underset{\sim}{\beta}^*$ according to (25.6). The resulting $\underset{\sim}{\beta}^*$-coefficients are given in Table 45/1.

46. Determination of the Extremal Identities

To illustrate the derivation of the extremal identities, one identity between each of the first, second and third order coefficients will be obtained from the first member of (44.3). If $\underset{\sim}{\beta}$ and $\underset{\sim}{Y}$ are expressed as sums of terms of the nth order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$,

$$\underset{\sim}{\beta} = \sum_{n=0}^{\infty}\underset{\sim}{\beta}^{(n)} \quad , \quad \underset{\sim}{Y} = \sum_{n=1}^{\infty}\underset{\sim}{y}^{(n)} \quad ,$$

the terms of the zeroth, first and second orders in the first member of (44.3) yield the following three equations, valid for all $\underset{\sim}{S}$ and $\underset{\sim}{T}$:

$$\Delta N(\beta_{,y}^{(1)}y_{,v}^{(1)}-\beta_{,v}^{(1)}y_{,y}^{(1)}) = 0 \quad ,$$

$$\Delta N(\beta_{,y}^{(2)}y_{,v}^{(1)}-\beta_{,v}^{(2)}y_{,y}^{(1)}+\beta_{,y}^{(1)}y_{,v}^{(2)}-\beta_{,v}^{(1)}y_{,y}^{(2)}) = 0 \quad ,$$

$$\Delta N(\beta_{,y}^{(3)}y_{,v}^{(1)}-\beta_{,v}^{(3)}y_{,y}^{(1)}+\beta_{,y}^{(2)}y_{,v}^{(2)}-\beta_{,v}^{(2)}y_{,y}^{(2)}+\beta_{,y}^{(1)}y_{,v}^{(3)}-$$
$$-\beta_{,v}^{(1)}y_{,y}^{(3)}+\gamma_{,y}^{(2)}z_{,v}^{(2)}-\gamma_{,v}^{(2)}z_{,y}^{(2)}) = 0 \quad , \qquad (46.1)$$

where, throughout the remainder of this Chapter, Δ is to be understood to mean Δ_j.

Now, $\underset{\sim}{\beta}^{(1)} = \underset{\sim}{\beta}_a\underset{\sim}{S} + \underset{\sim}{\beta}_b\underset{\sim}{T}$, $\underset{\sim}{y}^{(1)} = \underset{\sim}{y}_a\underset{\sim}{S} + \underset{\sim}{y}_b\underset{\sim}{T}$ and the first member of (46.1) may be written

$$\Delta N(\beta|y) = 0 \quad .$$

Since $\beta_a = \alpha_B^3 v_a$, $\beta_b = \alpha_B^3 v_b$, this becomes

$$\Delta N_y(y|v) = 0 \quad ,$$

which is identical with the first member of (14.3). The second member of (46.1) is linear in S_y and T_y and hence yields two identities. The coefficient of S_y gives

$$2\Delta N(y_b s_{\beta 1} - \beta_b s_{y1}) + \Delta N(\beta_a s_{y2} - y_a s_{\beta 2}) = 0 \quad . \qquad (46.2)$$

The expression on the left of this equality is typical of the manner in which the nth order G-coefficients enter into the identities between the nth order coefficients. The β-coefficients are given by

$$\beta_{\mu\nu\tau}^{(n)} = \alpha_B{}^3 v_{\mu\nu\tau}^{(n)} + \beta^*{}_{\mu\nu\tau}^{(n)} \quad , \quad \gamma_{\mu\nu\tau}^{(n)} = \alpha_B w_{\mu\nu\tau}^{(n)} + \gamma^*{}_{\mu\nu\tau}^{(n)} \quad .$$

Thus, (46.2) becomes

$$2N_y(y_b s_{v1} - v_b s_{y1}) + N_y(v_a s_{y2} - y_a s_{v2}) = -\Delta N(2y_b s^*{}_{\beta 1} - y_a s^*{}_{\beta 2}) \quad ,$$

since at the first surface all y- and v-coefficients of orders greater than one vanish. Moreover, in view of (40.5), the left hand member of this may be expressed simply in terms of the G-coefficients and, using (14.3) the above equation becomes

$$S_{gy2a} - 2S_{gy1b} = \Delta N(2y_b s^*{}_{\beta 1} - y_a s^*{}_{\beta 2}) = 3(N_{y1}/g)\Delta\alpha_B \beta_B v_a \quad .$$

Identities between the third order G-coefficients are obtained in a similar manner from the last member of (46.1). This is of the second order in S and T and transforms like Y under reflection in Ⓜ. Hence six third order identities are obtained. From the coefficient of S_y^2

$$3\Delta N(y_b t_{\beta 1} - \beta_B t_{y1}) + \Delta N(\beta_a t_{y2} - y_a t_{\beta 2}) = 2\Delta N(s_{y1} s_{\beta 2} - s_{y2} s_{\beta 1}) \quad ,$$

which, after rearranging the left hand terms as above, yields

$$3T_{gy1b} - T_{gy2a} = \Delta N(y_a t^*{}_{\beta 2} - 3y_b t^*{}_{\beta 1}) + 2N(s_{y1} s_{\beta 2} - s_{y2} s_{\beta 1}) \quad .$$

47. The Full Set of Extremal Identities

The full set of 12 second order and 28 third order extremal identities obtained in this manner from (44.3) are given in Tables 47/1,2.[1] In these equations Δ is to be understood to be Δ_j and all quantities without a surface indicator attached refer to the jth surface. According to (43.2) there should only be ten second order and 21 third order identities. Consequently, two of the second order identities and seven of the third order identities must in effect be combinations of the remaining identities. By inspection it is seen that (i) the sum of the fourth, fifth and eleventh members of Table 47/1 is zero, and (ii) the ninth member of Table 47/1 is the sum of the eighth and twelfth members. No other relations can be found between the second order identities and thus ten of the twelve are in fact mutually independent. Seven similar relations can be found between the third order identities, leaving only 21 mutually independent third order identities. The first order identities are just (14.3) and are not repeated here.

TABLE 47/1 Second Order Extremal Identities

$$S_{gy2a} - 2S_{gy1b} = \Delta N(2y_b s^*_{\beta 1} - y_a s^*_{\beta 2}) \quad , \quad 2S_{gy5a} - S_{gy2b} = \Delta N(y_b s^*_{\beta 2} - 2y_a s^*_{\beta 5}) \quad ,$$

$$S_{gz3a} - 2S_{gy8a} = \Delta N(2y_a s^*_{\beta 8} - z_a s^*_{\gamma 3}) \quad , \quad S_{gz4a} - S_{gy9a} = \Delta N(y_a s^*_{\beta 9} \; z_a s^*_{\gamma 4}) \quad ,$$

$$S_{gy9a} - S_{gz3b} = \Delta N(z_b s^*_{\gamma 3} - y_a s^*_{\beta 9}) \quad , \quad 2S_{gy10a} - S_{gz4b} = \Delta N(z_b s^*_{\gamma 4} - 2y_a s^*_{\beta 10}) ,$$

$$S_{gz6a} - 2S_{gy8b} = \Delta N(2y_b s^*_{\beta 8} - z_a s^*_{\gamma 6}) \quad , \quad S_{gz7a} - S_{gy9b} = \Delta N(y_b s^*_{\beta 9} - z_a s^*_{\gamma 7}) \quad ,$$

$$S_{gz6b} - S_{gy9b} = \Delta N(y_b s^*_{\beta 9} - z_b s^*_{\gamma 6}) \quad , \quad S_{gz7b} - 2S_{gy10b} = \Delta N(2y_b s^*_{\beta 10} - z_b s^*_{\gamma 7}) \quad ,$$

$$S_{gz3b} - S_{gz4a} = \Delta N(z_a s^*_{\gamma 4} - z_b s^*_{\gamma 3}) \quad , \quad S_{gz6b} - S_{gz7a} = \Delta N(z_a s^*_{\gamma 7} - z_b s^*_{\gamma 6}) \quad ,$$

---

1 In Table 47/1, the first line arises from the first integrability condition (44.3), the second from the second, etc. In Table 47/2, the first group arises from the first integrability condition, the second from the second, etc.

TABLE 47/2 Third Order Extremal Identities

$$T_{gy2a} - 3T_{gy1b} = \Delta N(3y_b t^*_{\beta 1} - y_a t^*_{\beta 2}) + 2N(s_{y2}s_{\beta 1} - s_{y1}s_{\beta 2}) \quad ,$$

$$T_{gy5a} - T_{gy2b} = \Delta N(y_b t^*_{\beta 2} - y_a t^*_{\beta 5}) + 2N(s_{y5}s_{\beta 1} - s_{y1}s_{\beta 5}) \quad ,$$

$$3T_{gy11a} - T_{gy5b} = \Delta N(y_b t^*_{\beta 5} - 3y_a t^*_{\beta 11}) + 2N(s_{y5}s_{\beta 2} - s_{y2}s_{\beta 5}) \quad ,$$

$$T_{gy14a} - T_{gy8b} = \Delta N(y_b t^*_{\beta 8} - y_a t^*_{\beta 14}) + N(s_{\gamma 3}s_{z6} - s_{\gamma 6}s_{z3}) \quad ,$$

$$T_{gy15a} - T_{gy9b} = \Delta N(y_b t^*_{\beta 9} - y_a t^*_{\beta 15}) + N(s_{\gamma 3}s_{z7} - s_{\gamma 7}s_{z3} + s_{\gamma 4}s_{z6} - s_{\gamma 6}s_{z4}) \quad ,$$

$$T_{gy16a} - T_{gy10b} = \Delta N(y_b t^*_{\beta 10} - y_a t^*_{\beta 16}) + N(s_{\gamma 4}s_{z7} - s_{\gamma 7}s_{z4}) \quad ,$$

$$T_{gz3a} - T_{gy8a} = \Delta N(y_a t^*_{\beta 8} - z_a t^*_{\gamma 3}) + 2N(s_{\beta 8}s_{y1} - s_{\beta 1}s_{y8}) \quad ,$$

$$2T_{gz4a} - T_{gy9a} = \Delta N(y_a t^*_{\beta 9} - 2z_a t^*_{\gamma 4}) + N(s_{\gamma 3}s_{z4} - s_{\gamma 4}s_{z3} + 2s_{\beta 9}s_{y1} - 2s_{\beta 1}s_{y9}) \quad ,$$

$$T_{gz6a} - 2T_{gy14a} = \Delta N(2y_a t^*_{\beta 14} - z_a t^*_{\gamma 6}) + N(s_{\gamma 6}s_{z3} - s_{\gamma 3}s_{z6} + 2s_{\beta 8}s_{y2} - s_{\beta 2}s_{y8}) \quad ,$$

$$T_{gy7a} - T_{gy15a} = \Delta N(y_a t^*_{\beta 15} - z_a t^*_{\gamma 7}) + N(s_{\gamma 6}s_{z4} - s_{\gamma 4}s_{z6} + s_{\beta 9}s_{y2} - s_{\beta 2}s_{y9}) \quad ,$$

$$T_{gy9a} - 2T_{gz3b} = \Delta N(2z_b t^*_{\gamma 3} - y_a t^*_{\beta 9}) + N(s_{\gamma 3}s_{z4} - s_{\gamma 4}s_{z3} + 2s_{\beta 1}s_{y9} - 2s_{\beta 9}s_{y1}) \quad ,$$

$$T_{gy10a} - T_{gz4b} = \Delta N(z_b t^*_{\gamma 4} - y_a t^*_{\beta 10}) + 2N(s_{\beta 1}s_{y10} - s_{\beta 10}s_{y1}) \quad ,$$

$$T_{gy15a} - T_{gz6b} = \Delta N(z_b t^*_{\gamma 6} - y_a t^*_{\beta 15}) + N(s_{\gamma 3}s_{z7} - s_{\gamma 7}s_{z3} + s_{\beta 2}s_{y9} - s_{\beta 9}s_{y2}) \quad ,$$

$$2T_{gy16a} - T_{gz7b} = \Delta N(z_b t^*_{\gamma 7} - 2y_a t^*_{\beta 16}) + N(s_{\gamma 4}s_{z7} - s_{\gamma 7}s_{z4} + 2s_{\beta 2}s_{y10} - 2s_{\beta 10}s_{y2}) \quad ,$$

TABLE 47/2 continued

$$T_{gz6a} - 2T_{gy8b} = \Delta N(2y_b t^*_{\beta 8} - z_a t^*_{\gamma 6}) + N(s_{\gamma 3}s_{z6} - s_{\gamma 6}s_{z3} + 2s_{\beta 8}s_{y2} - 2s_{\beta 2}s_{y8}) ,$$

$$T_{gz7a} - T_{gy9b} = \Delta N(y_b t^*_{\beta 9} - z_a t^*_{\gamma 7}) + N(s_{\gamma 3}s_{z7} - s_{\gamma 7}s_{z3} + s_{\beta 9}s_{y2} - s_{\beta 2}s_{y9}) ,$$

$$T_{gz12a} - T_{gy14b} = \Delta N(y_b t^*_{\beta 14} - z_a t^*_{\gamma 12}) + 2N(s_{\beta 8}s_{y5} - s_{\beta 5}s_{y8}) ,$$

$$2T_{gz13a} - T_{gy15b} = \Delta N(y_b t^*_{\beta 15} - 2z_a t^*_{\gamma 13}) + N(s_{\gamma 6}s_{z7} - s_{\gamma 7}s_{z6} + 2s_{\beta 9}s_{y5} - s_{\beta 5}s_{y9}) ,$$

$$T_{gy9b} - T_{gz6b} = \Delta N(z_b t^*_{\gamma 6} - y_b t^*_{\beta 9}) + N(s_{\gamma 6}s_{z4} - s_{\gamma 4}s_{z6} + s_{\beta 2}s_{y9} - s_{\beta 9}s_{y2}) ,$$

$$2T_{gy10b} - T_{gz7b} = \Delta N(z_b t^*_{\gamma 7} - 2y_b t^*_{\beta 10}) + N(s_{\gamma 7}s_{z4} - s_{\gamma 4}s_{z7} + 2s_{\beta 2}s_{y10} - 2s_{\beta 10}s_{y2}) ,$$

$$T_{gy15b} - 2T_{gz12b} = \Delta N(2z_b t^*_{\gamma 12} - y_b t^*_{\beta 15}) + N(s_{\gamma 6}s_{z7} - s_{\gamma 7}s_{z6} + 2s_{\beta 5}s_{y9} - 2s_{\beta 9}s_{y5}) ,$$

$$T_{gy16b} - T_{gz13b} = \Delta N(z_b t^*_{\gamma 13} - y_b t^*_{\beta 16}) + 2N(s_{\beta 5}s_{y10} - s_{\beta 10}s_{y5}) ,$$

$$T_{gz3b} - T_{gz4a} = \Delta N(z_a t^*_{\gamma 4} - z_b t^*_{\gamma 3}) + N(s_{\gamma 4}s_{z3} - s_{\gamma 3}s_{z4}) ,$$

$$T_{gz6b} - T_{gz7a} = \Delta N(z_a t^*_{\gamma 7} - z_b t^*_{\gamma 6}) + N(s_{\gamma 7}s_{z3} - s_{\gamma 3}s_{z7} + s_{\gamma 4}s_{z6} - s_{\gamma 6}s_{z4}) ,$$

$$T_{gz12b} - T_{gz13a} = \Delta N(z_a t^*_{\gamma 13} - z_b t^*_{\gamma 12}) + N(s_{\gamma 7}s_{z6} - s_{\gamma 6}s_{z7}) ,$$

$$3T_{gz17b} - T_{gz18a} = \Delta N(z_a t^*_{\gamma 18} - 3z_b t^*_{\gamma 17}) + 2N(s_{\beta 9}s_{y8} - s_{\beta 8}s_{y9}) ,$$

$$T_{gz18b} - T_{gz19a} = \Delta N(z_a t^*_{\gamma 19} - z_b t^*_{\gamma 18}) + 2N(s_{\beta 10}s_{y8} - s_{\beta 8}s_{y10}) ,$$

$$T_{gz19b} - 3T_{gz20a} = \Delta N(3z_a t^*_{\gamma 20} - z_b t^*_{\gamma 19}) + 2N(s_{\beta 10}s_{y9} - s_{\beta 9}s_{y10}) .$$

VIII. CLASSIFICATION AND INTERPRETATION OF COEFFICIENTS AND ABERRATIONS

48. The Classification of the Aberrations in the Axial Theory

(a) The aberration of a plane symmetric system has the formal paracanonical expansion

$$\underset{\sim}{\epsilon} = \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{\sim}{E}_{\mu\nu\tau}^{(n)} S_y^{\,n-\mu} T_y^{\,\mu-\nu} S_z^{\,\nu-\tau} T_z^{\,\tau} \quad . \tag{48.1}$$

Due to symmetry requirements ν is even for $E_{y\mu\nu\tau}^{(n)}$ and odd for $E_{z\mu\nu\tau}^{(n)}$. Formulae have been given which make it possible to compute the aberration coefficients $E_{\mu\nu\tau}^{(n)}$ when the surfaces are spherical. However nothing has been said of the meaning of the various coefficients. It is possible to classify the coefficients $\underset{\sim}{E}_{\mu\nu\tau}^{(n)}$ into small groups characterizing a particular type of aberration. Although the meaning of the corresponding coefficients for any quantity $\underset{\sim}{Q}$ bears no relation to that of the $\underset{\sim}{E}_{\mu\nu\tau}^{(n)}$; it is convenient to extend this classification formally to the coefficients of $\underset{\sim}{Q}$.

(b) In the context of the axial theory Steward[1] classified the terms of the aberration functions into two distinct classes: the S-group and the C-group. Any particular aberration of the S-group is an astigmatic- or S-type of aberration and is symmetric about the meridional plane through the ideal image and normal to the meridional plane. Any particular aberration of the C-group is a comatic- or C-type of aberration and is symmetric only about the meridional plane. Each term in the aberration functions depends on the object height H and the radius ρ of an annular

---

1 Steward (1958) §16.

zone of the entrance pupil through the factor $\rho^{2n-m+1}H^m$ where m = 0,1,2,... ...,2n+1 and is even for S-types and odd for C-types.
Buchdahl[2] further classified the aberrations in each type according to their dependence on ρ and H. Thus, an aberration depending on $\rho^{2n-m+1}h^m$ is said to be (2n+1)th order, mth degree coma if m is odd , (2n+1)th order, (2n-m+1)th degree astigmatism if m is even. It will be seen that this classification can be extended to asymmetrical systems.

49. S- and C-types of Aberrations

(a) When parabasal imagery was discussed, polar coordinates (ρ,θ) were introduced in the entrance pupil (§21b):

$$S_y = a\rho \cos\theta \quad , \quad S_z = \rho \sin\theta \quad .$$

It is now expedient to introduce polar coordinates in the object plane. With the basal point of O_1 as origin, set

$$H_{y1} = H\cos\phi \quad , \quad H_{z1} = H\sin\phi \quad ,$$

whence, if SPC are used,

$$T_y = \tau H\cos\phi \quad , \quad T_z = \tau H\sin\phi \quad .$$

When these sets of polar coordinates are substituted into (48.1), either the various powers of a and τ may be absorbed into the coefficients, or a and τ formally set equal to unity. Either way, the powers of these constants may be readily recovered if required and the nth order terms of (48.1) become

---

2 Buchdahl (1958) §7.

$$\underset{\sim}{\epsilon}^{(n)} = \sum_{\mu\nu\tau}^{n} \underset{\sim}{E}_{\mu\nu\tau}{}^{(n)} \rho^{n-\mu+\nu-\tau} H^{\mu-\nu+\tau} \cos^{n-\mu}\theta \sin^{\nu-\tau}\theta \cos^{\mu-\nu}\phi \sin^{\tau}\phi \quad . \quad (49.1)$$

Replace ν by $\bar{\nu} = \mu - \nu + \tau$. Since $\mu \geq \nu \geq \tau$, $\bar{\nu}$ satisfies $\mu \geq \bar{\nu} \geq \tau$. If

$$\bar{\underset{\sim}{E}}_{\mu\bar{\nu}\tau}{}^{(n)} = \underset{\sim}{E}_{\mu(\mu-\nu+\tau)}{}^{(n)} \quad ,$$

(49.1) becomes

$$\underset{\sim}{\epsilon}^{(n)} = \sum_{\mu\bar{\nu}\tau}^{n} \bar{\underset{\sim}{E}}_{\mu\bar{\nu}\tau}{}^{(n)} \rho^{n-\bar{\nu}} H^{\bar{\nu}} \cos^{n-\mu}\theta \sin^{\mu-\bar{\nu}}\theta \cos^{\bar{\nu}-\tau}\phi \sin^{\tau}\phi \quad . \quad (49.2)$$

Now

$$\sum_{\mu\bar{\nu}\tau}^{n} = \sum_{\mu=0}^{n} \sum_{\bar{\nu}=0}^{\mu} \sum_{\tau=0}^{\bar{\nu}} = \sum_{\bar{\nu}=0}^{n} \sum_{\mu=\bar{\nu}}^{n} \sum_{\tau=0}^{\bar{\nu}} \quad .$$

Then

$$\underset{\sim}{\epsilon}^{(n)} = \sum_{\bar{\nu}=0}^{n} \rho^{n-\bar{\nu}} H^{\bar{\nu}} \underset{\sim}{E}_{\bar{\nu}}{}^{(n)}(\theta,\phi) \quad , \quad (49.3)$$

where

$$\underset{\sim}{E}_{\bar{\nu}}{}^{(n)}(\theta,\phi) = \sum_{\mu=\bar{\nu}}^{n} \left[\sum_{\tau=0}^{\bar{\nu}} \bar{\underset{\sim}{E}}_{\mu\bar{\nu}\tau}{}^{(n)} \cos^{\bar{\nu}-\tau}\phi \sin^{\tau}\phi \right] \cos^{n-\mu}\theta \sin^{\mu-\bar{\nu}}\theta$$

$$= \sum_{\mu=\bar{\nu}}^{n} \underset{\sim}{E}_{\mu\bar{\nu}}{}^{(n)} \cos^{n-\mu}\theta \sin^{\mu-\bar{\nu}}\theta \quad . \quad (49.4)$$

The last equation defines $\underset{\sim}{E}_{\mu\bar{\nu}}{}^{(n)}$. The summation in (49.3) consists of two parts: first, when $n-\bar{\nu}$ is even and second, when $n-\bar{\nu}$ is odd. Consider (49.4) when $n-\bar{\nu}$ is even. In that case all terms in (49.4) are alternately of the form $\cos^{2s}\theta \sin^{2r}\theta$ or $\sin 2\theta \cos^{2s}\theta \sin^{2r}\theta$. Since $\cos 2\theta =$ $= 2\cos^2\theta - 1 = 1 - 2\sin^2\theta$, this gives $\underset{\sim}{E}_{\bar{\nu}}{}^{(n)}(\theta,\phi)$ as a function of 2θ of the form

$$\underset{\sim}{E}_{\bar{\nu}}{}^{(n)}(\theta,\phi) = \underset{\sim}{E}_{C\bar{\nu}}{}^{(n)}(2\theta,\phi) = \underset{\sim}{f}_{\bar{\nu}}{}^{(n)}(\cos 2\theta,\phi) + \sin 2\theta\, \bar{\underset{\sim}{f}}_{\bar{\nu}}{}^{(n)}(\cos 2\theta,\phi) \quad . \quad (49.5)$$

When $n-\bar{\nu}$ is odd the sum in (49.4) consists of terms $\cos\theta \cos^{2s}\theta \sin^{2r}\theta$ and $\sin\theta \cos^{2s}\theta \sin^{2r}\theta$ and $\underset{\sim}{E}_{\bar{\nu}}{}^{(n)}(\theta,\phi)$ is a function of θ of the form

$$\underset{\sim}{E}_{\bar{\nu}}{}^{(n)}(\theta,\phi) = \underset{\sim}{E}_{S\bar{\nu}}{}^{(n)}(\theta,\phi) = \cos\theta\, \underset{\sim}{g}_{\bar{\nu}}{}^{(n)}(\cos 2\theta,\phi) + \sin\theta\, \bar{\underset{\sim}{g}}_{\bar{\nu}}{}^{(n)}(\cos 2\theta,\phi) \quad . \quad (49.6)$$

Define

$$\underset{\sim}{\epsilon}_C{}^{(n)} = \sum_{\bar{\nu}=0}^{n} \underset{\sim}{E}_{C\bar{\nu}}{}^{(n)}(2\theta,\phi)\rho^{n-\bar{\nu}} H^{\bar{\nu}} \qquad (n-\nu \text{ even}) \quad ,$$

$$\underset{\sim}{\epsilon}_S{}^{(n)} = \sum_{\bar{\nu}=0}^{n} \underset{\sim}{E}_{S\bar{\nu}}{}^{(n)}(\theta,\phi)\rho^{n-\bar{\nu}} H^{\bar{\nu}} \qquad (n-\nu \text{ odd}) \quad .$$

and hence

$$\epsilon^{(n)} = \epsilon_C^{(n)} + \epsilon_S^{(n)} \quad , \tag{49.7}$$

where $\epsilon_C^{(n)}$ is a function 2θ containing only even powers of ρ and $\epsilon_S^{(n)}$ is a function of θ and contains only odd powers of ρ.

(b) By summing (49.7) from $n = 1$ to ∞, the total aberration becomes

$$\epsilon = \epsilon_C + \epsilon_S \quad ,$$

where ϵ_C and ϵ_S are sums of terms proportional to $\rho^{n-\bar{\nu}}H^{\bar{\nu}}$ where $n-\bar{\nu}$ is even and odd respectively; ϵ_C is the sum of all aberrations belonging to the C-group and ϵ_S is the sum of all aberrations belonging to the S-group. Each individual term $F_{C\bar{\nu}}^{(n)}(2\theta,\phi)\rho^{n-\bar{\nu}}H^{\bar{\nu}}$ in ϵ_C or $F_{S\bar{\nu}}^{(n)}(\theta,\phi)\rho^{n-\bar{\nu}}H^{\bar{\nu}}$ in ϵ_S is said to represent a distinct aberration-type. In particular, the aberration types in ϵ_C are comatic-types; those in ϵ_S are astigmatic-types. If two or more aberrations of the same type are combined, the resultant is an aberration of the same type, otherwise the aberration is said to be mixed. In particular, the term in (49.7) depending on ρ and H through the factor $\rho^{n-\bar{\nu}}H^{\bar{\nu}}$ will be called either

(i) nth order, $(n-\bar{\nu})$th degree astigmatism if $n-\bar{\nu}$ is odd

or (ii) nth order, $\bar{\nu}$th degree coma if $n-\bar{\nu}$ is even ,

and constitutes a fundamental or basic aberration-type. Evidently an aberration type is uniquely determined by the values of n and $\bar{\nu} = \mu - \nu + \tau$. The aberration coefficients of order n are thus divided into $n+1$ groups by this classification into aberration types and knowledge of all coefficients with a given value of n and $\bar{\nu}$ completely determines the contribution to ϵ by the corresponding aberration-type.

(c) This classification can be extended to the coefficients of quantities other than $\underset{\sim}{\epsilon}$ by simply examining the values of $\underline{n}$ and $\bar{\nu}$ irrespective of the fact that the coefficient is quite unrelated to coma or astigmatism. In particular, the aberration type with $\bar{\nu} = \underline{n}$ is frequently called <u>distortion</u> (see §7b) and for any quantity $\underset{\sim}{Q}$ the term $q_{n\tau\tau}^{(n)} T_y^{n-\tau} T_z^{\tau}$ is said to be a <u>distortion term</u>. Moreover, the surface coefficients $\hat{q}_{\mu\nu\tau}^{(n)}$ are also classified according to (i) and (ii) of §49b. Thus the coefficients $\hat{q}_{n\tau\tau}^{(n)}$ are (surface) coefficients of distortion terms.

The formal expansion (48.1) is valid even in the absence of plane symmetry provided the restriction on the values of ν is relaxed. Formally, (48.1) is also independent of the nature of the surfaces. It is evident that the classification of aberrations into S- and C-types is valid for an arbitrary optical system.

50. <u>Characteristic Features of Astigmatic and Comatic Aberrations</u>

(a) Some general conclusions regarding the shape of the image patch in the presence of a <u>single</u> aberration-type may be drawn from the form of (49.5,6). It will be assumed that vignetting is adequately represented by $\rho \leq \rho_{max}$. Consider first a particular astigmatic-type

$$\underset{\sim}{\epsilon} = \rho^{n-\nu} H^{\nu} \underset{\sim}{E}_{S\nu}^{(n)}(\theta,\phi) \quad .$$

It is evident from (49.6) that if ρ is kept fixed and θ allowed to vary ($0 \leq \theta \leq 2\pi$), a closed curve in the image plane is obtained from the above equation for $\underset{\sim}{\epsilon}$. Such a curve is called an aberration curve.[1] There are no terms independent of θ in $\underset{\sim}{E}_{SV}^{(n)}(\theta,\phi)$ and consequently, as ρ takes values $0 \leq \rho \leq \rho_{max}$, the origin $\underset{\sim}{\epsilon} = 0$ remains within the curve $0 \leq \theta \leq 2\pi$. Moreover, only a single power of ρ is involved in $\underset{\sim}{\epsilon}$ and all of the aberration curves $0 \leq \rho \leq \rho_{max}$ are geometrically similar. Thus the image is a patch of light bounded by the curve $\rho = \rho_{max}$.

Now consider a single comatic-type:

$$\underset{\sim}{\epsilon} = \rho^{n-\nu} H^{\nu} \underset{\sim}{E}_{C\nu}^{(n)}(2\theta,\phi) \quad .$$

As before, for fixed ρ, the curve $0 \leq \theta \leq 2\pi$ is closed but since $\underset{\sim}{E}_{C\nu}^{(n)}$ is of period π in θ, it is traversed twice while θ varies from 0 to 2π. By varying ρ, geometrically similar curves are again obtained. However, $\underset{\sim}{E}_{C\nu}^{(n)}$ contains a term independent of θ and the aberration curves are translated with respect to each other as ρ is varied. Two cases arise. First, the origin $\underset{\sim}{\epsilon} = 0$ remains either within or on the aberration curves and, second, the origin lies outside the curves. In the latter case the same power of ρ governs both the magnification and translation of the curves and each curve will touch two fixed straight lines through the origin. Thus the image patch is bounded by two straight lines and an arc of the curve $\rho = \rho_{max}$, clearly bearing a great resemblance to the typical case of linear coma for a rotationally symmetric system. The case where the origin remains

---

[1] For pencils of sufficiently small aperture, those rays for which ρ = const. will pass through an annular aperture in the diaphragm and the aberration curves are identical with the aperture curves (§78) obtained by vignetting all rays apart from those which pass through a specified annular aperture in the plane of the diaphragm.

within the aberration curves is a degenerate case of typical coma and arises when the slope of the tangent lines is imaginary. Consequently, no distinction will in general be drawn between these two cases of comatic aberration-types.

It is evident that both the comatic and astigmatic aberration-types are in general asymmetric. That is, they are neither symmetric about a line normal to Ⓜ, or about a line parallel to Ⓜ. In the presence of two or more astigmatic aberration-types (but no coma), the aberration curves are no longer geometrically similar since two or more distinct powers of ρ are involved. In the presence of two or more comatic-types (but no astigmatism) the tangent lines will become curves and the image patch may combine features of both cases of coma discussed above.

(b) In the above the object was in general not in the meridional plane. Thus, for a fixed object, Ⓚ would behave essentially as an asymmetric system. In fact, the general conclusions obtained above hold for asymmetric systems. Now suppose that the object lies in Ⓜ. Then $\phi = 0$ and only those coefficients for which $\tau = 0$ will contribute to the aberration. Consequently, since the system is plane symmetric, $\mu - \bar{\nu} = \nu$ is even in ϵ_y and odd in ϵ_z. In the light of this, the formation of the functions $\underset{\sim}{F}_{S\bar{\nu}}{}^{(n)}(\theta,\phi)$ and $\underset{\sim}{F}_{C\bar{\nu}}{}^{(n)}(2\theta,\phi)$ should be reconsidered. The individual terms in $F_{yC\bar{\nu}}{}^{(n)}$ are now of the form $\cos^{2r}\theta \sin^{2s}\theta$, and in $F_{zC\bar{\nu}}{}^{(n)}$ of the form $\sin 2\theta \cos^{2r}\theta \sin^{2s}\theta$. Thus[1]

$$F_{yC\bar{\nu}}{}^{(n)}(2\theta) = f_{y\bar{\nu}}{}^{(n)}(\cos 2\theta) \quad , \quad F_{zC\bar{\nu}}{}^{(n)} = \sin 2\theta\, \bar{f}_{y\bar{\nu}}{}^{(n)}(\cos 2\theta). \tag{50.1}$$

---

[1] Since $\phi = 0$, the formal dependence of $\underset{\sim}{F}_{C\bar{\nu}}{}^{(n)}$ etc., on ϕ will be dropped.

Similarly,

$$F_{yS\bar{\nu}}^{(n)}(\theta) = \cos\theta\, g_{y\bar{\nu}}^{(n)}(\cos 2\theta) \quad , \quad F_{zS\bar{\nu}}^{(n)}(2\theta) = \sin\theta\, \bar{g}_{z\bar{\nu}}^{(n)}(\cos 2\theta). \tag{50.2}$$

The general form of the aberration curves is unchanged. However, both the comatic and astigmatic aberration curves are symmetric about Ⓜ and the astigmatic aberration curve has the additional property of being symmetric about the line $\epsilon_y = 0$. A combination of any number of aberrations of the <u>same</u> type will preserve the symmetries associated with the appropriate aberration-type. However, a mixed aberration is only symmetric about Ⓜ. As the strength of the comatic aberration in the mixture is increased, so is the asymmetry about $\epsilon_y = 0$. (cf. figs. 12,13 and figs. 14,16.)

(c) A further specialisation is the case of a basal object. In this case H = 0 and only those terms for which $\tau = 0$, $\mu = \nu$ (i.e., $\bar{\nu} = 0$) will contribute to the aberration. $\underset{\sim}{F}_{c\bar{\nu}}^{(n)}$ and $\underset{\sim}{F}_{S\bar{\nu}}^{(n)}$ are again given by equations of the form (50.1,2) (with $\bar{\nu} = 0$) and the image patch has the same general appearance as when the object is meridional but <u>not</u> basal. This is to be expected since no additional symmetry property is invoked by constraining the object to lie on $Ⓡ_B$. When <u>n</u> is even, the aberration is comatic; when <u>n</u> is odd, the aberration is astigmatic. Thus the aberration of a basal object is <u>not</u> a generalisation of spherical aberration in the axial theory, since spherical aberration is exclusively astigmatic. (Compare these results with the analogous ones for a symmetric system in the axial theory. When the object is not axial the plane through the axis and the object is a plane of symmetry of the system. The form of the image is well known and in full accordance with §50b above. If the object is now chosen to lie on the axis, rotational symmetry is invoked and the image is rotationally symmetric. In fact is is exclusively astigmatic.)

51. Displacements and Rotations of the Image Plane

(a) Suppose the image plane is displaced along the x-axis by an amount x, measured positive if the displacement is in the positive x-direction. This effectively decreases both x_M and x_S by the amount x and the change in the image height is $x\underset{\sim}{V}$. In the new image plane the image height is given by

$$\bar{\underset{\sim}{H}} = \underset{\sim}{\epsilon} + \underset{\sim}{m}H_1 + x\underset{\sim}{V} = \underset{\sim}{\epsilon} + x\underset{\sim}{v}_a\underset{\sim}{S} + x\underset{\sim}{V}^\dagger + (\underset{\sim}{m}+x\underset{\sim}{v}_b\tau)\underset{\sim}{H}_1 \quad .$$[1]

If the magnification in the displaced image plane is defined to be

$$\bar{\underset{\sim}{m}} = \underset{\sim}{m} + x\underset{\sim}{v}_b\tau \quad ,$$

the new aberration is

$$\bar{\underset{\sim}{\epsilon}} = \underset{\sim}{\epsilon} + x\underset{\sim}{v}_a\underset{\sim}{S} + x\underset{\sim}{V}^\dagger \quad , \tag{51.1}$$

and the corresponding coefficients are

$$\bar{\underset{\sim}{E}}_a = \underset{\sim}{E}_a + x\underset{\sim}{v}_a \quad , \quad \bar{\underset{\sim}{E}}_{\mu\nu\tau}{}^{(n)} = \underset{\sim}{E}_{\mu\nu\tau}{}^{(n)} + x\underset{\sim}{v}_{\mu\nu\tau}{}^{(n)} \quad . \tag{51.2}$$

By varying x the values of these coefficients are changed. Thus the plane of best focus may be determined from a knowledge of the coefficients of the system. Suppose a certain $\underset{\sim}{v}$-coefficient is large compared to the others or in comparison with the corresponding aberration coefficient. Quite small displacements of the image plane then suffice to change the corresponding aberration coefficient by a significant amount whilst the others are virtually unaffected (see §60a). Since a single parameter is involved in a displacement of the image plane, only a single coefficient can be adjusted independently of the others by this means. The coefficient adjusted in this manner is usually either of the two coefficients of first

---

[1] For the remainder of this chapter, the magnification is assumed to be independent of $\underset{\sim}{H}_1$.

order first degree astigmatism.

Now suppose that the displacement x is so small that $x\underset{\sim}{V}^{\dagger}$ is negligible as compared to $x\underset{\sim}{v}_a\underset{\sim}{S}$. Then

$$\bar{\epsilon}_y = \epsilon_y + xv_a\rho\cos\theta \quad , \quad \bar{\epsilon}_z = \epsilon_z + xw_a\rho\sin\theta \quad . \tag{51.3}$$

Suppose the aberration $\underset{\sim}{\epsilon}$ is pure coma. Then ϵ_y and ϵ_z are functions of 2θ and if, in (51.3), θ is replaced by $\theta+\pi$ and x by -x, $\bar{\underset{\sim}{\epsilon}}$ is unchanged. It follows that the manifold of rays forming the image is symmetric about the undisplaced image plane. Consequently the area of the image in the undisplaced plane is less than that in any displaced plane. Thus a displacement of the image plane will not have any advantageous effect on a comatic aberration.[2] Moreover, the most objectionable effect of coma is its asymmetry. This arises because of the existence in the aberration of terms independent of θ. It is evident from (51.3) that small displacements of the image plane can have no effect on the asymmetry. (Usually asymmetry can be altered only by making changes to the system so as to introduce an asymmetry of the opposite sign to that already present.) If the aberration is not pure coma then the manifold of rays is asymmetric with respect to the undisplaced image plane and a displacement can be used to reduce the area of the image.[2] A classic example of this is spherical aberration.

It is important to note the distinction between transfer (§10) and displacement of the image plane. In the former no physical changes are involved, rather the changes are in the coordinates with which the image

---

[2] cf. Steward (1958) p.49 .

heights, etc., are specified. On the other hand, when the image plane is displaced, physical changes are being made and naturally the aberrations change. Note that Weinstein's[3] "transfer" does involve physical changes - his wave front after refraction at Ⓕ is physically distinct from that before refraction at $Ⓕ_+$. Thus it is not surprising that the transfer increments to his coefficients are non-zero.[3]

(b) Suppose that the aberration coefficients have been determined for the image plane normal to the x-axis and through the point O_B whose coordinates are $(\ell, \ell V_B, 0)$. Let the aberrations and magnifications associated with this plane be $\underset{\sim}{\epsilon}$ and $\underset{\sim}{m} = \tau \underset{\sim}{h}_b$ respectively. The image height is then

$$\underset{\sim}{H} = \underset{\sim}{Y} + \ell \underset{\sim}{V} = \underset{\sim}{\epsilon} + h_b \underset{\sim}{T} \quad .$$

Now rotate the plane about the z-axis so that its normal is $(\bar{\alpha},\bar{\beta},0)$. By (7.4) the image height associated with the rotated plane is

$$\bar{H}_y = \omega(\epsilon_y + h_{yb}T_y)/(1+\vartheta V) \quad , \quad \bar{H}_z = \epsilon_z + h_{zb}T_z - \bar{\beta}\bar{H}_y W \quad , \qquad (51.4)$$

where

$$\omega = (\bar{\alpha}+\bar{\beta}V_B)^{-1} \quad , \quad \vartheta = \bar{\beta}\omega \quad .$$

The first, second and third order terms of $\bar{\underset{\sim}{H}}$ are

$$\bar{H}_y^{(1)} = \omega[\epsilon_y^{(1)} + h_{yb}T_y] \quad ,$$

$$\bar{H}_y^{(2)} = \omega[\epsilon_y^{(2)} - \vartheta V^{(1)}(\epsilon_y^{(1)} + h_{yb}T_y)] \quad ,$$

$$\bar{H}_y^{(3)} = \omega[\epsilon_y^{(3)} - \vartheta\epsilon_y^{(2)}V^{(1)} + \vartheta(\epsilon_y^{(1)} + h_{yb}T_y)(\vartheta V^{(1)2} - V^{(2)})] \quad , \qquad (51.5)$$

---

3 Weinstein (1949) §7,8 .

and

$$\bar{H}_z^{(1)} = \epsilon_z^{(1)} + h_{zb}T_z \quad ,$$

$$\bar{H}_z^{(2)} = \epsilon_z^{(2)} - \vartheta W^{(1)}(\epsilon_y^{(1)}+h_{yb}T_y) \quad ,$$

$$\bar{H}_z^{(3)} = \epsilon_z^{(3)} - \vartheta W^{(1)}[\epsilon_y^{(2)}-\vartheta V^{(1)}(\epsilon_y^{(1)}+h_{yb}T_y)] - \vartheta W^{(2)}(\epsilon_y^{(1)}h_{yb}T_y) \quad . \tag{51.6}$$

Define the magnification associated with the rotated image plane to be $\bar{\underline{m}}$:

$$\bar{\underline{m}} = (\omega m_y, m_z) = (\omega\tau h_{yb}, \tau h_{zb}) \quad . \tag{51.7}$$

The aberration in the new plane is then

$$\bar{\underline{\epsilon}} = \bar{\underline{H}} - \bar{\underline{m}}H_1 \quad ,$$

and the corresponding aberration coefficients may be read off (51.5,6). For example, the first and second order coefficients are

$$\bar{E}_{ya} = \omega E_{ya} \qquad \bar{E}_{za} = E_{za} \quad ,$$

$$\bar{S}_{ey1} = \omega[S_{ey1}-\vartheta v_a E_{ya}] \qquad \bar{S}_{ez3} = S_{ez3} - \vartheta w_a E_{ya}$$

$$\bar{S}_{ey2} = \omega[S_{ey2}-\vartheta(v_a h_{yb}+v_b E_{ya})] \qquad \bar{S}_{ez4} = S_{ez4} - \vartheta w_b E_{ya}$$

$$\bar{S}_{ey5} = \omega[S_{ey5}-\vartheta v_b h_{yb}] \qquad \bar{S}_{ez6} = S_{ez6} - \vartheta h_{yb} w_a$$

$$\bar{S}_{ey\alpha} = \omega S_{ey\alpha} \quad , \quad \alpha = 8,9,10 \qquad \bar{S}_{ez7} = S_{ez7} - \vartheta h_{yb} w_b \quad . \tag{51.8}$$

Apart from the factor ω, the first order coefficients are unaffected. However, the second order coefficients S_{ey1}, S_{ey2}, S_{ey5} and $S_{ez\alpha}$ are affected and can be adjusted by a rotation of the image plane. Since only a single parameter is involved in such a rotation,[4] only one coefficient can be adjusted independently of the others by this means.

It is customary to discuss the significance of a particular aberration-type under the assumption that other types are zero. Consequently assume that first order astigmatism is zero, whence (51.8) reduce to

$$\bar{S}_{ey2} = \omega(S_{ey2} - \vartheta v_a h_{yb}) \quad , \quad \bar{S}_{ey5} = \omega(S_{ey5} - \vartheta v_b h_{yb}) \quad ,$$

$$\bar{S}_{ez6} = S_{ez6} - \vartheta h_{yb} w_a \quad , \quad \bar{S}_{ez7} = S_{ez7} - \vartheta h_{yb} w_b \quad , \qquad (51.9)$$

with other coefficients unaltered apart from the factor ω for the E_y-coefficients. The coefficients S_{ey2}, S_{ez6} are coefficients of <u>second order, first degree astigmatism</u> (see §52b) and it is seen that they are associated with a curvature of field of the form of an inclination of the image plane.[5] The rotation of the image plane has also altered two of the coefficients of distortion (see §52d). The third order coefficients are of little interest. All coefficients except those of the third order third degree astigmatism are affected, even if first and second order aberrations were absent in the original image plane.

From (51.7) it is evident that a rotation of the image plane can be used to produce equal magnifications in the y- and z-directions.[5] This would probably be their principle use. Otherwise, rotations are of no use in controlling the first order aberrations but could be used to control one second order coefficient. Combined with displacements of the image plane, two coefficients are under control. Any further improvement necessitates changes to Ⓚ.

---

4 The image plane always conforms with the plane symmetry of the system.
5 See Montel (1953) §6.2 .

52. Second Order Aberrations

(a) The detailed nature of any particular aberration-type can, in principle, be obtained from (49.4,5,6). In the context of the axial theory accounts of these appear in most texts.[1] For general systems three works will be mentioned. First, Nijboer,[2] and later Montel,[2] classify the aberrations according to the manner in which the associated coefficients arise in the expansion of the wave-front deformation. Each aberration-type is associated with a single coefficient and the form of the aberration is particularly simple. Finally, Barakat and Houston[3] consider the aberrations of a general system, a plane symmetric system and a doubly symmetric system.

All three of these papers discuss the nature of the aberrations of any order by considering the aberration associated with a single coefficient. In this and the next section the form of the aberration corresponding to each aberration-type of the second and third order will be discussed and, whenever possible, all the coefficients associated with the type under discussion will be simultaneously taken into account. The

---

1 For example, Steward (1958), chapter III.

2 Nijboer (1943); Montel (1953).

3 Barakat and Houston (1966). This paper will be referred to in the next few sections as BH and it should be noted that the captions of some of their diagrams are incorrect. The correct captions to figures 5,7 and 12 are those attached, in BH, to figures 7,12 and 5 respectively. In making references to the diagrams of BH, the captions will be ignored, thus a reference to BH figure 5 refers to the diagram labelled by fig.5 in BH.

total aberration is the sum of the individual types. However, combinations of two or more distinct types will not be discussed and accordingly interesting features will probably go unobserved. For example, the characteristic focal line of a zonal aperture[4] arises when primary coma and astigmatism are combined.

(b) The second order terms of the aberration function are given by

$$\epsilon_y^{(2)} = \rho^2(S_{ey1}\cos^2\theta + S_{ey8}\sin^2\theta) + \rho(S_{ey2}H_{y1}\cos\theta + S_{ey9}H_{z1}\sin\theta) + S_{ey5}H_{y1}^2 + S_{ey10}H_{z1}^2 \quad ,$$

$$\epsilon_z^{(2)} = \rho^2 S_{ez3}\cos\theta\sin\theta + \rho(S_{ez4}H_{z1}\cos\theta + S_{ez6}H_{y1}\sin\theta) + S_{ez7}H_{y1}H_{z1} \quad , \tag{52.1}$$

where H_1 is the object height and powers of a and τ have been absorbed into the coefficients. In discussing the aberrations it is convenient to introduce local coordinates in the image plane such that the origin is the point $\underset{\sim}{\epsilon} = 0$ and the y-axis lies in Ⓜ. The terms in (52.1) depending on ρ^2 are <u>second order, zeroth degree coma</u> and may be written

$$\epsilon_y = (\sigma_1+\sigma_2\cos 2\theta)\rho^2 \quad , \quad \epsilon_z = \sigma_3\rho^2 \sin 2\theta \quad ,$$

where

$$\sigma_1 = \tfrac{1}{2}(S_{ey1}+S_{ey8}) \quad , \quad \sigma_2 = \tfrac{1}{2}(S_{ey1}-S_{ey8}) \quad , \quad \sigma_3 = \tfrac{1}{2}S_{ez3} \quad .$$

Thus

$$(\epsilon_y-\sigma_1\rho^2)^2/\sigma_2^2 + \epsilon_z^2/\sigma_3^2 = \rho^4 \quad , \tag{52.2}$$

which represents an ellipse with semi-axes $\sigma_2\rho^2$, $\sigma_3\rho^2$ centred at $(\sigma_1\rho^2,0)$. Two cases arise. First, if S_{ey1} and S_{ey8} are of the same sign then $|\sigma_1| > |\sigma_2|$ and the image patch is bounded by an arc of the ellipse (52.2)

---

[4] Conrady (1957) §55c.

for which $\rho = \rho_{max}$ and by the two straight lines through the origin and tangent to this ellipse. The angle between the straight bounds is 2θ when

$$\theta = \arctan[\sigma_3(\sigma_1^2-\sigma_2^2)^{-\frac{1}{2}}] \quad .$$

The image patch is symmetric about the line $\epsilon_z = 0$ and has the familiar comatic appearance. Second, if S_{ey1} and S_{ey8} are not of the same sign the image is wholly bounded by the ellipse $\rho = \rho_{max}$.

(c) Second order linear astigmatism is given by

$$\epsilon_y = \rho[(\sigma_4+x)\cos\theta+\sigma_5\sin\theta] \quad , \quad \epsilon_z = \rho[\sigma_6\cos\theta+(\sigma_7+x)\sin\theta] \ , \quad (52.3)$$

where x represents a displacement of the image plane as in §51b. Here

$$\sigma_4 = S_{ey2}H_{y1} \quad , \quad \sigma_5 = S_{ey9}H_{z1} \quad , \quad \sigma_6 = S_{ez4}H_{z1} \quad , \quad \sigma_7 = S_{ez6}H_{y1} \quad .$$

In general (52.3) represents an ellipse whose orientation is determined by the object and the coefficients. Assume that $H_{z1} \neq 0$. Then (52.3) will become

$$\epsilon_y = \alpha\epsilon_z \quad , \quad \alpha = (\sigma_4+x)/\sigma_6 \quad ,$$

in planes such that

$$(\sigma_4+x)(\sigma_7+x) - \sigma_5\sigma_6 = 0 \quad . \quad (52.4)$$

In general therefore there are two distinct planes in which the image is a straight line and these <u>focal lines</u> are not necessarily normal to each other. Let the positions of these image planes be x_1 and x_2 and the corresponding values of α be α_1 and α_2. The angle between the focal lines will be 90° if and only if $\alpha_1\alpha_2 = -1$. Now

$$x_1 + x_2 = -(\sigma_4+\sigma_7) \quad , \quad x_1x_2 = \sigma_4\sigma_7 - \sigma_5\sigma_6 \quad ,$$

thus

$$-\alpha_1\alpha_2 = -[\sigma_4^2+\sigma_4(x_1+x_2)+x_1x_2]/\sigma_6^2 = \sigma_5/\sigma_6 \quad .$$

This is unity if and only if $S_{ey9} = S_{ez4}$, independently of the location of the object. Now suppose $H_{z1} = 0$, that is, the object lies in Ⓜ, in which case $\sigma_5 = \sigma_6 = 0$. It is evident from (52.3) that the two focal lines are normal to each other and pass through the points

$$x_1 = -S_{ey2}H_{y1} \quad , \quad x_2 = -S_{ez6}H_{y1} \quad .$$

When the object is not in Ⓜ, plane symmetry is destroyed. Consequently it is expected that imagery of this object would bear some resemblance to that produced by an arbitrary system. This is indeed the case since the astigmatism discussed above closely resembles first order imagery of non-orthogonal systems.[5]

(d) The terms of (52.1) containing only $\underline{H}_1$ are those of <u>second order</u>, <u>second degree coma</u>, or <u>second order distortion</u>. In the presence of this aberration, the image height is

$$H_y = m_yH_{y1} + S_{ey5}H_{y1}^2 + S_{ey10}H_{z1}^2 \quad , \quad H_z = m_zH_{z1} + S_{ez7}H_{y1}H_{z1} \quad . \quad (52.5)$$

The coefficient affecting meridional rays is S_{ey5} which is therefore termed the coefficient of (second order) meridional distortion. S_{ey10} is the only coefficient involved for sagittal rays and is accordingly the coefficient of (second order) sagittal distortion. The most convenient way to depict the effects of distortion is through the image of a square grid in the object plane. If all three of the coefficients in (52.5) are non-zero the square grid is imaged as a grid of parabolic lines, symmetric about Ⓜ. If the coefficient S_{ey5} is the only non zero coefficient, the square grid is imaged as BH fig.12 (<u>Note</u>: ϵ_y-axis is

---

5 Luneburg (1964) §36.

along the y-axis). The images produced when S_{ey10} and S_{ez7} are respectively the only non zero coefficients are BH figs. 6 and 5 (Note: ϵ_y-axis is along the x-axis).

53. Third Order Aberrations. Comatic Asymmetry.

(a) The terms in $\underset{\sim}{\epsilon}^{(3)}$ proportional to ρ^3 constitute third order, third degree astigmatism and are

$$\epsilon_y = T_{ey1}\rho^3\cos^3\theta + T_{ey8}\rho^3\cos\theta\sin^2\theta \ , \ \epsilon_z = T_{ez3}\rho^3\cos^2\theta\sin\theta + T_{ez17}\rho^3\sin^3\theta \ .$$

For a symmetric system this aberration-type is none other than spherical aberration. Accordingly, write the above as

$$\epsilon_y = \tau_1\rho^3\cos\theta + \tau_2\rho^3\cos\theta\sin^2\theta \ , \ \epsilon_z = \tau_3\rho^3\sin\theta + \tau_4\rho^3\cos^2\theta\sin\theta \ , \quad (53.1)$$

where

$$\tau_1 = T_{ey1} \ , \quad \tau_2 = T_{ey8} - T_{ey1} \ , \quad \tau_3 = T_{ez17} \ , \quad \tau_4 = T_{ez3} - T_{ez17} \ .$$

The coefficients τ_1 and τ_3 govern an aberration very similar to primary spherical aberration except that the aberration curves are ellipses rather than circles. Consequently it is reasonable to take T_{ey1} and T_{ez17} as the coefficients analogous to the coefficient A_a of primary spherical aberration in the axial theory. The remaining terms of (53.1) are governed by τ_2 and τ_4 and represent an aberration curve similar to that of BH fig.8 except that the scale may be different along the two axes.

(b) Third order, linear coma is given by

$$\epsilon_y = T_{ey2}\rho^2H_{y1}\cos^2\theta + T_{ey9}\rho^2H_{z1}\cos\theta\sin\theta + T_{ey14}\rho^2H_{y1}\sin^2\theta \quad ,$$

$$\epsilon_z = T_{ez4}\rho^2H_{z1}\cos^2\theta + T_{ez6}\rho^2H_{y1}\cos\theta\sin\theta + T_{ez18}\rho^2H_{z1}\sin^2\theta \quad .$$

Define

$$\tau_5 = \tfrac{1}{2}(T_{ey2}+T_{ey14}) \quad , \quad \tau_6 = \tfrac{1}{2}(T_{ey2}-T_{ey14}) \quad , \quad \tau_7 = \tfrac{1}{2}T_{ey9} \quad ,$$

$$\tau_8 = \tfrac{1}{2}(T_{ez4}+T_{ez18}) \quad , \quad \tau_9 = \tfrac{1}{2}(T_{ez4}-T_{ez18}) \quad , \quad \tau_{10} = \tfrac{1}{2}T_{ez6} \quad .$$

Then

$$\epsilon_y = \rho^2[(\tau_5+\tau_6\cos 2\theta)H_{y1}+\tau_7H_{z1}\sin 2\theta] \ , \ \epsilon_z = \rho^2[(\tau_8+\tau_9\cos 2\theta)H_{z1}+\tau_{10}H_{y1}\sin 2\theta]. \tag{53.2}$$

For a meridional object this reduces to an aberration-type formally similar to second order, zeroth degree coma, except that it is proportional to $\rho^2 H_{y1}$ and thus vanishes for a basal object. Similarly, if the object is sagittal, the resulting aberration is formally similar to the case just considered but with the roles of the ϵ_y- and ϵ_z-axes interchanged. In the general case, the aberration curve ρ = constant is, by (53.2), an ellipse (centred away from the origin) whose orientation in the image plane is determined by the object. Again two cases arise depending on the values of the coefficients: either the image patch is bounded by the ellipse $\rho = \rho_{max}$, or by an arc of this ellipse and the two tangents to it and through the origin.

(c) Third order, linear astigmatism is analogous to second order linear astigmatism but depends on the second power of $\underset{\sim}{H}$. It is given by

$$\epsilon_y = [(T_{ey5}H_{y1}^{\ 2}+T_{ey10}H_{z1}^{\ 2}+x)\cos\theta+T_{ey15}H_{y1}H_{z1}\sin\theta]\rho \quad ,$$

$$\epsilon_z = [T_{ez7}H_{y1}H_{z1}\cos\theta+(T_{ez12}H_{y1}^{\ 2}+T_{ez19}H_{z1}^{\ 2}+x)\sin\theta]\rho \quad , \tag{53.3}$$

where x represents a displacement of the image plane. T_{ey5}, T_{ey10}, T_{ez12} and T_{ez19} are clearly coefficients of curvature of field and the associated image surface is tangent to the image plane x = 0. This aberration needs no further discussion.

(d) The distortion terms in $\underline{\epsilon}^{(3)}$, that is, third order distortion, are governed by T_{ey11}, T_{ey16}, T_{ez13} and T_{ez20}. In the presence of third order distortion

$$H_y = m_y H_{y1} + T_{ey11} H_{y1}^{\;3} + T_{ey16} H_{y1} H_{z1}^{\;2} \;, \; H_z = m_z H_{z1} + T_{ez13} H_{y1}^{\;2} H_{z1} + T_{ez20} H_{z1}^{\;3} . \tag{53.4}$$

The above four coefficients of distortion are respectively the coefficients D_{26}, D_{28}, D_{30} and D_{23} of BH and reference may be made to this paper, especially to BH figs. 7 and 15 for a description of the significance of each coefficient.

(e) The objectionable feature of coma is its asymmetry with respect to the principal ray and it is of importance that this be reduced as much as possible. Before this can be done, some measure of the asymmetry is required. For this purpose the definitions in M§42a of the meridional and sagittal (comatic) asymmetries K_M and K_S will be adopted:

$$K_M = \tfrac{1}{2}[\epsilon_y(\rho,0)+\epsilon_y(\rho,\pi)] - \epsilon_y(0,0) \;, \; K_S = \epsilon_y(\rho,\tfrac{1}{2}\pi) - \epsilon_y(0,0) . \tag{53.5}$$

Two important properties of these are (i) if the image is coma free, $K_M = K_S = 0$, and (ii) if K_M and K_S are of opposite sign, the point of intersection of the principal ray with the image plane is enclosed by the aberration curve. If $K_M = -K_S$ the image will have a symmetric meridional section. If the expression for the aberration is substituted into (53.5), it is found that

$$K_M = (S_{ey1}+T_{ey2}H_{y1})\rho^2 \quad , \quad K_S = (S_{ey8}+T_{ey14}H_{y1})\rho^2 \quad , \tag{53.6}$$

from which the asymmetries are easily calculated (see §74a).

IX. SYMMETRIC SYSTEMS

54. Rotational Symmetry

(a) The optical system Ⓚ of the previous chapters consisted exclusively of spherical surfaces and had a single plane of symmetry. The restriction to spherical surfaces will be maintained. However, additional symmetries consistent with the plane of symmetry may be introduced. The presence of these symmetries does not invalidate an application of the theory presented so far, provided that the base-ray lies in a plane of symmetry of Ⓚ. Beyond a single plane of symmetry, the next level of symmetry is two planes of symmetry. The centres of the surfaces must then lie in the intersection of the two planes of symmetry and, consequently, in addition to being plane symmetric Ⓚ has rotational symmetry, that is, Ⓚ is a symmetric system. Ⓚ may also be reversible or concentric but these cases will not be considered here. It is of interest to investigate what consequences, if any, rotational symmetry has on the theory. The principal result will be the existence of additional identities, the rotation identities, between the G-coefficients (§57-59).

Any plane containing the optical axis of a symmetric system $Ⓚ_S$ is a plane of symmetry of $Ⓚ_S$. The base-ray must lie in such a plane and this plane will be the meridional plane of $Ⓚ_S$. The pseudo-axis is taken to be the optical axis so that the x-axes of the basal coordinate systems Ⓒ are parallel to the optical axis and the object and image planes are normal to the x-axes. The axial coordinate system $Ⓒ_A$ associated with any surface has its origin at the pole of the surface, its x-axis along the optical axis and its y-axis in the meridional plane. With this choice of coordinates the axial and basal canonical variables are linearly related (§55) and the axial theory is recovered if the base-ray is taken along the optical axis (§60). It would then be possible to compute the primary coefficients of the system by applying the basal theory to the system. However, by means of the axial theory, coefficients up to the seventh order can be computed. Consequently there is nothing to be gained by taking the base-ray along the optical axis, in fact there is little point in taking the base-ray to be in the region where the fifth order axial theory gives good predictions of the aberrations produced by the system.

It is evident that the most likely application of the basal theory to a symmetric system would be to a wide angle system with the base-ray in the outer regions of the field. Provided a suitable base-ray is traced, the theory may be applied to investigate the image forming properties of an extreme wide angle system in regions of the field corresponding to field angles of greater than 180^o. Due to the singularity in $\underset{\sim}{V}$ at a field angle of 180^o, the coefficients computed for a base-ray lying on one side of this field angle cannot be used to predict the image qualities of the system for field angles on the other side of 180^o. Although the ray trace scheme of Ford[1] is suitable for tracing the base-ray when the field angle is less than 180^o, it may not be applied for rays of field angle greater than 180^o. However, only minor modifications[2] are required in order to make it valid in this case. Alternatively, a trigonometric trace may be used.

(b) To avoid confusion between basal variables and the corresponding axial variables the following convention is used. Quantities in the axial theory will be distinguished by a subscript "A": thus Q_A. Q_B will be the

---

1 Ford (1960)

2 See footnote 1, §87.

corresponding quantity referring to the base-ray and Q will be reserved for the basal equivalent of Q_A. Usually Q_A, Q_B and Q are simply related.

55. Relationships Between the Canonical Variables and Coordinates of the Basal and Axial Theories

(a) At a particular surface, the *axial* canonical variable $\underset{\sim}{Y}_A$ of a ray is the coordinate of the point of intersection of the ray with the polar tangent plane of the surface and the variable $\underset{\sim}{V}_A$ is given by

$$\underset{\sim}{V}_A = \underset{\sim}{\beta}_A/\alpha_A \quad .$$

Since Ⓒ$_A$ and Ⓒ are related by a translation, $\underset{\sim}{V}_A$ is just the $\underset{\sim}{V}_A$ defined in (5.1). Let the base-ray intersect Ⓕ at Ⓟ$_B$ whose coordinates in Ⓒ$_A$ are $(X_B, y_B, 0)$. Then, if the canonical variables of the base-ray are $\underset{\sim}{Y}_B$ and $\underset{\sim}{V}_B$,

$$\underset{\sim}{y}_B = \underset{\sim}{Y}_B + X_B\underset{\sim}{V}_B \quad . \tag{55.1}$$

Let Ⓡ be an arbitrary ray with axial canonical variables $\underset{\sim}{Y}_A$ and $\underset{\sim}{V}_A$ and basal canonical variables $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. By definition $\underset{\sim}{Y}$ is the coordinate (referred to Ⓟ$_B$ as origin) of the point of intersection of Ⓡ with the plane $x_A = X_B$. Thus

$$\underset{\sim}{Y} = \underset{\sim}{Y}_A + X_B\underset{\sim}{V}_A - \underset{\sim}{y}_B = \underset{\sim}{Y}_A - \underset{\sim}{Y}_B + X_B(\underset{\sim}{V}_A - \underset{\sim}{V}_B) \quad . \tag{55.2}$$

Since $\underset{\sim}{V}$ is given by (5.2), the axial and basal canonical variables are related by

$$\underset{\sim}{V} = \underset{\sim}{V}_A - \underset{\sim}{V}_B \quad , \quad \underset{\sim}{Y} = \underset{\sim}{Y}_A - \underset{\sim}{Y}_B + X_B\underset{\sim}{V} \quad . \tag{55.3}$$

(b) If Ⓡ has axial paracanonical coordinates $\underset{\sim}{S}_A$ and $\underset{\sim}{T}_A$, it is convenient for practical purposes to define the basal paracanonical coordinates of Ⓡ to be

$$\underset{\sim}{S} = \underset{\sim}{S}_A - \underset{\sim}{S}_B \quad , \quad \underset{\sim}{T} = \underset{\sim}{T}_A - \underset{\sim}{T}_B \quad , \tag{55.4}$$

where $\underset{\sim}{S}_B$ and $\underset{\sim}{T}_B$ are the axial paracanonical coordinates of the base-ray. Since

$$\underset{\sim}{S}_A = \sigma_A \underset{\sim}{Y}_{A1} + \bar{\sigma}_A \underset{\sim}{V}_{A1} \quad , \quad \underset{\sim}{T} = \tau_A \underset{\sim}{Y}_{A1} + \bar{\tau}_A \underset{\sim}{V}_{A1} \quad ,$$

(55.4) gives

$$\underset{\sim}{S} = \sigma_A(\underset{\sim}{Y}_{A1}-\underset{\sim}{Y}_{B1}) + \bar{\sigma}_A(\underset{\sim}{V}_{A1}-\underset{\sim}{V}_{B1}) = \sigma_A \underset{\sim}{Y}_1 + (\bar{\sigma}_A - X_{B1}\sigma_A)\underset{\sim}{V}_1 \quad ,$$

$$\underset{\sim}{T} = \tau_A(\underset{\sim}{Y}_{A1}-\underset{\sim}{Y}_{B1}) + \bar{\tau}_A(\underset{\sim}{V}_{A1}-\underset{\sim}{V}_{B1}) = \tau_A \underset{\sim}{Y}_1 + (\bar{\tau}_A - X_{B1}\tau_A)\underset{\sim}{V}_1 \quad . \tag{55.5}$$

If SPC are used in the axial theory $\bar{\sigma}_A = p_A\sigma_A$, $\bar{\tau}_A = \ell_{A1}\tau_A$ and hence

$$\underset{\sim}{S} = \sigma\underset{\sim}{Y}_1 + \bar{\sigma}\underset{\sim}{V}_1 \quad , \quad \underset{\sim}{T} = \tau\underset{\sim}{Y}_1 + \bar{\tau}\underset{\sim}{V}_1 \quad , \tag{55.6}$$

where

$$\sigma = \sigma_A \quad , \quad \bar{\sigma} = (p_A - X_{B1})\sigma \quad , \quad \tau = \tau_A \quad , \quad \bar{\tau} = (\ell_{A1} - X_{B1})\tau \quad . \tag{55.7}$$

The object and pupil planes of the axial theory will be used in the basal theory and are represented in the latter by $x_1 = p$, $x_1 = \ell_1$ respectively (see §6b). Thus

$$p_A = p + X_{B1} \quad , \quad \ell_{A1} = \ell_1 + X_{B1} \quad .$$

(55.7) is in fact just the pair of equations following (6.6) and SPC in the basal theory are automatically used. If the axial SPC are OE,OT etc., coordinates, then the basal SPC (55.4) are also said to be OE,OT etc., coordinates.

56. Aberrations

In the neighbourhood of the optical axis of a symmetric system imagery is perfect and pairs of conjugate planes and points are defined. The image plane and ideal image corresponding to a given object plane and point are uniquely defined and it is natural to employ these in the basal theory. The basal aberration will then be identical with the aberration defined in the axial theory and may be expressed in terms of the basal variables as follows:[1]

$$\underset{\sim}{\epsilon}_A = \underset{\sim}{H}_A - m_A\underset{\sim}{H}_{A1} = \underset{\sim}{H} - m_A\underset{\sim}{H}_1 + \underset{\sim}{H}_B - m_A\underset{\sim}{H}_{B1} \quad ,$$

where $\underset{\sim}{H}_B$ is the point of intersection of the base-ray with the image plane and m_A is the paraxial magnification. The ideal image of $\underset{\sim}{H}_{B1}$ is $m_A\underset{\sim}{H}_{B1}$.

Thus

$$\underset{\sim}{\epsilon}_A = \underset{\sim}{H} - m_A\underset{\sim}{H}_1 + \underset{\sim}{\epsilon}_B = \underset{\sim}{\epsilon} + \underset{\sim}{\epsilon}_B \quad ,$$

where $\underset{\sim}{\epsilon}_B$ is the aberration of Ⓡ$_B$ in Ⓗ, the appropriate object being O_{B1}. According to (7.3), $\underset{\sim}{\epsilon}$ is what would normally be taken as the basal aberration for distortion free imagery in the absence of (explicit) rotational symmetry.

It has already been mentioned that the most likely application of the basal theory to a symmetric system would be to a wide or extreme wide angle system. For various reasons, distortion is large and negative for such systems. First, the ideal image height tends to infinity with large field angles and the usable area of the image plane is quite small. Second, if the illumination of the image is to be uniform, negative distortion must be introduced.[2] What is required is that the blur of the image, represented by aberrations other than distortion, be small. Thus it is necessary to subtract out the distortion. This is unsatisfactory since in practice it can involve taking the differences between numbers which are quite large as compared to their residues. In such cases it is better to refer the aberration to some suitable member Ⓡ$_P$ of the pencil of rays by determining all the image heights and then subtracting that of

---

1 <u>Note</u>: $\underset{\sim}{H}_A = \underset{\sim}{Y}_A + \ell_A\underset{\sim}{V}_A = \underset{\sim}{Y} + \underset{\sim}{Y}_B - X_B\underset{\sim}{V} + \ell_A\underset{\sim}{V} + \ell_A\underset{\sim}{V}_B = \underset{\sim}{H}_B + \underset{\sim}{H}$,

where $\underset{\sim}{H}$ is given by (7.5).

2 Slussareff (1941).

Ⓡ$_P$. The location of the image is given by the image height of Ⓡ$_P$. It is desirable that Ⓡ$_P$ be taken as the principal ray of the pencil (§62).

In terms of the basal variables, the axial image height of any ray is given by

$$\underset{\sim}{H}_A = \underset{\sim}{H}_B + \underset{\sim}{h}_a\underset{\sim}{S} + \underset{\sim}{h}_b\underset{\sim}{T} + \sum_{n=2}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{\sim}{h}_{\mu\nu\tau}{}^{(n)} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} \quad . \tag{56.3}$$

Any ray is specified by four numbers, say $\underset{\sim}{S}$ and $\underset{\sim}{T}$. If the ray is known to be principal, two suffice, $\underset{\sim}{T}$ say. For a principal ray $\underset{\sim}{S}$ can be expressed in terms of $\underset{\sim}{T}$ (§65) and eliminated from (56.3) to give the image height of a principal ray as (§66)

$$\underset{\sim}{H}_P = \underset{\sim}{H}_B + \sum_{n=1}^{\infty} \sum_{\tau=0}^{n} \underset{\sim}{h}_{Pn\tau\tau}{}^{(n)} T_y{}^{n-\tau} T_z{}^{\tau} \quad . \tag{56.4}$$

If this image height is subtracted from $\underset{\sim}{H}_A$, the effect on (56.3) is simply to change the coefficients of distortion and exact values are obtained for the aberration coefficients. Thus the aberration referred to the principal ray is determined directly.

57. Invariants and Identities

(a) Throughout physics the existence of rotational symmetry of a system is associated with an invariant. For example, the angular momentum of a dynamical system about an axis of rotational symmetry of the system is a constant of the motion. Such is the case in optics; it is well known that an optical invariant exists for symmetric systems.[1] In the context

---

1 For example, M(4.8,3) and Herzberger (1958) equation (4.1). It is of interest to note that this invariant also exists for semi-symmetric systems, that is, systems with rotational symmetry but lacking a plane of symmetry, for example, an electron microscope.

of the axial theory, the optical invariant is

$$\underline{\underline{E}} = N\alpha(Z_A V_A - W_A Y_A) \quad .$$

As it appears above, $\underline{\underline{E}}$ is a function of the axial variables. However, by replacing $\underset{\sim}{Y}_A$ and $\underset{\sim}{V}_A$ by (55.2)

$$\underline{\underline{E}} = N\alpha(ZV - WY + V_B Z - Y_B W) \quad . \tag{57.1}$$

Since this is an invariant

$$\Delta\underline{\underline{E}} = 0 \quad , \tag{57.2}$$

where the Δ has the generalised meaning of §44-47.

(b) The existence of the invariant, or equivalently the associated symmetry property, is a constraint on the system and as such should manifest itself through the coefficients characterising the system. The number of degrees of freedom of the system is reduced. Hence it is to be expected that the number of independent $\underset{\sim}{G}$-coefficients is also reduced, that is, that there exist additional identities between the $\underset{\sim}{G}$-coefficients. These identities will be called the rotation identities of the system and may be obtained simultaneously with the extremal identities by endowing the characteristic function with the appropriate symmetry properties.[2] However, in §59 the identities are derived from the invariance of $\underline{\underline{E}}$.

(c) If the characteristic function $\underline{\underline{V}}$ of a symmetric system is referred to an arbitrary base-ray, $\underline{\underline{V}}$ will be formally indistinguishable from the characteristic of an asymmetric system. However, the rotational symmetry brings about certain identities between the coefficients of $\underline{\underline{V}}$. This may

---

[2] This was done in the context of the axial theory in Buchdahl (1965) §9b.

be seen as follows. Let $\underline{\underline{V}}$ be the point characteristic of an asymmetric system Ⓚ. Suppose a base-ray has been traced through Ⓚ and take its points of intersection with the first and last surfaces as base-points. The base-planes are any two planes, one through each of the two base-points. Rays are specified by $\underset{\sim}{y}$ and $\underset{\sim}{y}'$, their points of intersection with the base-planes, referred to coordinate systems whose origins are the base-points and whose x-axes are normal to the base-planes. The point characteristic is then a function of $\underset{\sim}{y}$ and $\underset{\sim}{y}'$ and has the formal expansion

$$\underline{\underline{V}} = \underline{\underline{V}}_0 + \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \underline{\underline{v}}_{\mu\nu\tau}^{(n)} y'^{\,n-\mu} y^{\mu-\nu} z'^{\,\nu-\tau} z^{\tau} \quad . \tag{57.3}$$

(<u>Note</u>: The coefficients $\underline{\underline{v}}_{\mu\nu\tau}^{(n)}$ are <u>not</u> pseudo-coefficients.)

Now suppose the system is rotationally symmetric and the ray $(\underset{\sim}{y}_0,\underset{\sim}{y}_0')$ is the axis of symmetry Ⓐ of Ⓚ. In this case the base-planes are conveniently chosen normal to Ⓐ, so that the x-axes of the coordinate systems associated with object and image spaces of Ⓚ are parallel to Ⓐ. If Ⓚ is rotated about Ⓐ, the coordinates of any ray Ⓡ become $\bar{\underset{\sim}{y}},\bar{\underset{\sim}{y}}'$ where

$$\bar{y} = y_0 + (y-y_0)\cos\theta + (z-z_0)\sin\theta$$

$$\bar{z} = z_0 - (y-y_0)\sin\theta + (z-z_0)\cos\theta \quad , \tag{57.4}$$

with similar equations for $\bar{\underset{\sim}{y}}'$. $\underset{\sim}{y}$ and $\underset{\sim}{y}'$ are the coordinates of Ⓡ before the rotation is performed. If the rotation is infinitesimal,

$$\bar{y} = y + (z-z_0)d\theta \quad , \quad \bar{z} = z - (y-y_0)d\theta \quad .$$

The rotational symmetry of Ⓚ implies that

$$\underline{\underline{V}}(\bar{\underset{\sim}{y}},\bar{\underset{\sim}{y}}') = \underline{\underline{V}}(\underset{\sim}{y},\underset{\sim}{y}') \quad ,$$

and for an infinitesimal rotation

$$\underline{\underline{V}}(y+(z-z_0)d\theta \;,\; y'+(z'-z_0')d\theta \;,\; z-(y-y_0)d\theta \;,\; z'-(y'-y_0')d\theta) = \underline{\underline{V}}(\underset{\sim}{y},\underset{\sim}{y}') \;.$$

By expanding the left hand member,

$$z_0\underline{\underline{V}}_{,y} - y_0\underline{\underline{V}}_{,z} + z_0'\underline{\underline{V}}_{,y'} - y_0'\underline{\underline{V}}_{,z'} = z\underline{\underline{V}}_{,y} - y\underline{\underline{V}}_{,z} + z'\underline{\underline{V}}_{,y'} - y'\underline{\underline{V}}_{,z'} \quad . \tag{57.5}$$

If (57.3) is substituted into this and the coefficients of equal powers of $\underset{\sim}{y}$ and $\underset{\sim}{y}'$ are equated, relationships between the coefficients of $\underline{\underline{V}}$ are obtained. The $\underline{n}$th order terms on the left will contain only the ($\underline{n}$+1)th order coefficients of $\underline{\underline{V}}$, whereas the right hand member contains the $\underline{n}$th order coefficients. Thus certain relations are found between the $\underline{n}$th and ($\underline{n}$+1)th order coefficients of $\underline{\underline{V}}$. In other words, identities are found between the coefficients of $\underline{\underline{V}}$, and these are due solely to the rotational symmetry of K.

If Ⓚ is plane symmetric, the y- and y'-axes of the coordinate systems associated with the base-planes may be taken to lie in the plane of symmetry. Thus $z_0 = z_0' = 0$ and (57.5) becomes

$$y_0\underline{\underline{V}}_{,z} + y_0'\underline{\underline{V}}_{,z'} = y\underline{\underline{V}}_{,z} - z\underline{\underline{V}}_{,y} + y'\underline{\underline{V}}_{,z'} - z'\underline{\underline{V}}_{,y'} \quad . \tag{57.6}$$

Since $\underline{\underline{V}}$ is invariant under reflection in Ⓜ, $\underline{\underline{V}}_{,z}$ must change sign under such a reflection and the number of terms of the $\underline{n}$th order on the left of (57.6) is $\underline{\underline{N}}_{zn}$. Thus there are $\underline{\underline{N}}_{zn}$ identities between the ($\underline{n}$+1)th order coefficients of $\underline{\underline{V}}$. Since these determine the $\underline{n}$th order imagery of Ⓚ, it is to be expected that the number of rotation identities between the $\underline{n}$th order $\underset{\sim}{G}$-coefficients is just $\underline{\underline{N}}_{zn}$. This is confirmed in the next section.

58. Enumeration of the Rotation Identities

The rotation identities may be obtained by forming the paracanonical expansions of $\underline{\underline{E}}$ and $\underline{\underline{E}}_1$ and substituting these into (57.2). It follows that the coefficients of each power of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ must be zero and the resulting set of equations constitute the rotation identities. It is readily verified by inspection of (57.1) that the $\underline{n}$th order $\underset{\sim}{G}$-coefficients turn up in the $\underline{n}$th order terms of $\underline{\underline{E}}$ solely through those terms in $\underline{\underline{E}}$ linear in W and Z. Consequently, the rotation identities express the $\underline{n}$th order $\underset{\sim}{G}_z$-coefficients in terms of the lower order $\underset{\sim}{G}$-coefficients. Moreover $\underline{\underline{E}}$ transforms like Z under a reflection in Ⓜ and hence there are $\underline{\underline{N}}_{zn}$ identities obtained from the invariance of $\underline{\underline{E}}$. Thus the number of rotation identities is $\underline{\underline{N}}_{zn}$, in agreement with the number obtained in §57c.

It has already been seen that there are $\underline{\underline{N}}_{In}$ extremal identities (§43) between the $\underline{n}$th order $\underset{\sim}{G}$-coefficients. Thus, for a symmetric system there are $\underline{\underline{N}}_{gn} - \underline{\underline{N}}_{In} - \underline{\underline{N}}_{zn}$ independent $\underline{n}$th order $\underset{\sim}{G}$-coefficients. Substitution of the appropriate numbers according to (43.1,2) and (24.5) gives the number of <u>independent, nth order G-coefficients of a symmetric system with a meridional base-ray</u> to be:

$$\tfrac{1}{4}(n+2)(n+4) \quad \text{when } \underline{n} \text{ is } \underline{\text{even}}, \text{ or } \quad \tfrac{1}{4}(n+3)^2 \quad \text{when } \underline{n} \text{ is } \underline{\text{odd}} .$$

Reference to M§19 shows that this is in <u>excess</u> of the corresponding axial $\underset{\sim}{G}$-coefficients. Thus the more general base-ray has produced a proliferation of coefficients. The number of independent coefficients of the first three orders are compared in Table 58/1 for the following three cases: (i) a symmetric system in the axial theory, (ii) a symmetric system with a meridional base-ray, and (iii) a general plane symmetric system with a meridional base-ray.

<u>TABLE 58/1 Numbers of Independent Coefficients</u>

	First Order	Second Order	Third Order
Axial theory	3	0	6
Symmetric system, meridional base-ray	4	6	9
Plane symmetric system, meridional base-ray	6	10	19

59. Determination of the Rotation Identities

(a) It is of advantage both when deriving the rotation identities and when working with them numerically, to introduce two new quantities, U and a, defined by

$$U = N_z(ZV-YW) \quad , \quad a = \alpha/\alpha_B \ . \tag{59.1}$$

At the first surface U and $\underline{\underline{E}}$ may be expressed in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ by applying (6.3) to (59.1) and (57.1). Hence

$$U_1 = (N_{z1}/g)(S_zT_y-S_yT_z) \quad ,$$

$$\underline{\underline{E}}_1 = (N_1\alpha_1/g)(S_zT_y-S_yT_z+S_zT_{By}-S_{By}T_z) \quad .$$

In terms of U and a

$$\underline{\underline{E}} = aU + N_za(ZV_B-Y_BW) \quad ,$$

$$\underline{\underline{E}}_1 = a_1U_1 + (N_{z1}a_1/g)(S_zT_{By}-S_{By}Y_z) \quad . \tag{59.2}$$

U, a and $\underline{\underline{E}}$ may be expanded in series of the form (24.1). It should be noted that U and $\underline{\underline{E}}$ change sign under reflection in Ⓜ, whereas a does not. Write

$$\underline{\underline{E}} = \sum_{n=1}^{\infty} \underline{\underline{e}}^{(n)} \quad , \quad U = \sum_{n=2}^{\infty} u^{(n)} \quad , \quad a = 1 + \sum_{n=1}^{\infty} a^{(n)} .$$

It follows from (57.2) that

$$\underline{\underline{e}}^{(n)} - \underline{\underline{e}}_1^{(n)} = 0 \quad . \tag{59.3}$$

The $\underline{\underline{e}}^{(n)}$ can be written down from the first member of (59.2), and in particular at the first surface, from the second member of (59.2). For the first three orders these are

$$\underline{\underline{e}}^{(1)} = N_z(z^{(1)}V_B - Y_B w^{(1)}) \quad ,$$

$$\underline{\underline{e}}^{(2)} = N_z(z^{(2)}V_B - Y_B w^{(2)}) + a^{(1)}\underline{\underline{e}}^{(1)} + u^{(2)} \quad ,$$

$$\underline{\underline{e}}^{(3)} = N_z(z^{(3)}V_B - Y_B w^{(3)}) + a^{(1)}\underline{\underline{e}}^{(2)} + (a^{(2)} - a^{(1)2})\underline{\underline{e}}^{(1)} + u^{(3)} \quad ,$$

$$\underline{\underline{e}}_1^{(1)} = (N_{z1}/g)(S_z T_{By} - S_{By} T_z) \quad ,$$

$$\underline{\underline{e}}_1^{(2)} = a_1^{(1)}\underline{\underline{e}}_1^{(1)} + U_1 \quad ,$$

$$\underline{\underline{e}}_1^{(3)} = a_1^{(1)}\underline{\underline{e}}_1^{(2)} + (a_1^{(2)} - a_1^{(1)2})\underline{\underline{e}}_1^{(1)} \quad . \tag{59.4}$$

It follows that the first order identities are contained in

$$N_z(z^{(1)}V_B - Y_B w^{(1)}) = (N_{z1}/g)(S_z T_{By} - S_{By} T_z) \quad ,$$

where $z^{(1)}$ and $w^{(1)}$ are given by (14.2). Hence the <u>parabasal rotation identities</u> are

$$N_z(z_a V_B - w_a Y_B) = (N_{z1}/g)T_{By} \quad , \quad N_z(z_b V_B - w_b Y_B) = -(N_{z1}/g)S_{By} \quad .$$

Consider the terms $N_z(z^{(n)}V_B - Y_B w^{(n)})$. According to (40.1)

$$N_z(z^{(n)}V_B - Y_B w^{(n)}) = N_z[(z_a V_B - w_a Y_B)\delta_z^{(n)} + (z_b V_B - w_b Y_B)\delta_w^{(n)}]$$

$$= (N_{z1}/g)[\delta_z^{(n)} T_{By} - \delta_w^{(n)} S_{By}] \quad , \tag{59.5}$$

where the parabasal identities have been used. The second and third order terms (59.3) can now be written in the form

$$(N_{z1}/g)(\delta_w^{(2)} S_{By} - \delta_z^{(2)} T_{By}) = \underline{\underline{e}}_1^{(1)}\Delta a^{(1)} + \Delta u^{(2)} \quad ,$$

$$(N_{z1}/g)(\delta_w^{(3)} S_{By} - \delta_z^{(3)} T_{By}) = \underline{\underline{e}}_1^{(1)}(\Delta a^{(2)} - a^{(1)}\Delta a^{(1)}) + u_1^{(2)}\Delta a^{(1)} + u^{(3)} \quad , \tag{59.6}$$

by substituting (59.5) into (59.4) and the latter into (59.3). Use was made of (59.3) to replace $\underline{\underline{e}}^{(n)}$ by $\underline{\underline{e}}_1^{(n)}$ in the above and $\underline{\underline{e}}_1^{(2)}$ was eliminated from the third order equations by the fifth member of (59.4). It is evident that the $\underline{n}$th order equation corresponding to (59.6) will be of the form

$$(N_{z1}/g)(\delta_w{}^{(n)}S_{By}-\delta_z{}^{(n)}T_{By}) = \text{sums of products of lower order terms.}$$

The coefficient of a typical term on the left hand side of this is $(N_{z1}/g)(`G_{z\mu\nu\tau a}{}^{(n)}S_{By}+`G_{z\mu\nu\tau a}{}^{(n)}T_{By})$ and thus the general $\underline{n}$th order rotation identity is of the form

$$G_{z\mu\nu\tau a}{}^{(n)}S_{By} + G_{z\mu\nu\tau b}{}^{(n)}T_{By} = \text{sums of products of lower order coefficients.}$$

The lower order identities may be read off from (59.6). For example, the coefficient of S_yS_z in the first of (59.6) gives

$$S_{gz3a}S_{By} + S_{gz3b}T_{By} = (N_{z1}/g)T_{By}\Delta a_a + s_{u3} \quad ,$$

and the coefficient of $S_y{}^2S_z$ in the second of (59.6) gives

$$T_{gz3a}S_{By} + T_{gz3b}T_{By} = (N_{z1}/g)(\Delta s_{a1}-a_a\Delta a_a)T_{By} + t_{u1} \quad .$$

(b) The rotation identities of any order may in principle be written down in this manner and for the first three orders are presented in Table 59/2. However, the u- and a-coefficients have yet to be found in terms of the $\underset{\sim}{G}$-coefficients. From (45.1)

$$a = 1 - \alpha_B\beta_BV + \tfrac{1}{2}\alpha_B{}^2(3\beta_B{}^2-1)V^2 - \tfrac{1}{2}\alpha_B{}^2W^2 + 0(3) \quad .$$

By the usual method of substituting for $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ according to (25.6), the a-coefficients are found, and appear in Table 59/1. In a similar manner the u-coefficients may be read off

$$u^{(2)} = N_z(z^{(1)}v^{(1)}-w^{(1)}y^{(1)}) \quad ,$$

$$u^{(3)} = N_z(z^{(2)}v^{(1)}-w^{(2)}y^{(1)}+z^{(1)}v^{(2)}-w^{(1)}y^{(2)}) \quad .$$

TABLE 59/1 The a- and u-coefficients

$$a_a = -\alpha_B\beta_B v_a \qquad a_b = -\alpha_B\beta_B v_b ,$$

$$s_{a1} = -\alpha_B\beta_B s_{v1} + \tfrac{1}{2}\alpha_B^{\,2}(3\beta_B^{\,2}-1)v_a^{\,2} \qquad s_{a8} = -\alpha_B\beta_B s_{v8} - \tfrac{1}{2}\alpha_B^{\,2} w_a^{\,2}$$

$$s_{a2} = -\alpha_B\beta_B s_{v2} + \alpha_B^{\,2}(3\beta_B^{\,2}-1)v_a v_b \qquad s_{a9} = -\alpha_B\beta_B s_{v9} - \alpha_B^{\,2} w_a w_b$$

$$s_{a5} = -\alpha_B\beta_B s_{v5} + \tfrac{1}{2}\alpha_B^{\,2}(3\beta_B^{\,2}-1)v_b^{\,2} \qquad s_{a10} = -\alpha_B\beta_B s_{v10} - \tfrac{1}{2}\alpha_B^{\,2} w_b^{\,2} ,$$

$$s_{u3} = N_z(z_a v_a - w_a y_a)$$

$$s_{u4} = N_z(z_b v_a - w_b y_a)$$

$$s_{u6} = N_z(z_a v_b - w_a y_b)$$

$$s_{u7} = N_z(z_b v_b - w_b y_b) ,$$

$$t_{u3} = N_z(s_{z3}v_a - s_{w3}y_a + z_a s_{v1} - w_a s_{y1})$$

$$t_{u4} = N_z(s_{z4}v_a - s_{w4}y_a + z_b s_{v1} - w_b s_{y1})$$

$$t_{u6} = N_z(s_{z6}v_a - s_{w6}y_a + z_a s_{v2} - w_a s_{y2} + s_{z3}v_b - s_{w3}y_b)$$

$$t_{u7} = N_z(s_{z7}v_a - s_{w7}y_a + z_b s_{v2} - w_b s_{y2} + s_{z4}v_b - s_{w4}y_b)$$

$$t_{u12} = N_z(s_{z6}v_b - s_{w6}y_b + z_a s_{v5} - w_a s_{y5})$$

$$t_{u13} = N_z(s_{z7}v_b - s_{w7}y_b + z_b s_{v5} - w_b s_{y5})$$

$$t_{u17} = N_z(z_a s_{v8} - w_a s_{y8})$$

$$t_{u18} = N_z(z_a s_{v9} - w_a s_{y9} + z_b s_{v8} - w_b s_{y8})$$

$$t_{u19} = N_z(z_a s_{v10} - w_a s_{y10} + z_b s_{v9} - w_b s_{y9})$$

$$t_{u20} = N_z(z_b s_{v10} - w_b s_{y10}) .$$

TABLE 59/2 The Rotation Identities

$$N_z(z_aV_B-w_aY_B) = (N_{z1}/g)T_{By}$$

$$N_z(z_bV_B-w_bY_B) = -(N_{z1}/g)T_{By} ,$$

$$S_{gz3a}S_{By} + S_{gz3b}T_{By} = s_{u3} + (N_{z1}/g)T_{By}\Delta a_a$$

$$S_{gz4a}S_{By} + S_{gz4b}T_{By} = s_{u4} + (N_{z1}/g)(1-S_{By}\Delta a_a)$$

$$S_{gz6a}S_{By} + S_{gz6b}T_{By} = s_{u6} - (N_{z1}/g)(1-T_{By}\Delta a_b)$$

$$S_{gz7a}S_{By} + S_{gz7b}T_{By} = s_{u7} - (N_{z1}/g)S_{By}\Delta a_b ,$$

$$T_{gz3a}S_{By} + T_{gz3b}T_{By} = t_{u3} + (N_{z1}/g)T_{By}(\Delta s_{a1}-a_a\Delta a_a)$$

$$T_{gz4a}S_{By} + T_{gz4b}T_{By} = t_{u4} - (N_{z1}/g)[\Delta a_a+S_{By}(\Delta s_{a1}-a_a\Delta a_a)]$$

$$T_{gz6a}S_{By} + T_{gz6b}T_{By} = t_{u6} + (N_{z1}/g)[\Delta a_a+T_{By}(\Delta s_{a2}-a_a\Delta a_b-a_b\Delta a_a)]$$

$$T_{gz7a}S_{By} + T_{gz7b}T_{By} = t_{u7} - (N_{z1}/g)[\Delta a_b+S_{By}(\Delta s_{a2}-a_a\Delta a_b-a_b\Delta a_a)]$$

$$T_{gz12a}S_{By} + T_{gz12b}T_{By} = t_{u12} + (N_{z1}/g)[\Delta a_b+T_{By}(\Delta s_{a5}-a_b\Delta a_b)]$$

$$T_{gz13a}S_{By} + T_{gz13b}T_{By} = t_{u13} - (N_{z1}/g)S_{By}(\Delta s_{a5}-a_b\Delta a_b)$$

$$T_{gz17a}S_{By} + T_{gz17b}T_{By} = t_{u17} + (N_{z1}/g)T_{By}\Delta s_{a8}$$

$$T_{gz18a}S_{By} + T_{gz18b}T_{By} = t_{u18} - (N_{z1}/g)(S_{By}\Delta s_{a8}-T_{By}\Delta s_{a9})$$

$$T_{gz19a}S_{By} + T_{gz19b}T_{By} = t_{u19} - (N_{z1}/g)(S_{By}\Delta s_{a9}-T_{By}\Delta s_{a10})$$

$$T_{gz20a}S_{By} + T_{gz20b}T_{By} = t_{u20} - (N_{z1}/g)S_{By}\Delta s_{a10} .$$

At the first surface, the <u>only non-zero</u> u-coefficients are

$$s_{u4j} = -s_{u6j} = -N_{zj}/g \quad , \quad j = 1 \quad . \tag{59.7}$$

The second and third order u-coefficients at any other surface are also given in Table 59/1.

The rotation identities presented in Table 59/2 could, if desired, be expressed directly in terms of the $\underset{\sim}{G}$-coefficients by replacing the a- and u-coefficients by the appropriate expressions in terms of the $\underset{\sim}{G}$-coefficients. Although the second order identities retain their simple form, those of the third order do not and there seems little to gain by doing this. Further, the rotation identities can in some cases be rearranged by means of the extremal identities, but once again this seems of little value.

60. <u>The Axial Limit</u>

Consider the case where the base-ray of the basal theory is taken along the optical axis. From the results of §55 it is clear that the basal and axial canonical variables, and similarly the paracanonical coordinates, are then identical. Moreover, the pseudo-parameters take simple values: the angles of incidence and refraction are both 0°, the normal $\vec{n}_B$ and direction-cosines $\vec{\beta}_B$ of the base-ray are (1,0,0), $\underset{\sim}{V}_B$ and $\underset{\sim}{I}_B$ are zero and $\bar{R}_B$ is simply (1-k). If these values are inserted into the basic equations of the basal theory it is seen that these equations reduce to the corresponding equations of the axial theory. (10.12) becomes M(4.6,7), (27.3,4) become M(62.4,5) and so on. Thus, <u>when the base-ray lies along the optical axis of a symmetric system the basal theory reduces to the axial theory</u>. In particular the expansions of corresponding quantities in the two theories, for example, $\Delta\underset{\sim}{G}_A$ and $\Delta\underset{\sim}{G}$, must be identical. Thus all even order basal coefficients must vanish and the third order terms must be identical with the primary terms of the axial theory.

In the <u>axial theory</u> $\Delta\underset{\sim}{G}_a$ is given by (see M(8.3))

$$\Delta\underset{\sim}{G}_a = a\underset{\sim}{S}\xi + \bar{a}\underset{\sim}{T}\xi + b\underset{\sim}{S}\eta + \bar{b}\underset{\sim}{T}\eta + c\underset{\sim}{S}\zeta + \bar{c}\underset{\sim}{T}\zeta$$

where

$$\xi = \underset{\sim}{S}\cdot\underset{\sim}{S} \quad , \quad \eta = \underset{\sim}{S}\cdot\underset{\sim}{T} \quad , \quad \zeta = \underset{\sim}{T}\cdot\underset{\sim}{T} \quad .$$

If this is expressed explicitly in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$

$$\Delta G_{ya} = aS_y^3 + (\bar{a}+b)S_y^2T_y + (\bar{b}+c)S_yT_y^2 + aS_yS_z^2 + bS_yS_zT_z + cS_yT_y^2 + \bar{c}T_y^3 + + \bar{a}T_yS_z^2 + \bar{b}T_yS_zT_z + \bar{c}T_yT_z^2 + 0(5) \quad ,$$

$$\Delta G_{za} = aS_y^2S_z + \bar{a}S_y^2T_z + bS_yT_yS_z + \bar{b}S_yT_yT_z + cT_y^2S_z + \bar{c}T_y^2T_z + aS_z^3 + + (\bar{a}+b)S_z^2T_z + (\bar{b}+c)S_zT_z^2 + \bar{c}T_z^3 + 0(5) \quad .$$

In the case under consideration this series must be identical with the corresponding series for $\Delta\underset{\sim}{G}_a$ in the <u>basal theory</u>. Simple relations must therefore exist between the primary coefficients and the appropriate third order basal coefficients. By summing over surfaces the same relationships will hold between intermediate and final coefficients. These are given in Table 60/1 which, it should be noted, is valid for both the <u>a</u>- and the <u>b</u>-components of the coefficients. Similar though more complex relationships will exist between the higher order coefficients of the two theories.

When the base-ray is not the optical axis the basal coefficients are formally functions of the paracanonical coordinates of the base-ray. Moreover, in the absence of rotational symmetry they are also functions of the decentring parameters. In either case the coefficients may be formally expanded as series in the appropriate variables. This is, in effect, what Stephan and Maréchal[1] (and others) have done for the case of weak decentring.

TABLE 60/1 Axial Limit of the Third Order Basal Coefficients

$T_{gy1} = A$	$T_{gz3} = A$
$T_{gy2} = (\bar{A}+B)$	$T_{gz4} = \bar{A}$
$T_{gy5} = (\bar{B}+C)$	$T_{gz6} = B$
$T_{gy8} = A$	$T_{gz7} = \bar{B}$
$T_{gy9} = B$	$T_{gz12} = C$
$T_{gy10} = C$	$T_{gz13} = \bar{C}$
$T_{gy11} = \bar{C}$	$T_{gz17} = A$
$T_{gy14} = \bar{A}$	$T_{gz18} = (\bar{A}+B)$
$T_{gy15} = \bar{B}$	$T_{gz19} = (\bar{B}+C)$
$T_{gy16} = \bar{C}$	$T_{gz20} = \bar{C}$

---

1 Stephan (1949); Maréchal (1949).

X. PRINCIPAL RAYS, VIGNETTING AND THE ENTRANCE PUPIL

61. On the Entrance Pupil

(a) Of paramount importance in the design of lens systems, especially of wide angle systems, is the amount of energy transmitted by the system in the form of light from a given object. Consider a pencil of light from a point object. This pencil will be of limited extent due to vignetting either by the surfaces or by the diaphragm. The transmitted energy will depend on the solid angle defined by the pencil, which in turn is determined by the vignetting. Thus it is necessary to find some manner in which the vignetting of the pencil may be conveniently obtained and represented. The most suitable solution to this problem is based on the idea of an entrance pupil for the system. Apart from a few comments in §67 vignetting by the surfaces is not considered. However, vignetting by the diaphragm is considered in detail and is by far the most important case.

Consider a perfect optical system Ⓚ with a plane, circular diaphragm Ⓓ in its image space[1] and neglect vignetting by the surfaces. Pencils of rays from a given object are then limited solely by Ⓓ. Since the system is perfect there exists a plane (in the object space) conjugate to Ⓓ and in this plane the (perfect) image of Ⓓ is a circle Ⓔ. All rays which pass through the rim of Ⓓ pass through the rim of Ⓔ and all rays transmitted by Ⓚ (i.e., which pass through Ⓓ) pass through Ⓔ and conversely. Consequently a ray will be transmitted by Ⓚ if and only if it passes through Ⓔ. This is formally expressed by saying that

---

[1] That part of Ⓚ which is on the image-side of Ⓓ is of no concern in the ensuing discussion. Only the part of Ⓚ preceding Ⓓ is required to be perfect.

the rays are vignetted by Ⓔ. Ⓔ is called the entrance pupil of Ⓚ and accurately represents the vignetting of Ⓚ. Whether or not a particular ray is vignetted by Ⓚ is determined simply by its point of intersection with the entrance pupil.

The entrance pupil is the same for all pencils and is an absolute property of the system. As the inclination of the pencil to the plane of the entrance pupil decreases, the solid angle defined by the pencil must decrease and the illumination of the image will also decrease. This is expressed by the cosine-fourth-power law of illumination,[2] and may be counteracted by one or more of the following.[3] (i) The image surface could be taken so as to be always normal to the principal rays in the image space of Ⓚ. (ii) Aberrations between Ⓔ and Ⓓ could be introduced so that the area of Ⓔ increases for oblique pencils. (iii) Negative distortion of the image could be introduced. In particular, the illumination dE arising from an area dS of the entrance pupil of a symmetric system is given by Reiss to be[4]

$$dE = (B \cos^4\psi_1 h_k'/f^2 H_k') \frac{dh_k'}{dH_k'} \, dS \quad , \qquad (61.1)$$

where B is the intrinsic brightness of the object, ψ_1 the inclination of the (elementary) pencil in the object space, f the focal length of the system, h_k' the ideal image height and H_k' the distorted image height. It has already been seen how the distortion can be determined. It remains to find the area of the entrance pupil.

---

2 For example, see Born and Wolf (1964) §4.8.3 and Slussareff (1941).

3 Slussareff (1941). The question of distortion is discussed in detail by Reiss (1948).

4 Reiss (1948), equation 17.

(b) It is now assumed that aberrations are present in the imagery of Ⓓ by Ⓚ. Accordingly a sharp image of Ⓓ is no longer formed and a more general definition of the entrance pupil must be found. First the plane of Ⓔ must be chosen. If there exists a base-ray for the system such that parabasal imagery of Ⓓ is stigmatic, it is reasonable to take the plane of the corresponding parabasal image of Ⓓ as the plane of the entrance pupil for all base-rays. In particular, this will be done for symmetric systems (cf. §56 in relation to the choice of image plane and ideal image). Otherwise the choice is more or less arbitrary and a plane through one of the foci of the basal point of Ⓓ would suffice. The entrance pupil Ⓔ corresponding to a given pencil is then defined to be the set of points of intersection of the rays of this pencil with the chosen plane.

Any ray from a point object will intersect both Ⓔ and Ⓓ in a single point (birefringent media having been excluded from the outset). Thus a pencil of rays from a given point object maps points in Ⓓ into points in Ⓔ. This mapping is one-to-one. Since Ⓔ is in the object space there is one and only one ray through each point of Ⓔ, and each of these rays can intersect the diaphragm in one and only one point. Now suppose that two distinct rays from the object intersect in a point in Ⓓ, mapping two points of Ⓔ into a single point of Ⓓ. In this case there exists a base-ray in whose parabasal region the object has a focus in or very close to Ⓓ. It is obvious that in such a case Ⓓ will not act as a stop and some other choice for Ⓓ must be made. Consequently the mapping of points in Ⓔ into points in Ⓓ is one-to-one. Moreover, there being no surfaces with singular points (such as the vertex of a cone), the mapping is continuous.

Let ℓ_1 and ℓ_2 be any two closed curves in Ⓓ such that ℓ_2 is interior to and distinct from ℓ_1. ℓ_1 and ℓ_2 are mapped into two closed curves in Ⓔ: ℓ_1' and ℓ_2' respectively. Since ℓ_1 and ℓ_2 are non-intersecting, so are ℓ_1' and ℓ_2'. Suppose ℓ_2' is not interior to ℓ_1'. Then if ℓ_2 is continuously shrunk to a point, ℓ_2' cannot be shrunk to a point without violating either the continuity or the one-to-oneness of the mapping. Therefore ℓ_2' must be interior to ℓ_1' and any point in Ⓓ interior to a given closed curve ℓ in Ⓓ will be mapped into a point interior to the map of ℓ. Ⓔ is a connected area. It follows that those rays which intersect Ⓓ in points on its rim will intersect the plane of the entrance pupil in a closed curve and those rays which are transmitted by Ⓚ will intersect the plane of the entrance pupil in points enclosed by this curve.

(c) It is frequently stated[5] that it is impossible to determine the shape of the entrance pupil by any other means apart from tracing a large number of rays through the system and determining whether or not they are vignetted by the diaphragm. Suppose the $\underset{\sim}{G}$-coefficients are known at either of the surfaces adjacent to Ⓓ. It is then possible (§63) to express the coordinates $\underset{\sim}{Y}_d$ of the points of intersection of rays with Ⓓ as series in the paracanonical coordinates of the rays and examination of $\underset{\sim}{Y}_d$ in the usual manner reveals whether or not the ray is vignetted. If SPC are used, $\underset{\sim}{S}$ is the point of intersection of the ray $\underset{\sim}{S}$, $\underset{\sim}{T}$ with Ⓔ. Since the mapping (by a pencil from a given point object) of points in Ⓓ into points in Ⓔ is one-to-one, $\underset{\sim}{Y}_d$ is a single valued function of $\underset{\sim}{S}$ and vice versa. Thus the series for $\underset{\sim}{Y}_d$ may be inverted, and $\underset{\sim}{S}$ expressed as a

---

5 Gardner (1947) Part VIII; Herzberger (1958) p.105; Stavroudis and Sutton (1965) p.2 .

series in $\underset{\sim}{Y}_d$ and $\underset{\sim}{T}$ (§64). This enables the boundary of Ⓔ to be determined directly.

62. On Principal Rays

In the literature the term "principal ray" has various connotations. Some authors apply the term "principal", "chief" or "central" to any ray to which neighbouring rays are referred; that is, to the base-ray of this thesis. A principal ray is also defined to be a ray passing through the centre of the diaphragm. However, the most common definition of a principal ray is a ray passing through the centre of the entrance pupil. If imagery between Ⓔ and Ⓓ is perfect these last two definitions are of course equivalent.

Consider for the moment a symmetric system. Any paraxial ray through the centre of the diaphragm also passes through the centre of the entrance pupil, and vice versa. Due to the presence of aberrations this is not the case for general rays. For small field angles the aberrations associated with the entrance pupil and diaphragm can be frequently neglected (see M§41c) and principal rays then pass through the centres of both the entrance pupil and the diaphragm. Suppose the system is a wide angle or extreme wide angle system, for example, a sky lens.[1] Assuming that, except perhaps in the extreme outer regions of the field, Ⓓ is solely responsible for vignetting, the ray through the centre of the diaphragm is certainly transmitted by the system. However, quite crude ray tracing suffices to show that for such a system rays which are greatly inclined to the axis

---

[1] Hill (1924); Havlicék (1951).

and which pass through the centre of the paraxial entrance pupil are not transmitted by the system (§69,71). If not vignetted, they are either totally internally reflected or miss a surface altogether. For a particular system this is demonstrated by accurate ray traces in §71. The usual definition of principal ray is inappropriate for such systems and since the ray through the centre of the diaphragm is transmitted, the following terminology is offered. A ray is a proper principal ray, or simply a principal ray, if it passes through the centre of the diaphragm. A ray passing through the centre of the paraxial entrance pupil is an improper principal ray. The term improper principal ray only has meaning for systems possessing a base-ray in whose parabasal region imagery is stigmatic, for example, a symmetric system.

A general definition of entrance pupil was given in §61b. The centre of the entrance pupil is now defined to be the point of intersection of the principal ray with the entrance pupil. In the absence of aberrations between Ⓓ and Ⓔ this reduces to the usual definition of the centre of the entrance pupil.

63. Vignetting by the Diaphragm

In the plane $x = d_d$ of the object space of the ith surface insert a circular diaphragm Ⓓ of radius ρ_d centred on the point $\vec{y}_d = (d_d, y_d, 0)$. (Note: (i) By definition d_d must be negative, and (ii) a subscript d in the position of a surface indicator (rank 5) implies association with the diaphragm.) If Ⓚ is symmetric $y_d = -y_{Bi}$ where y_{Bi} is the y-coordinate (in axial coordinates) of the point of intersection of the base-ray with $Ⓕ_i$. In the plane of Ⓓ set up a coordinate system $Ⓒ_d$ whose origin is the centre of Ⓓ and whose y- and z-axes are respectively parallel to the y- and z-axes of Ⓒ$_i$. The diaphragm coordinates $\underset{\sim}{Y}_d$ of the ray ($\underset{\sim}{S}$,$\underset{\sim}{T}$) are the coordinates in Ⓒ$_d$ of the point of intersection of the ray with Ⓓ. The diaphragm coordinates of the base-ray are

$$\underset{\sim}{Y}_{Bd} = d_d \underset{\sim}{V}_{Bi} - \underset{\sim}{y}_d \quad . \tag{63.1}$$

Thus, for any ray whose canonical variables at the ith surface are $\underset{\sim}{Y}_i$,$\underset{\sim}{V}_i$

$$\underset{\sim}{Y}_d = \underset{\sim}{Y}_{Bd} + \underset{\sim}{Y}_i + d_d \underset{\sim}{V}_i \quad . \tag{63.2}$$

If the ray is a principal ray, $\underset{\sim}{Y}_d = 0$, and in particular if the base-ray is itself a principal ray, $\underset{\sim}{Y}_{Bd} = 0$.

Define the diaphragm coefficients by

$$\underset{\sim}{y}_{ad} = \underset{\sim}{y}_{ai} + d_d \underset{\sim}{v}_{ai} \quad , \quad \underset{\sim}{y}_{bd} = \underset{\sim}{y}_{bi} + d_d \underset{\sim}{v}_{bi} \quad ,$$

$$\underset{\sim}{y}_{\mu\nu\tau d}{}^{(n)} = ({}^{\backprime}\underset{\sim}{G}_{\mu\nu\tau i}{}^{(n)} | \underset{\sim}{y}_d) \quad . \tag{63.3}$$

In terms of these coefficients (63.2) becomes

$$\underset{\sim}{Y}_d = \underset{\sim}{Y}_{Bd} + \underset{\sim}{y}_{ad}(\underset{\sim}{S}+\underset{\sim}{\delta}_y) + \underset{\sim}{y}_{bd}(\underset{\sim}{T}+\underset{\sim}{\delta}_v)$$

$$= \underset{\sim}{Y}_{Bd} + \underset{\sim}{y}_{ad}\underset{\sim}{S} + \underset{\sim}{y}_{bd}\underset{\sim}{T} + \sum_{n=2}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{\sim}{y}_{\mu\nu\tau d}{}^{(n)} S_y{}^{n-\mu} T_y{}^{\mu-\nu} S_z{}^{\nu-\tau} T_z{}^{\tau} \quad . \tag{63.4}$$

Consequently, given the $\underset{\sim}{G}$-coefficients of the ith surface up to the nth order, the diaphragm coordinates of the ray ($\underset{\sim}{S}$,$\underset{\sim}{T}$) may be computed to the nth order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ by means of (63.4). This ray will be vignetted by the diaphragm if and only if

$$Y_d{}^2 + Z_d{}^2 > \rho_d{}^2 \quad . \tag{63.5}$$

If $\underset{\sim}{Y}_d$ was calculated to the nth order and (63.5) is satisfied, the ray ($\underset{\sim}{S}$,$\underset{\sim}{T}$) is vignetted (by the diaphragm) to the nth order.

64. The Shape of the Entrance Pupil

(a) If SPC are used a pencil of rays from a point object is specified by the two quantities $\underset{\sim}{T}$ determined by the object. Each member of the pencil intersects the entrance pupil in some point specified by $\underset{\sim}{S}$. (63.4) enables the point of intersection with Ⓓ of the ray $\underset{\sim}{S}$ of the pencil $\underset{\sim}{T}$ to be calculated. The condition that this ray pass through the rim of the diaphragm is

$$Y_d^{\ 2} + Z_d^{\ 2} = \rho_d^{\ 2} \quad , \qquad (64.1)$$

and according to the results of §61b, the rim of Ⓔ is given by those values of $\underset{\sim}{S}$ which satisfy (64.1). Also, according to §61c the series (63.4) is a single valued function of $\underset{\sim}{S}$ for given $\underset{\sim}{T}$. Moreover (63.4) may be inverted so as to express $\underset{\sim}{S}$ as a series in $\underset{\sim}{T}$ and $\underset{\sim}{Y}_d$. Formally, this series is

$$\underset{\sim}{S} = \underset{\sim}{S}_0 + \sum_{n=1}^{\infty} \underset{\sim}{s}^{(n)} = \underset{\sim}{S}_0 + \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \underset{\sim}{s}_{\mu\nu\tau}^{(n)} Y_d^{\ n-\mu} T_y^{\ \mu-\nu} Z_d^{\ \nu-\tau} T_z^{\ \tau} \quad , \qquad (64.2)$$

where the $\underset{\sim}{s}_{\mu\nu\tau}^{(n)}$ are called entrance pupil coefficients and $\underset{\sim}{S}_0$ is zero if the base-ray is a principal ray. Introduce polar coordinates (ρ,θ) into $Ⓒ_d$ so that

$$Y_d = \rho \cos \theta \quad , \quad Z_d = \rho \sin \theta \quad . \qquad (64.3)$$

Then, (64.1) becomes

$$\rho = \rho_d \quad ,$$

and if this and (64.3) are substituted into (64.2), the latter is the parametric equation of the rim of the entrance pupil.

It will now be assumed that the base-ray is a principal ray. Then $\underset{\sim}{Y}_{Bd} = 0$. The entrance pupil coefficients are determined by substituting (64.2) into (63.4), and equating the coefficients of the resulting identity in $\underset{\sim}{Y}_d$, $\underset{\sim}{T}$. The first, second and third order terms of these identities are

$$y_{ad}s_y^{(1)} + y_{bd}T_y = Y_d \qquad z_{ad}s_z^{(1)} + z_{bd}T_z = Z_d \quad , \qquad (64.4)$$

$$y_{ad}s_y^{(2)} + s_{y1d}s_y^{(1)2} + s_{y2d}s_y^{(1)}T_y + s_{y5d}T_y^{\,2} + s_{y8d}s_z^{(1)2} + s_{y9d}s_z^{(1)}T_z + \\ + s_{y10d}T_z^{\,2} = 0$$

$$z_{ad}s_z^{(2)} + s_{z3d}s_y^{(1)}s_z^{(1)} + s_{z4d}s_y^{(1)}T_z + s_{z6}T_y s_z^{(1)} + s_{z7d}T_yT_z = 0 \,, \quad (64.5)$$

$$y_{ad}s_y^{(3)} + 2s_{y1d}s_y^{(2)}s_y^{(1)} + s_{y2d}s_y^{(2)}T_y + 2s_{y8d}s_z^{(2)}s_z^{(1)} + s_{y9d}s_z^{(2)}T_z + \\ + t_{y1d}s_y^{(1)3} + t_{y2d}s_y^{(1)2}T_y + t_{y5d}s_y^{(1)}T_y^{\,2} + t_{y8d}s_y^{(1)}s_z^{(1)2} + \\ + t_{y9d}s_y^{(1)}s_z^{(1)}T_z + t_{y10d}s_y^{(1)}T_z^{\,2} + t_{y11d}T_y^{\,3} + t_{y14d}T_ys_z^{(1)2} + \\ + t_{y15d}T_ys_z^{(1)}T_z + t_{y16d}T_yT_z^{\,2} = 0$$

$$z_{ad}s_z^{(3)} + s_{z3d}(s_y^{(2)}s_z^{(1)} + s_z^{(2)}s_y^{(1)}) + s_{z4d}s_y^{(2)}T_z + s_{z6d}T_ys_z^{(2)} + \\ + t_{z3d}s_y^{(1)2}s_z^{(1)} + t_{z4d}s_y^{(1)2}T_z + t_{z6d}s_y^{(1)}T_ys_z^{(1)} + \\ + t_{z7d}s_y^{(1)}T_yT_z + t_{z12d}T_y^{\,2}s_z^{(1)} + t_{z13d}T_y^{\,2}T_z + t_{z17d}s_z^{(1)3} + \\ + t_{z18d}s_z^{(1)2}T_z + t_{z19d}s_z^{(1)}T_z^{\,2} + t_{z20d}T_z^{\,3} = 0 \quad . \qquad (64.6)$$

Consider the first order terms (64.4). If $\underset{\sim}{s}^{(1)}$ is substituted into these according to (64.2) two identities are obtained:

$$y_{ad}(p_{sy1}Y_d + p_{sy2}T_y) + y_{bd}T_y \equiv Y_d \quad , \quad z_{ad}(p_{sz3}Z_d + p_{sz4}T_z) + z_{bd}T_z \equiv Z_d \; .$$

Thus the first order entrance pupil coefficients are given by

$$y_{ad}p_{s1} - 1 = 0 \; , \; y_{ad}p_{sy2} + y_{bd} = 0 \; , \; z_{ad}p_{sz3} - 1 = 0 \; , \; z_{ad}p_{sz4} + z_{bd} = 0 \; .$$

In a similar manner the corresponding second order entrance pupil coefficients are determined from (64.5,6) and are given in Table 64/1. The third order coefficients are likewise obtained from (64.6). The first two of these are

$$y_{ad}t_{sy1} = -2s_{y1d}s_{sy1}p_{sy1} - t_{y1d}p_{sy1}^{3} ,$$

$$z_{ad}t_{sz3} = -s_{z3d}(s_{sy1}p_{sz3}+s_{sz3}p_{sy1}) + t_{z3d}p_{sy1}^{2}p_{sz3} . \qquad (64.7)$$

Table 64/1 also contains the third order coefficients required in §65.

By means of the series (64.2) the entrance pupil may be determined, in principle, for any pencil and for any aperture of the system to any order in the coordinates of the pencil and the aperture. Apart from the tracing of a principal ray as base-ray, no ray tracing is involved. It is seen that a knowledge of the $\underset{\sim}{G}$-coefficients of the system enables the shape of the entrance pupil to be determined.[1]

(b) Let the parametric equation of a closed plane curve be $x = x(t)$, $y = y(t)$. If the parameter t takes values $t_0 \leq t \leq t_1$ whilst the curve is traversed once, the area enclosed by the curve is the magnitude of[2]

$$A = \int_{t_0}^{t_1} y(t)x'(t)dt$$

where $x'(t) = dx(t)/dt$. If (64.3), with $\rho = \rho_d$, is substituted into (64.2) the resulting series for $\underset{\sim}{S}$ in terms of $\underset{\sim}{T}$ and ρ_d is a parametric

---

[1] The corresponding analysis can be carried out in the context of the axial theory. If the derivatives of the axial $\underset{\sim}{G}$-coefficients with respect to the system parameters are computed along the lines of M part III, the change in the entrance pupil as a result of a change in the construction of the system may be determined.

[2] Courant (1959) p.273 .

TABLE 64/1 Entrance Pupil Coefficients

$$y_{ad}p_{sy1} = 1 \qquad z_{ad}p_{sz3} = 1$$

$$y_{ad}p_{sy2} = -y_{bd} \qquad z_{ad}p_{sz4} = -z_{bd} \quad ,$$

$$y_{ad}s_{sy1} = -s_{y1d}p_{sy1}^{2} \qquad y_{ad}s_{sy8} = -s_{y8d}p_{sz3}^{2}$$

$$y_{ad}s_{sy2} = -p_{sy1}(2s_{y1d}p_{sy2}+s_{y2d}) \qquad y_{ad}s_{sy9} = -p_{sz3}(2s_{y8d}p_{sz4}+s_{y9d})$$

$$y_{ad}s_{sy5} = -s_{y1d}p_{sy2}^{2}-s_{y2d}p_{sy2}-s_{y5d} \qquad y_{ad}s_{sy10} = -s_{y8d}p_{sz4}^{2}-s_{y9d}p_{sz4}-s_{y10d} \quad ,$$

$$z_{ad}s_{sz3} = -s_{z3d}p_{sy1}p_{sz3}$$

$$z_{ad}s_{sz4} = -p_{sy1}(s_{z3d}p_{sz4}+s_{z4d})$$

$$z_{ad}s_{sz6} = -p_{sz3}(s_{z3d}p_{sy2}+s_{z6d})$$

$$z_{ad}s_{sz7} = -p_{sy2}(s_{z3d}p_{sz4}+s_{z4d}) - s_{z6d}p_{sz4} - s_{z7d} \quad ,$$

$$y_{ad}t_{sy11} = -(2s_{y1d}p_{sy2}+s_{y2d})s_{sy5} - t_{y1d}p_{sy2}^{3} - t_{y2d}p_{sy2}^{2} - t_{y5d}p_{sy2} - t_{y11d}$$

$$y_{ad}t_{sy16} = -(2s_{y1d}p_{sy2}+s_{y2d})s_{sy10} - (2s_{y8d}p_{sz4}+s_{y9d})s_{sz7} - (t_{y8d}p_{sz4}^{2}+$$
$$+t_{y9d}p_{sz4}+t_{y10d})p_{sy2} - t_{y14d}p_{sz4}^{2} - t_{y15d}p_{sz4} - t_{y16d} \quad ,$$

$$z_{ad}t_{sz13} = -(s_{z3d}p_{sz4}+s_{z4d})s_{sy5} - (s_{z3d}p_{sy2}+s_{z6d})s_{sz7} - (t_{z3d}p_{sz4}+t_{z4d})\cdot$$
$$\cdot p_{sy2}^{2} - (t_{z6d}p_{sz4}+t_{z7d})p_{sy2} - t_{z12d}p_{sz4} - t_{z13d}$$

$$z_{ad}t_{sz20} = -(s_{z3d}p_{sz4}+s_{z4d})s_{sy10} - t_{z17d}p_{sz4}^{3} - t_{z18d}p_{sz4}^{2} - t_{z19d}p_{sz4} -$$
$$- t_{z20d} \quad .$$

representation of the boundary of Ⓔ. The parameter is θ and takes values $0 \leq \theta \leq 2\pi$. Correct to the third order in ρ and $\underset{\sim}{T}$, S_y and $\partial S_z/\partial\theta$ are given by

$$S_y = p_{sy1}\rho \cos\theta + p_{sy2}T_y + s_{sy1}\rho^2\cos^2\theta + s_{sy2}\rho T_y \cos\theta + s_{sy5}T_y^2 + \\ + s_{sy8}\rho^2\sin^2\theta + s_{sy9}\rho T_z \sin\theta + s_{sy10}T_z^2 + 0(3)$$

$$\partial S_z/\partial\theta = p_{sz3}\rho \cos\theta + s_{sz3}\rho^2(\cos^2\theta - \sin^2\theta) - s_{sz4}\rho T_z \sin\theta + s_{sz6}\rho T_y \cos\theta + \\ + 0(3) \quad .$$

Since $\int_0^{2\pi} \cos^m\theta \sin^n\theta \, d\theta$ vanishes unless m and n are both even, the area of Ⓔ is given to the fourth order by

$$A_E = \int_0^{2\pi} S_y(\partial S_z/\partial\theta)d\theta = \pi\rho_d^2[p_{sy1}p_{sz3} + (p_{sy1}s_{sz6} + p_{sz3}s_{sy2})T_y] + 0(4) \quad (64.8)$$

The third order terms in $\underset{\sim}{S}$ enable A_E to be calculated correct to the fifth order. The additional terms in (64.8) are then

$$A^{(4)} = \pi\rho_d^2\{\tfrac{1}{4}\rho_d^2[p_{sz3}(3t_{sy1}+t_{sy8}) + p_{sy1}(3t_{sz17}+t_{sz3}) + 2s_{sz3}(s_{sy1}-s_{sy8})] + \\ + (p_{sz3}t_{sy5} + s_{sz6}s_{sy2} + p_{sy1}t_{sz12})T_y^2 + (p_{sz3}t_{sy10} - s_{sz4}s_{sy9} + p_{sy1}t_{sz19})T_z^2].$$

More important than the area of Ⓔ is the cross-section of the pencil measured normal to the principal ray of the pencil. If the SPC of the principal ray are $\underset{\sim}{S}_P, \underset{\sim}{T}$, then its inclination in the object space is given by the direction tangents

$$\underset{\sim}{V}_{AP} = \underset{\sim}{V}_{B1} + \underset{\sim}{v}_{a1}\underset{\sim}{S}_P + \underset{\sim}{v}_{b1}\underset{\sim}{T} \quad .$$

The area A of the projection of Ⓔ onto a plane normal to the principal ray is then

$$A = A_E(1 + \underset{\sim}{V}_{AP}\cdot\underset{\sim}{V}_{AP})^{-\frac{1}{2}} \quad . \quad (64.9)$$

If only the first order terms in (63.4) or (64.2) are used to determine vignetting, it is evident from (64.8) that the area of Ⓔ will be predicted to be independent of the pencil. Consequently, in order to ensure that the illumination of the image does not decrease with increasing field angle it is necessary that there be aberrations of the second or higher orders associated with imagery between Ⓔ and Ⓓ.

65. SPC of Proper Principal Rays

(a) By definition proper principal rays pass through the centre of the diaphragm. Substitution of $\rho = 0$ into (64.2) gives a series for $\underset{\sim}{S}_P$ in terms of $\underset{\sim}{T}$:

$$\underset{\sim}{S}_P = \underset{\sim}{S}_0 + \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \underset{\sim}{s}_{n\mu\mu}^{(n)} T_y^{\,n-\mu} T_z^{\,\mu} \quad . \tag{65.1}$$

If the entrance pupil coefficients are known, the paracanonical coordinates $\underset{\sim}{S}_P$ of the principal ray of the pencil $\underset{\sim}{T}$ may be determined. If the base-ray is itself principal, $\underset{\sim}{S}_P = 0$, $\underset{\sim}{T} = 0$ satisfies (65.1). Then $\underset{\sim}{S}_0 = 0$ and the coefficients $s_{n\mu\mu}^{(n)}$ of the first three orders are given in Table 64/1. To the fourth order

$$S_{Py} = p_{sy2}T_y + s_{sy5}T_y^{\,2} + s_{sy10}T_z^{\,2} + t_{sy11}T_y^{\,3} + t_{sy16}T_yT_z^{\,2} + 0(4) \quad ,$$

$$S_{Pz} = p_{sz4}T_z + s_{sz7}T_yT_z + t_{sz13}T_y^{\,2}T_z + t_{sz20}T_z^{\,3} + 0(4) \quad , \tag{65.2}$$

if the base-ray is a principal ray.

(b) In general, however, the base-ray need not be a principal ray. Then $\underset{\sim}{S}_0 = (S_0, 0)$ where S_0 is to be determined. Identities corresponding to (64.4-6) are obtained by substituting (64.2) (with $S_0 \neq 0$) into (63.4) and differ from (64.4-6) solely by terms proportional to S_0. There is in addition a zeroth order term:

$$Y_{Byd} + y_{ad}S_0 + \sum_{n=2}^{\infty}\sum_{\mu=0}^{n} \underset{\sim}{y}_{\mu 00d}{}^{(n)}S_0{}^n = 0 \quad , \tag{65.3}$$

the solution of which gives S_0. Note that the diaphragm coefficients of <u>all</u> orders are required in order to determine S_0. This is also the case for each coefficient $\underset{\sim}{s}_{n\mu\mu}{}^{(n)}$. By way of illustration, the first order identity corresponding to the first of (64.4) is

$$(y_{ad}+2s_{y1d}S_0+3t_{y1d}S_0{}^2+\dots)s_y{}^{(1)} + (y_{bd}+s_{y2d}S_0+t_{y2d}S_0{}^2+\dots)T_y \equiv Y_d \quad .$$

Thus

$$(y_{ad}+2s_{y1d}S_0+3t_{y1d}S_0{}^2+\dots)p_{sy1} = 1 \quad ,$$

$$(y_{ad}+2s_{y1d}S_0+3t_{y1d}S_0{}^2+\dots)p_{sy2} + (y_{bd}+s_{y2d}S_0+t_{y2d}S_0{}^2+\dots) = 0 \quad . \tag{65.4}$$

Due to the fact that (65.3,4) and the corresponding higher order equations contain the $\underset{\sim}{G}$-coefficients of all orders it is necessary to make an approximation before the coefficients in (65.1) may be computed. Two cases present themselves. Either S_0 is sufficiently small so that powers of S_0 greater than the third, say, may be neglected, or the $\underset{\sim}{G}$-coefficients of order higher than the third, say, are negligible. Both cases may be accommodated by neglecting terms proportional to $S_0{}^{n-\mu}T_y{}^{\mu-\nu}T_z{}^{\nu}$ for $n > 3$, say. By way of an example, the coordinates of <u>meridional</u> principal rays are determined from

$$S_{Py} = S_0 + p_{sy2}T_y + s_{sy5}T_y{}^2 + t_{sy11}T_y{}^3$$

where

$$t_{y1d}S_0{}^3 + s_{y1d}S_0{}^2 + y_{ad}S_0 + y_{Bd} = 0 \quad ,$$

$$p_{sy2} = \Gamma[y_{bd}+S_0s_{y2d}+S_0{}^2t_{y2d}] \quad ,$$

$$s_{sy5} = \Gamma[s_{y5d}+p_{sy2}s_{y2d}+p_{sy2}{}^2s_{y1d}+S_0(t_{y5d}+2p_{sy2}t_{y2d}+3p_{sy2}{}^2t_{y1d})]$$

$$t_{sy11} = \Gamma[s_{sy5}(s_{y2d}+2p_{sy2}s_{y1d})+t_{y11d}+p_{sy2}t_{y5d}+p_{sy2}{}^2t_{y2d}+p_{sy2}{}^3+t_{y1d}] \quad , \tag{65.5}$$

and

$$\Gamma = -(y_{ad}+2S_0s_{y1d}+3S_0{}^2t_{y1d})^{-1} \quad . \tag{65.6}$$

66. Aberrations Referred to Principal Rays

(a) In view of the difficulties of §65b, the base-ray will be assumed to be a principal ray, in which case $\underset{\sim}{S}_P$ is given by (65.1) with $S_0 = 0$. The image height of any ray is given by the series (42.6). Since principal rays are specified by the two numbers $\underset{\sim}{T}$, the image height $\underset{\sim}{H}_P$ of a principal ray must also be determined by $\underset{\sim}{T}$. The same may be said for the direction tangents $\underset{\sim}{V}_P$ of principal rays. Formally, $\underset{\sim}{H}_P$ and $\underset{\sim}{V}_P$ are given by

$$\underset{\sim}{H}_P = \underset{\sim}{h}_{Pb}T + \sum_{n=2}^{\infty}\sum_{\mu=0}^{n} \underset{\sim}{h}_{Pn\mu\mu}{}^{(n)} T_y{}^{n-\mu} T_z{}^{\mu} \quad ,$$

$$\underset{\sim}{V}_P = \underset{\sim}{v}_{Pb}\underset{\sim}{T} + \sum_{n=2}^{\infty}\sum_{\mu=0}^{n} \underset{\sim}{v}_{Pn\mu\mu}{}^{(n)} T_y{}^{n-\mu} T_z{}^{\mu} \quad , \qquad (66.1)$$

where "P" is a subscript of rank 1. The $\underset{\sim}{h}_P$- and $\underset{\sim}{v}_P$-coefficients are determined by eliminating $\underset{\sim}{S}$ from the paracanonical expansions of $\underset{\sim}{H}$ and $\underset{\sim}{V}$ by means of (65.1). Table 66/1 gives the $\underset{\sim}{h}_P$-coefficients. The $\underset{\sim}{v}_P$-coefficients are given by a similar set of equations where the $\underset{\sim}{h}$- and $\underset{\sim}{h}_P$-coefficients are replaced by the corresponding $\underset{\sim}{v}$- and $\underset{\sim}{v}_P$-coefficients.

In an image plane displaced by an amount x from the position in which the $\underset{\sim}{h}$-coefficients are $\underset{\sim}{h}_{\mu\nu\tau}{}^{(n)}$, the corresponding $\underset{\sim}{h}$-coefficients are

$$\bar{\underset{\sim}{h}}_{\mu\nu\tau}{}^{(n)} = \underset{\sim}{h}_{\mu\nu\tau}{}^{(n)} + x\underset{\sim}{v}_{\mu\nu\tau}{}^{(n)} \quad .$$

The image height of the principal ray in the displaced image plane will be $\bar{\underset{\sim}{H}}_P$ and the $\bar{\underset{\sim}{h}}_P$-coefficients are determined by substituting $\bar{\underset{\sim}{h}}_{\mu\nu\tau}{}^{(n)}$ for $\underset{\sim}{h}_{\mu\nu\tau}{}^{(n)}$ in Table 66/1. Since the $\underset{\sim}{h}_P$-coefficients are linear in the corresponding $\underset{\sim}{h}$-coefficients,

$$\bar{\underset{\sim}{h}}_{P\mu\nu\tau}{}^{(n)} = \underset{\sim}{h}_{P\mu\nu\tau}{}^{(n)} + x\underset{\sim}{v}_{P\mu\nu\tau}{}^{(n)} \quad . \qquad (66.2)$$

(b) It is now possible to refer the aberration of the members of a pencil of rays to the principal ray of the pencil by defining the

TABLE 66/1 The $\underline{h}_p$-coefficients

$$p_{Phy2} = h_{ya}p_{sy2} + h_{yb} \quad ,$$

$$s_{Phy5} = h_{ya}s_{sy5} + s_{hy1}p_{sy2}^{2} + s_{hy2}p_{sy2} + s_{hy5}$$

$$s_{Phy10} = h_{ya}s_{sy10} + s_{hy8}p_{sz4}^{2} + s_{hy9}p_{sz4} + s_{hy10} \quad ,$$

$$t_{Phy11} = h_{ya}t_{sy11} + (2s_{hy1}p_{sy2}+s_{hy2})s_{sy5} + t_{hy1}p_{sy2}^{3} + t_{hy2}p_{sy2}^{2} + t_{hy5}p_{sy2} + t_{hy11}$$

$$t_{Phy16} = h_{ya}t_{sy16} + (2s_{hy1}p_{sy2}+s_{hy2})s_{sy10} + (2s_{hy8}p_{sz4}+s_{hy9})s_{sz7} + (t_{hy8}p_{sz4}^{2}+t_{hy9}p_{sz4}+t_{hy10})p_{sy2} + t_{hy14}p_{sz4}^{2} + t_{hy15}p_{sz4} + t_{hy16} \quad .$$

$$p_{Phz4} = h_{za}p_{sz4} + h_{zb} \quad ,$$

$$s_{Phz7} = h_{za}s_{sz7} + (s_{hz3}p_{sz4}+s_{hz4})p_{sy2} + s_{hz6}p_{sz4} + s_{hz7} \quad ,$$

$$t_{Phz13} = h_{za}t_{sz13} + (s_{hz3}p_{sz4}+s_{hz4})s_{sy5} + (s_{hz3}p_{sy2}+s_{hz6})s_{sz7} + (t_{hz3}p_{sz4}+t_{hz4})p_{sy2}^{2} + (t_{hz6}p_{sz4}+t_{hz7})p_{sy2} + t_{hz12}p_{sz4} + t_{hz13}$$

$$t_{Phz20} = h_{za}t_{sz20} + (s_{hz3}p_{sz4}+s_{hz4})s_{sy10} + t_{hz17}p_{sz4}^{3} + t_{hz18}p_{sz4}^{2} + t_{hz19}p_{sz4} + t_{hz20} \quad .$$

aberration of any ray to be

$$\underset{\sim}{\varepsilon} = \underset{\sim}{H} - \underset{\sim}{H}_P \quad . \tag{66.3}$$

The corresponding aberration coefficients are

$$\underset{\sim}{E}_{\mu\nu\tau}{}^{(n)} = \underset{\sim}{h}_{\mu\nu\tau}{}^{(n)} - \underset{\sim}{h}_{P\mu\nu\tau}{}^{(n)} \quad , \tag{66.4}$$

it being understood that $\underset{\sim}{h}_{P\mu\nu\tau}{}^{(n)} = 0$ unless $\nu = \tau$, $\mu = n$. Only the distortion coefficients are affected. Note that there are coefficients corresponding to first order distortion:

$$E_{yb} = h_{yb} - h_{Pyb} = -h_{ya}p_{sy2} \quad , \quad E_{zb} = -h_{za}p_{sz4} \quad .$$

67. Vignetting by the Surfaces

Although it is assumed that vignetting by the surfaces is not important it is as well to consider briefly how vignetting by the surfaces could be analysed in the context of the basal theory. The $\underset{\sim}{y}$- and $\underset{\sim}{v}$-coefficients at each surface can be readily obtained from the corresponding intermediate $\underset{\sim}{G}$-coefficients. Thus the canonical variables $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ for any ray can be determined at each surface. If the point of incidence of the ray with the $\underline{i}$th surface is (X_i, y_i, z_i) then

$$\underset{\sim}{y}_i = \underset{\sim}{Y}_i + X_i\underset{\sim}{V}_i + X_i\underset{\sim}{V}_{Bi} \quad .$$

In the course of computing the $\underset{\sim}{g}$-coefficients the $\hat{x}$-coefficients were computed. These are available for use and X can be determined to the same order in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ as $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ themselves are known. The equation above then yields y and z at each surface. If the radius of the rim of Ⓕ$_i$ is ρ_i the ray is vignetted by Ⓕ$_i$ when

$$(\underset{\sim}{y}_i - \underset{\sim}{n}_i/c_i)^2 > \rho_i^2 \quad .$$

If $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ were calculated to the <u>n</u>th order, and this inequality is satisfied, the ray is said to be <u>vignetted by the ith surface to the nth order</u>. If vignetting by a surface does take place, the symmetry of astigmatic aberrations is usually destroyed.

68. <u>On the Choice of Base-Ray</u>

It is natural that the base-ray should be chosen so as to have some special significance for the system. The choice of a proper principal ray as a base-ray has several advantages. First, it is almost certainly transmitted by the system and lies in the centre of the aperture of a pencil of rays about the base-ray. Suppose the aperture of the system is such that the predicted aberrations of such a pencil (using the basal theory) agree satisfactorily with ray traces over all but a small portion of the aperture of the pencil. The overall agreement can then be expected to be quite good. Now suppose the base-ray is still a member of this pencil but lies in the outer regions of the aperture of the pencil. The extent of the neighbourhood of the base-ray over which satisfactory agreement can be obtained will have changed only a little. However, a large number of the rays in this region will now be vignetted whereas a corresponding number of rays outside the region of good agreement will be transmitted by the system and contribute to the image. The overall agreement will consequently not be as good as when the base-ray was in the centre of the aperture of the pencil.

Second, if the base-ray is not a principal ray the formulae of the preceding sections of this chapter will not be applicable and the more general but approximate formulae must be developed according to §65(b). In view of the importance of obtaining a reliable estimate of the cross section of the pencil it is desirable that the base-ray be a principal ray. In the numerical examples to follow the base-rays have all been taken to be principal rays with the exception of the one chosen so as to demonstrate the effect of the choice of base-ray on the predicted aberrations (§81). It is important to note that the general theory developed prior to the present chapter is valid for <u>any base-ray lying in the plane of symmetry of the system</u>, and not just to principal rays.

XI. THE COEFFICIENTS OF A SKY-LENS

69. The Sky-Lens of Havlicěk

The remainder of this thesis is devoted to a numerical illustration of the application of the equations of the preceding chapters to a wide angle, symmetric lens system and to an investigation of the quality of the basal predictions (i.e., predictions made using the basal coefficients) of the aberrations of the system. The system Ⓚ chosen for this purpose is a two-component sky-lens due to Havlicěk,[1] designed to operate at f/10 with a focal length of 11 mm. and a full field of 200°. The system is presented diagrammatically in fig.1 and its constructional data, after reduction to unit focal length, is given in Table 69/1. The data given by Havlicěk[1] and Flügge[2] corresponds to a focal length of 10.3 mm. However, when considering the effect of a particular aberration on the state of correction of the system it is sufficient to assume a nominal focal length of 1 cm.

It is assumed that the second surface has a clear aperture equal to its radius. The ray passing through the centre of the diaphragm and with the maximum possible field angle consistent with the given system data then makes an angle ψ_3 with the optical axis in the air space between the second and third components, where

$$V_d = \tan \psi_3 = [c_2(d_2'+d_d)-1]^{-1} = 1.184 \quad . \qquad (69.1)$$

This corresponds to a full field of about 192° (see Table 71/1), somewhat less than the prescribed 200°. The maximum field angle may be increased

---

1 Havlicěk (1951).

2 Flügge (1955) p.318.

by either moving the diaphragm (D,fig.1) nearer the front element (C) of Ⓚ or by making the active surface of $Ⓕ_2$ somewhat more than a hemisphere (see §70,71). The path of the extreme ray through the system is shown as ray (d) in fig.1.

The front element (C) of Ⓚ acts as a field compressor, reducing the field from 200° down to 100°, and produces a virtual image of the hemisphere. The second member (V) is a positive element and produces on the photographic plate (P) a real image of the compressed field.

TABLE 69/1 Constructional Details of the Sky-Lens

j	1	2	3	4	k
c_j	0.41189	2.5054	-0.61003	-2.5054	---
N_j	1	1.5230	1	1.5230	1
d_j	---	0.19422	0.88372	0.11653	---

The diaphragm is situated before the third surface: $d_d = -0.14761$. At an aperture of f/10, the radius of the diaphragm is $\rho_d = 0.08760$.

70. The Axial Coefficients of the Sky-Lens

(a) For systems such as the above the object is assumed to be at infinity, in which case there is no difference between $O\bar{T}$ and $O\bar{E}$ coordinates (M§13,38a). Assume for the moment that p = 0, that is, canonical coordinates are being used. The corresponding paraxial and primary *a*-

coefficients are presented in Tables 70/1,2 and, incidentally, were computed using the same computer programme as that used to obtain basal coefficients (see §60). In order to eliminate primary coma, p may be chosen such that $\bar{A}_{ak}' = 0$. From M(34.41) this requires

$$p = -\bar{A}_{ak}'/A_{ak}' = 0.5606 \quad ,$$

for which the corresponding position of the diaphragm is

$$d_d = -(y_{b3}-y_{a3}p)/(v_{b3}-v_{a3}p) = -0.1555 \quad ,$$

and differs slightly from the value -0.1476 obtained from Flügge. With $d_d = -0.1555$, (69.1) gives $V_d = 1.213$ and a total field of 196^o. Corresponding to $d_d = -0.1476$, p = 0.5632. Using this value of p and $O\bar{T}$ coordinates, the paraxial, primary, secondary and tertiary coefficients of Ⓚ are given in Tables 70/3,4. (Note: The sign conventions relevant to these coefficients are those of this thesis.)

(b) Inspection of Table 70/4 reveals the following features. Spherical aberration is governed principally by the last surface with small amounts of the opposite sign contributed by the second surface. In fact, these two surfaces, the steepest, govern virtually all the coefficients. The exception is distortion which is governed weakly by the first and third surfaces. The higher the power of ρ (the aperture), associated with a certain coefficient, the stronger is the control of the second and last surfaces, with the last dominating. There is no evidence at all from Table 70/4 that any attempt was made during the design of the system to balance the aberrations of one surface with those of another. The interesting feature is that the angles of incidence and refraction of principal rays are always small at the second and last surfaces, of the order of 3^o, whereas at the others they are large (see

Table 71/2). Moreover, for rays of a given inclination in the object space, the angles of incidence and refraction remain virtually constant at the first surface, whereas they differ widely at the last (cf. Tables 71/2,3). Fig. 2 compares the spot diagram for $V_1 = 0.5$ and 0.75 obtained using the primary, secondary and tertiary coefficients of Ⓚ with the corresponding diagrams obtained by ray tracing. It is seen that the axial predictions break down before $V_1 = 0.75$ ($\psi_1 = 37^\circ$) but are very good at $V_1 = 0.5$ ($\psi_1 = 27^\circ$).

(c) Although it is not intended to use this system with values of p other than 0.5632, it is of interest to investigate the consequences of the choice of some other value for p, for instance 0.5606, such that $\bar{A}_{ak}' = 0$. A glance at Table 70/4 shows that primary and secondary linear coma are both of the same sign. However, by taking $p < 0.5606$, primary coma is negative and could be used to balance secondary coma. (Note that such a value of p would further increase the field angle.) Distortion is quite insensitive to the changes of p contemplated (of the order of 0.005) and astigmatism changes by a very small amount, viz. the coefficient $\bar{\bar{B}}_{ak}'$ governing the astigmatic difference has its maximum value of -0.0051 when $\bar{A}_{ak}' = 0$ as compared to the value -0.0058 in Table 70/4. Finally, spherical aberration, which is independent of p, is so large that it will mask other types of aberration for pencils near the axis of the system, though not necessarily in the outer regions of the field (§72). Thus the only apparent consequence of small changes of p is to alter the total field angle of Ⓚ. However, since the focal length is short, it is doubtful whether manufacturing tolerances would allow a distinction between values of p such as 0.5606 and 0.5632, say.

TABLE 70/1 Paraxial Coefficients for Canonical Coordinates

	1	2	3	4	final
y_a	1.0000	0.97253	1.9083	2.0359	2.0359
y_b	0.0000	0.12753	1.1589	1.2765	1.2765
v_a	0.0000	-0.14144	1.0589	1.0951	-1.0000
v_b	1.0000	0.65660	1.1671	1.0091	-0.13582

TABLE 70/2 Primary Coefficients for Canonical Coordinates

$A_k' = -58.300$ $\bar{A}_k' = -32.687$

$B_k' = -65.373$ $\bar{B}_k' = -36.657$

$C_k' = -18.295$ $\bar{C}_k' = -10.598$

TABLE 70/3 Paraxial Coefficients for the Case p = 0.56315

	1	2	3	4	final
y_a	1.0000	0.97253	1.9083	2.0359	2.0359
y_b	-0.56315	-0.42015	0.08425	0.12998	0.12998
v_a	0.00000	-0.14144	1.0589	1.0951	-1.0000
v_b	1.0000	0.73625	0.57077	0.39242	0.42733

TABLE 70/4 Primary, Secondary and Tertiary Aberration Coefficients

	1	2	3	4	final
A	-0.00788	6.8433	-0.00359	-65.132	-58.300
$\bar{A}$	-0.01469	-0.94342	0.01772	1.0854	0.14498
B	-0.02938	-1.88684	0.03545	2.1708	0.28996
$\bar{B}$	-0.05478	0.26012	-0.17498	-0.03617	-0.00582
C	-0.09811	0.56024	0.01725	-0.44827	0.03111
$\bar{C}$	-0.18295	-0.07724	-0.08517	0.00747	-0.33789
S_1	-0.00078	48.127	0.34997	-1516.2	-1467.8
$\bar{S}_1$	-0.00145	-6.2815	-1.1729	39.927	32.471
S_2	-0.00468	-19.348	-1.0053	145.68	125.32
$\bar{S}_2$	-0.00872	1.9491	-0.68447	6.7636	8.0195
S_3	-0.00797	2.0171	-0.57906	8.2931	9.7232
$\bar{S}_3$	-0.01486	-0.69439	0.23015	4.7600	4.2809
S_4	0.00124	-2.9137	0.28072	29.572	26.941
$\bar{S}_4$	0.00230	0.57286	0.29784	-0.81541	0.05759
S_5	0.01354	0.00905	0.07507	8.3720	8.4696
$\bar{S}_5$	0.02525	0.07068	0.07897	-0.23278	-0.05787
S_6	0.04541	-0.19888	-0.05511	0.19161	-0.01697
$\bar{S}_6$	0.08467	-0.00903	0.09193	0.02983	0.19740
T_1	-0.00009	397.70	-4.0367	-39,971	-39,578
$\bar{T}_1$	-0.00017	-52.312	-12.383	1084.8	1020.1
T_2	-0.00082	-240.35	-11.623	6582.3	6330.4
$\bar{T}_2$	-0.00152	27.904	1.8718	56.722	86.496
T_3	-0.00107	15.916	-3.5809	371.83	384.16
$\bar{T}_3$	-0.00200	-5.1303	1.5908	101.61	98.065
T_4	-0.00097	-16.122	6.4668	1170.8	1161.1
$\bar{T}_4$	-0.00181	3.5569	2.8067	-54.397	-48.035
T_5	-0.00074	-5.8887	0.89597	349.23	344.24
$\bar{T}_5$	-0.00138	1.8625	0.27801	-24.432	-22.293
T_6	0.00323	-0.08645	0.36218	-14.861	-14.582
$\bar{T}_6$	0.00603	0.15901	-0.04508	-3.2909	-3.1709
T_7	0.00200	10.097	-0.25541	-69.570	-59.726
$\bar{T}_7$	0.00372	-1.0639	0.24339	-2.9902	-3.8070
T_8	0.00608	3.2135	0.32296	-43.594	-40.052
$\bar{T}_8$	0.01134	-0.26964	-0.44252	-2.5904	-3.2912
T_9	-0.00146	-0.07914	-0.07938	-8.8972	-9.0572
$\bar{T}_9$	-0.00272	-0.05026	-0.10633	-0.43356	-0.59287
T_{10}	-0.02580	0.13849	0.05831	-0.49064	-0.31965
$\bar{T}_{10}$	-0.04811	0.016881	-0.07718	-0.02567	-0.13408

71. The Projection of the Hemisphere. Principal Rays.

(a) The principal uses of sky lenses seem to be for cloud surveys, aerial mapping and the tracking of rockets. As such the lens has to produce a mapping of the hemisphere, or at least a large part of it, onto some other surface which, in the case of the system under consideration, is a plane. According to Hill[1] the most satisfactory projection of the hemisphere onto a plane is either the stereographic projection, for which the ratio $\tan \frac{1}{2}\psi_1/\tan \psi_k'$ is constant, or the equidistant projection, for which the ratio $\psi_1/\tan \psi_k'$ is constant. ψ_i is the angle a ray makes with the optical axis before refraction at the ith surface. In view of the fact that the diaphragm is not in the image space of this system it seems pertinent to replace $\tan \psi_k'$ in the above ratios by H_k', the image height of the principal ray corresponding to inclination ψ_1 in the object space. Proper principal rays were traced through the system and some details obtained from these traces are presented in Table 71/1. Values of V_k' (= $\tan \psi_k'$) and H_k' were plotted against ψ_1 (see fig.3). The plot of V_k' against ψ_1, shown as -----, departs to a greater extent from a linear relation than does the plot of H_k', shown as ——— . It is evident from the plot of H_k' that the projection of a hemisphere afforded by the system is equidistant up to 75° half-field and over the remainder of the half-field up to 92° it is a good approximation to an equidistant projection.

Also given in Table 71/1 are V_d, the direction tangent of the principal ray at the diaphragm; S_{Ay} and T_{Ay}, the axial SPC of the principal ray; and P, the abscisse, in the object space, of the point to which the principal ray appears to be aimed. For rays in the paraxial region P = p.

---

[1] Hill (1924).

Rays for which $V_d > 1.184$ (§69) are vignetted by the second surface unless its active surface is more than a hemisphere, in which case the limiting ray is tangent to the first surface with a field angle of over 120^o. The SPC of principal rays are plotted in fig.4. For small T_y (cf. M(33.6))

$$S_y \propto T_y^{\ 3} \quad .$$

However, for large T_y the relationship becomes linear and is asymptotic to the line

$$S_y = 0.339T_y - 0.611 \quad , \qquad (71.1)$$

shown as – – – – in fig.4. The lower left hand quadrant of fig.4 corresponds to values of ψ_1 greater than 90^o. The variation in P and in S_y clearly illustrates the need for distinguishing between proper and improper principal rays (§62). The difference between the curves of H_k' and V_k' against ψ_1 (fig.3) is due to the fact that in the image space of Ⓚ the point of intersection of principal rays with the axis also varies slightly with ψ_1.

(b) From the configuration of the system (fig.1), it is clear that improper principal rays of a sufficiently large field angle will miss the second surface altogether. In fact, neglecting vignetting, total internal reflection limits the half-field of improper principal rays to a little over 62^o. However, in this case vignetting by the diaphragm sets in at about 45^o. Three proper principal rays are shown in fig.1 as the rays (a), (b) and (c). The improper principal ray (a'), corresponding to (a), is vignetted by the diaphragm and (b'), corresponding to (b), suffers total internal reflection at the fourth surface. The improper principal ray corresponding to (c) is not shown but is totally internally reflected at the second surface. Data relating to these rays and obtained from ray traces is given in Tables 71/2,3. The various columns are respectively: the surface indicator, the inclination of the ray to the axis, the inclination to the axis of the normal to the surface at the point of incidence of the ray, the angles of incidence and refraction and the coordinates of the point of intersection of the ray with the surface.

TABLE 71/1 Details Relating to Proper Principal Rays

V_d	ψ_1	T_{Ay}	S_{Ay}	P	V_k'	H_k'
0.1	10.02	0.17661	0.00044	0.56064	0.07467	0.17478
0.2	19.90	0.36200	0.00365	0.55306	0.14818	0.34710
0.3	29.53	0.56665	0.01298	0.54023	0.21945	0.51466
0.4	38.83	0.80467	0.03319	0.52190	0.28754	0.67546
0.5	47.60	1.0989	0.07192	0.49772	0.35171	0.82794
0.6	56.12	1.4893	0.14274	0.46730	0.41146	0.97097
0.7	64.08	2.0570	0.27340	0.43024	0.46650	1.1039
0.8	71.57	2.9991	0.53105	0.38608	0.51674	1.2265
0.86	75.83	3.9620	0.82082	0.35598	0.54460	1.2952
0.9	78.60	4.9582	1.1345	0.33434	0.56225	1.3389
0.95	81.95	7.0758	1.8234	0.30546	0.58329	1.3914
1.0	85.22	11.943	3.4478	0.27446	0.60323	1.4415
1.1	91.45	-39.547	-14.136	0.20570	0.63998	1.5349
1.2	97.35	-7.7499	-3.3803	0.12698	0.67283	1.6196

TABLE 71/2 The Trace of Some Principal Rays

Ray (a):-

$S_{Ay} = 0.0719$, $T_{Ay} = 1.0958$

i	ψ_i	θ_i	I_i	I_i'	X_{Ai}	y_{Ai}
1	47.7	11.7	36.0	22.7	0.05031	-0.49166
2	34.4	48.7	-14.4	-22.2	0.13595	-0.30007
3	26.6	2.6	24.0	15.5	-0.00163	0.07300
4	18.0	15.5	2.5	3.9	-0.01455	0.10676

$\psi_{Ak}' = 19.4$, $H_{Ak}' = 0.82793$, $Y_d < 10^{-5}$

Ray (b):-

$S_{Ay} = 0.2734$, $T_{Ay} = 2.0570$

i	ψ_i	θ_i	I_i	I_i'	X_{Ai}	y_{Ai}
1	64.1	16.3	47.7	29.1	0.09811	-0.68319
2	45.4	63.9	-18.5	-29.0	0.22384	-0.35858
3	35.0	3.5	31.5	20.0	-0.00313	0.10115
4	23.6	20.8	2.7	4.2	-0.02611	0.14197

$\psi_{Ak}' = 25.0$, $H_{Ak}' = 1.1039$, $Y_d < 10^{-5}$

Ray (c):-

$S_{Ay} = 1.1345$, $T_{By} = 4.9582$

i	ψ_i	θ_i	I_i	I_i'	X_{Ai}	y_{Ai}
1	78.6	20.9	57.7	33.7	0.15969	-0.86592
2	54.6	76.4	-21.8	-34.4	0.30510	-0.38790
3	42.0	4.5	37.5	23.6	-0.00503	0.12835
4	28.0	25.6	2.5	3.8	-0.03910	0.17229

$\psi_{Ak}' = 29.3$, $H_{Ak}' = 1.3389$, $Y_d = 2.8210^{-5}$

TABLE 71/3 The Trace of Some Improper Principal Rays

Ray (a'):-

$S_{Ay} = 0.0$, $T_{Ay} = 1.0989$

i	ψ_i	θ_i	I_i	I_i'	X_{Ai}	y_{Ai}
1	47.7	13.1	34.6	21.9	0.06303	-0.54958
2	35.0	56.5	-21.5	-33.9	0.17869	-0.33273
3	22.6	-1.4	24.0	15.5	-0.00049	-0.04010
4	14.1	-1.6	15.6	24.2	-0.00015	-0.01082

$\psi_{Ak}' = 22.6$, $H_{Ak}' = 0.83874$, $Y_d = -0.10120$

Ray is vignetted by the diaphragm.

Ray (b'):-

$S_{Ay} = 0.0$, $T_{Ay} = 2.0570$

i	ψ_i	θ_i	I_i	I_i'	X_{Ai}	y_{Ai}
1	64.1	20.4	43.7	27.0	0.15204	-0.84565
2	47.4	85.8	-38.5	-71.3	0.37003	-0.39807
3	14.5	-9.5	24.0	15.3	-0.02257	-0.27106

$Y_d = -0.30339$

Ray is vignetted by the diaphragm and totally internally reflected at the fourth surface.

Ray (c'):-

$S_{Ay} = 0.0$, $T_{Ay} = 4.9582$

i	ψ_i	θ_i	I_i	I_i'	X_{Ai}	y_{Ai}
1	78.6	29.8	48.8	29.6	0.32012	-1.2050

Ray is totally internally reflected at the second surface.

72. The Basal G-coefficients of the Sky Lens

(a) Three base-rays were chosen for illustrating the basal coefficients of the system. They are designated (a), (b) and (c) and their axial SPC S_{By} and T_{By} ($O\bar{T}$ coordinates) are given in Table 72/1.

TABLE 72/1 Axial SPC of the Base-Rays, Augmenting Factors and Astigmatic Distances

	S_{By}	T_{By}	μ_y	μ_z	x_M	x_S
(a)	0.0719	1.0989	0.68059	0.88034	0.02352	0.03469
(b)	0.2734	2.0570	0.45286	0.78876	0.04303	0.07786
(c)	1.1345	4.9582	0.21469	0.69276	0.05103	0.13529

These are in fact the three proper principal rays shown in fig.1 as (a), (b) and (c) respectively, details of which appear in Table 71/2. These base-rays were chosen so as, together with the axial coefficients, to cover a wide range of field angles, in this case up to about 80° half-field. Moreover, base-rays (a) and (c) respectively lie in regions of the field where the dominant aberrations are spherical aberration on the one hand and coma and astigmatism on the other (see §74a,80c). For rays with half-field angles approaching 90°, large changes in T_{Ay} correspond to quite small changes in the field angle. This is a handicap to economically covering the region of the half-field between 80° and 90°. However, it is not expected that systems primarily designed to cover such large fields would exhibit any sudden changes in behaviour over a small range of field angles. In the case of the system under consideration, it is felt that an additional base-ray at a half-field angle of over 90° would have been desirable, making it possible to investigate regions of the field corresponding to half-fields beyond 90°.

In the case of the base-ray (c) the maximum values of the angles of incidence I_i and of the inclination to the axis ψ_i are 57.7° and 78.6° respectively (Table 71/2). It is known[1] that if these angles are large then predictions based on the axial coefficients cannot be relied upon. However, in the case of the basal theory, the base-ray, which itself may have large values for these angles, is traced through the system. Neighbouring rays will have I, $\underset{\sim}{V}_A$, etc., differing by some amount from the corresponding values for the base-ray. It is reasonable to expect that it is these differences that will primarily govern the quality of the predictions based on the basal coefficients, and not the angles themselves. The point is this: although the angles of incidence, etc., may become large for some rays, judicious choice of a base-ray (such that for it these very angles are large) should enable reliable predictions to be made.

Table 72/1 also gives the positions (x_M and x_S) of the meridional and sagittal foci F_M and F_S and the two augmenting factors $\underset{\sim}{\mu}$. The positions of F_M and F_S are shown in fig.1. Note that they always lie behind the ideal image plane. However, it does not follow that the surface of best overall focus also lies behind the ideal image plane; for instance, this would not be the case if longitudinal spherical aberration was large and negative.

(b) Appendix G consists of tables of the surface contributions to the final $\underset{\sim}{G}_a$-, $\underset{\sim}{G}_b$- and $\underset{\sim}{h}$-coefficients for base-ray (a). In each case

---

1 Ford (1962) §14.

the contribution by the $\underline{i}$th surface is given in the $\underline{i}$th column. In these tables the coefficients $s_{gy1},\ldots,t_{gz20}$ are written sgy 1,...,tgz20. The SPC used were $O\bar{T}$ coordinates which, since the system is symmetric, are obtained from (55.6,7) provided the values of σ_A, τ_A, $\bar{\sigma}_A$ and $\bar{\tau}_A$ appropriate to axial $O\bar{T}$-coordinates are substituted into (55.7).

Consider first the $\underset{\sim}{G}_a$-coefficients and bear in mind the observations of §70 on the axial coefficients and their surface contributions. It is noted that there is no evidence of any balance between the contributions of the various surfaces. Certainly, if the system had exhibited a strong degree of balance between the primary coefficients and their surface contributions, it would be expected that this would carry over in some degree to the third order basal coefficients. In fact, it is expected that the coefficients of odd orders would almost certainly reflect the balance of the corresponding basal coefficients and that this would diminish in the outer regions of the field. It is also expected that to a lesser extent the coefficients of even orders would reflect the balance of the axial coefficients of the corresponding aberration-type.

In contrast to the axial and third order basal coefficients, the second order basal coefficients are not dominated by the fourth surface. In fact they are always of the same sign as the contribution by the second surface and opposite in sign to that of the fourth surface. With a few exceptions the third order coefficients are governed by the fourth surface, the contributions from which are large but balanced to a small extent by contributions of the opposite sign from the second surface. The exceptions are the distortion coefficients for which no obvious trends exist. The second order $\underset{\sim}{G}_b$-coefficients also exhibit no obvious trends and in some cases are even larger than the corresponding $\underset{\sim}{G}_a$-coefficients. The third order $\underset{\sim}{G}_b$-coefficients are generally an order of magnitude smaller than the corresponding $\underset{\sim}{G}_a$-coefficients and exhibit the same general characteristics as the $\underset{\sim}{G}_a$-coefficients. The final $\underset{\sim}{G}_a$- and $\underset{\sim}{G}_b$-coefficients given in Appendix G satisfy each of the identities given in Tables 47/1,2 and 59/2.

Inspection of the coefficients for base-rays (b) and (c) reveals the following trends. The second order coefficients increase slightly and then decrease with increasing field angle. The third order coefficients decrease with the exception of those for which $\mu-\nu$ is 2 or 3 (i.e., those coefficients which multiply terms depending on S_z through its second or third power). All $\underset{\sim}{G}_b$-coefficients become opposite in sign to the corresponding $\underset{\sim}{G}_a$-coefficients.

73. The $\underset{\sim}{h}$-coefficients of the Sky-Lens

(a) The final $\underset{\sim}{h}$- and $\underset{\sim}{v}$-coefficients for the three base-rays (a), (b) and (c) are given in Tables 73/1,2. The outstanding features of the $\underset{\sim}{h}$-coefficients are (i) the smallness of the coefficients of linear astigmatism, (ii) the increase in magnitude of the second order coefficients and decrease in magnitude of the third order coefficients as T_{By} is increased, (iii) the relatively great magnitude of the coefficients of third order third degree astigmatism and, in particular, the fact that the coefficients of terms depending on S_z through the powers S_z^2 and S_z^3 do not share the general decrease of the $\underset{\sim}{h}$-coefficients as observed in (ii), and (iv) the apparent independence of the ratios of corresponding $\underset{\sim}{h}$- and $\underset{\sim}{v}$-coefficients on the coefficients concerned (with a few exceptions) and on the base-ray.

TABLE 73/1 h-coefficients

Base-ray	(a)	(b)	(c)
h_{ya}	0.01907	0.02618	0.01654
h_{yb}	0.44062	0.17476	0.02999
s_{hy1}	-0.16944	-0.35862	-0.20742
s_{hy2}	0.07023	0.18020	0.12952
s_{hy5}	-0.23671	-0.09737	-0.02728
s_{hy8}	-0.12302	-0.40752	-0.77995
s_{hy9}	0.00398	0.08811	0.33857
s_{hy10}	-0.14149	-0.09163	-0.06046
t_{hy1}	-28.867	-11.364	-1.5561
t_{hy2}	13.022	8.5375	1.5411
t_{hy5}	-1.9427	-2.1205	-0.50630
t_{hy8}	-43.578	-29.532	-13.462
t_{hy9}	5.7885	8.1195	6.4847
t_{hy10}	-0.15929	-0.52701	-0.77177
t_{hy11}	0.17895	0.20188	0.05658
t_{hy14}	6.4377	7.1770	4.23765
t_{hy15}	-0.84897	-1.9324	-2.0124
t_{hy16}	0.07194	0.17193	0.24485

Base-ray	(a)	(b)	(c)
h_{za}	0.03238	0.06777	0.10752
h_{zb}	0.75130	0.52765	0.24544
s_{hz3}	-0.18699	-0.45557	-0.46174
s_{hz4}	0.00012	0.04033	0.08730
s_{hz6}	0.07340	0.14258	0.15369
s_{hz7}	-0.28752	-0.19050	-0.07862
t_{hz3}	-33.552	-16.778	-4.0838
t_{hz4}	2.2114	2.2771	0.98573
t_{hz6}	9.9823	8.2929	2.6482
t_{hz7}	-0.65616	-1.1035	-0.62842
t_{hz12}	-0.73005	-1.0191	-0.42872
t_{hz13}	0.09400	0.18072	0.10845
t_{hz17}	-50.703	-43.783	-35.773
t_{hz18}	9.6707	17.038	24.305
t_{hz19}	-0.59750	-2.1987	-5.5030
t_{hz20}	-0.11710	0.05169	0.41037

TABLE 73/2 $\underset{\sim}{y}$-coefficients

Base-ray	(a)	(b)	(c)
v_a	-0.81070	-0.60846	-0.32415
v_b	0.31184	0.22757	0.11775
s_{v1}	0.26190	0.08125	-0.01598
s_{v2}	0.16218	0.13144	0.06383
s_{v5}	-0.06870	-0.05043	-0.02038
s_{v8}	0.15955	0.10807	-0.00899
s_{v9}	0.09059	0.09859	0.09902
s_{v10}	-0.03448	-0.03448	-0.02923
t_{v1}	-12.766	-5.0756	-0.69942
t_{v2}	5.7141	3.7059	0.66781
t_{v5}	-0.90469	-0.94565	-0.22109
t_{v8}	-19.335	-13.362	-6.2487
t_{v9}	2.5794	3.5258	2.8757
t_{v10}	-0.09805	-0.25726	-0.34428
t_{v11}	0.05346	0.08476	0.02542
t_{v14}	2.9799	3.2951	1.9730
t_{v15}	-0.37691	-0.88430	-0.91868
t_{v16}	0.01158	0.06621	0.11149

Base-ray	(a)	(b)	(c)
w_a	-0.93320	-0.87036	-0.79475
w_b	0.38111	0.34246	0.29524
s_{w3}	0.15063	0.00034	-0.09380
s_{w4}	0.10161	0.12727	0.11637
s_{w6}	0.05725	0.05009	0.04292
s_{w7}	-0.06678	-0.06251	-0.04562
t_{w3}	-14.578	-7.2570	-1.7460
t_{w4}	1.0551	1.0039	0.41521
t_{w6}	4.3427	3.5390	1.1143
t_{w7}	-0.28113	-0.49270	-0.27421
t_{w12}	-0.34518	-0.44538	-0.17975
t_{w13}	0.02085	0.06507	0.04622
t_{w17}	-22.085	-19.133	-15.683
t_{w18}	4.6175	7.6819	10.744
t_{w19}	-0.18863	-0.95395	-2.4325
t_{w20}	-0.02660	0.02655	0.18050

(b) The smallness of the coefficients of linear astigmatism is of course due to the fact that imagery is almost stigmatic in the chosen image plane (in this case the paraxial ideal image plane),

$$h_{ya} = y_a + \ell v_a = v_a(\ell - \ell_M) = -v_a x_M \quad ,$$

and since v_a is of the order of -1, h_{ya} will be of the order of the displacement of the meridional focus from the image plane. If parabasal imagery is nearly stigmatic, this is small. The analysis of h_{za} proceeds in an analogous manner. The second order coefficients vanish if the system is symmetric, hence their presence indicates a departure from rotational symmetry.[1] For this reason it could be expected that the second order terms would increase in magnitude as the base-ray is taken more and more into the outer regions of the field. It has been seen (§60) that the third order coefficients are formally the sum of the corresponding axial coefficients (according to Table 60/1) and terms which vanish when the base-ray is along the optical axis. In view of the fact that the third order basal coefficients decrease in magnitude as the field angle is increased, it is probable that these terms are of the opposite sign to the corresponding axial coefficients and increase in magnitude as the field angle is increased.

---

1 Strictly it is sufficient for a system to have two planes of symmetry and for the base-ray to lie along the common section of these two planes in order that the second order coefficients vanish. The presence of second order coefficients therefore indicates either that the base-ray does not lie along this section or that the system is not doubly symmetric. Since a doubly symmetric system of spherical surfaces is symmetric, this distinction need be of no concern at present.

Be that as it may, the relative change in the magnitudes of the coefficients naturally has some effect on the general nature of the aberrations at any particular field angle. It will be seen, for instance in §74, that the aberration is predominantly spherical for half-field angles less than about 55^{o}, whereupon a gradual transition to coma and astigmatism takes place. It might be expected that in the outer regions of the field the large values of the coefficients t_{hz17} and t_{hz18} in H_z would produce an image patch that was greatly extended in the direction normal to Ⓜ. This appearance of a "meridional focal line" would be somewhat offset by the coefficient t_{hy8} in H_y. However, the entrance pupil is highly elliptical for large field angles. The permissible range of values for S_y is roughly three times that for S_z and this should greatly reduce any appearance of "meridional astigmatism". (See also §80.)

(c) It was noticed that the ratio $r = \underset{\sim}{h}_{\mu\nu\tau}^{(3)}/\underset{\sim}{v}_{\mu\nu\tau}^{(3)}$ was virtually independent of μ, ν and τ and of the base-ray. Moreover, there was no noticeable distinction between the values of r obtained from h_y- and v-coefficients on the one hand, or h_z- and w-coefficients on the other. The ratio was usually between 2.25 and 2.35. Consider for the moment the quantity $\bar{\ell} = H_{yk}'/V_k'$ for a meridional ray Ⓡ. This, it is readily seen, is the distance from the image plane of the point of intersection of Ⓡ and $Ⓡ_B$. Furthermore, the base-ray will intersect the optical axis at a point whose distance from the image plane is $\bar{\ell}_B$, where

$$\bar{\ell}_B = H_{Byk}'/V_{Bk}' \quad .$$

From Table 71/1 it is found that the values of $\bar{\ell}_B$ for the three base-rays (a), (b) and (c) are 2.354, 2.366 and 2.381 respectively.[2] The constancy of both $\bar{\ell}_B$ and the ratio r and the fact that their respective values are quite close strongly suggests that they are somehow or other related. However, this cannot be the case. In the context of the axial theory, and for axial rays, $\bar{\ell} = \bar{\ell}_B$. The ratios r were investigated for primary and secondary spherical aberration and found to be 2.31 and 2.37 respectively. However, for axial rays, $\bar{\ell}$ is longitudinal spherical aberration, which is certainly nowhere near 2.3 .

Consider the ratio r for the third order coefficients of H_y and V. From (24.11)

$$r = t_{hy\alpha}/t_{v\alpha} = T_{gy\alpha a}/\mu_{ya}({}^{\backprime}T_{gy\alpha}|v) - x_M$$

$$= -(y|v)T_{gy\alpha a}/v_a(T_{gy\alpha a}v_b - T_{gy\alpha b}v_a) - x_M \quad .$$

In order that this be independent of the third order coefficients, the denominator of the first term must be proportional to $T_{gy\alpha a}$. This will be the case if the coefficient $T_{gy\alpha b}$ is negligible. Then

$$r = \frac{y_b}{v_b} - \frac{y_a}{v_a} - x_M = \ell + \frac{y_b}{v_b} \quad , \tag{73.1}$$

by (22.2). Note that the analysis above holds for all orders of coefficients. This is also exactly the ratio of the first order coefficients h_{yb} and v_b and is the distance from the image plane of the meridional focus of the basal point of the entrance pupil. Similarly, the ratio r has the desired properties for the coefficients of H_z and W provided the G_{zb}-coefficients are small. It was checked numerically that those coefficients

---

2 For paraxial rays $\bar{\ell}_B = 2.34$. The small variation in $\bar{\ell}_B$ is characteristic of the fact that the aberrations associated with the diaphragm and the exit pupil are small.

whose b-components were <u>not</u> small did <u>not</u> produce a constant value of r. Further, the isolated coefficients for which the b-components were very small gave as the value for r that obtained from (73.1). This affords a partial explanation of the constancy of the ratio r.

(d) If it is accepted that the ratio r is known, some estimate may be made of the effect of the out-of-focus terms $-x_M V$ and $-x_S W$. From (24.11) it follows that

$$t_{hy\alpha} = rt_{v\alpha} = T_{gy\alpha a}/\mu_y - x_M t_{v\alpha} \quad .$$

Thus

$$t_{v\alpha} = T_{gy\alpha a}/\mu_y(r+x_M)$$

and hence

$$t_{hy\alpha} = T_{gy\alpha a}[1-x_M/(r+x_M)]/\mu_y \quad . \tag{73.2}$$

If $x_M \ll r$ the difference between this and the value $T_{gy\alpha a}/\mu_y$ which $t_{hy\alpha}$ would have if parabasal imagery was stigmatic is truly negligible. A similar analysis can be carried out for the coefficients of H_z. It should be obvious that (73.2) can be immediately generalized to coefficients of any order.

74. <u>Preliminary Comments on the Aberrations and Their Determination</u>

(a) An experienced optical designer can make a good assessment of the aberrations of a system from the relative and absolute magnitudes of the aberration coefficients, without resorting to the production of spot diagrams and the like. Lacking such experience, a preliminary analysis of the (meridional) aberrations of the system will be attempted on the following premises. Firstly, let ϵ_S and ϵ_C be the sum of all astigmatic and all comatic aberrations respectively for a member of some pencil of rays. Then the total aberration is

$$\underset{\sim}{\epsilon} = \underset{\sim}{\epsilon}_S + \underset{\sim}{\epsilon}_C \quad .$$

If $|\underset{\sim}{\epsilon}_S| \gg |\underset{\sim}{\epsilon}_C|$ throughout all but possibly a small region of the aperture, the image will be characteristic of an astigmatic aberration type. If $|\underset{\sim}{\epsilon}_S| \gg |\underset{\sim}{\epsilon}_C|$ under similar conditions, the image will be highly comatic. Secondly, if the meridional and sagittal asymmetries are small by comparison with the total aberration, asymmetry will be of no importance even if the image itself is substantially comatic.

Only the meridional fans of pencils about each base-ray will be considered. Then $S_z = T_y = T_z = 0$ and

$$\epsilon_{Cy} = s_{hyl}S_y^2 \quad , \quad \epsilon_{Sy} = h_{ya}S_y + t_{hyl}S_y^3 \quad .$$

By considering the relative magnitude of $\underset{\sim}{\epsilon}_S$ and $\underset{\sim}{\epsilon}_C$ for the meridional fan about base-ray (a) it is found that the aberration is astigmatic except for a narrow range of values of S_y about 0.026 where the comatic aberration is of the order of 10^{-4}. Due to vignetting (§77) the range of permissible values of S_y is $-0.06 < S_y < 0.06$. The maximum asymmetry is 6×10^{-4} , about a tenth of the aberration at full aperture. For base-ray (b) the range of S_y is $-0.08 < S_y < 0.08$ and for a meridional fan about the base-ray, astigmatic aberrations dominate if $|S_y| < 0.03$ and in the extreme outer regions of the aperture. When $\epsilon_{Sy} = 0$, the comatic aberration is 8×10^{-4} . At full aperture the asymmetry is 2.5×10^{-3} or about half the aberration. For a quarter of the full aperture, the asymmetry is about 1.6×10^{-4} or about an eighth of the aberration. Finally, for base-ray (c), $-0.15 < S_y < 0.15$ and the aberration is comatic for $|S_y| > 0.08$ and astigmatic for $|S_y| < 0.03$. The last region represents only a small fraction of the incident pencil.

When $\epsilon_{Sy} = 0$, the comatic aberration is of the order of 2×10^{-3}. The asymmetry is large throughout the aperture.

In summary, the aberrations of the meridional fan about base-ray (a) are astigmatic, the fan about the base-ray (b) is astigmatic except for the outer regions of the aperture and the fan about base-ray (c) is highly comatic. These considerations can be extended to pencils inclined to the base-ray. It is then possible to find values of $\underset{\sim}{T}$ such that the aberration undergoes a definite transformation in its nature - for example, from an astigmatic to a comatic aberration. In the present case the values of $\underset{\sim}{T}$ found such that this occurs are so large that it is most unlikely that they are in the region of convergence of the basal series for $\underset{\sim}{H}$.

(b) When producing spot diagrams it is usual to select from a pencil of rays only those members which pass through the grid points of a grid placed in the plane of the entrance pupil. However, it has already been seen that it is the cross-sectional area of the pencil which is important in determining the illumination of the image. It therefore seems appropriate to place the grid normal to the principal ray of the pencil. Suppose the SPC of the principal ray have been determined by some means (for instance, from the basal coefficients) and that its value of $\underset{\sim}{S}_A$ is $\underset{\sim}{S}_P$. Further, let its coordinate $\underset{\sim}{V}_{A1}$ be $\underset{\sim}{V}_P$. Let the grid be placed normal to the principal ray and through the point of intersection of the latter with Ⓔ (i.e., through the centre of the entrance pupil for the pencil). With this point as origin, set up in the grid a coordinate system whose y-axis lies in Ⓜ. If the grid mesh is δ, only those rays which intersect the grid in the points $(n\delta, m\delta)$, where n and m are integers, will be selected when constructing the spot diagram. The axial SPC $\underset{\sim}{S}_A$ of these rays can be determined from

$$S_{Ay} = S_{Py} + n\delta(1+\underset{\sim}{V}_P\cdot\underset{\sim}{V}_P)^{\frac{1}{2}} \quad , \quad S_{Az} = S_{Pz} + m\delta \quad . \qquad (74.1)$$

If the pencil is <u>meridional</u> - its principal ray is a meridional ray - (74.1) is exact. If the pencil is not meridional, (74.1) is a good approximation since, in the region of convergence of the basal series for $\underset{\sim}{H}$, W_P will be small compared to V_P, especially if the field angle of the base-ray is large.

When the aberrations of the members of a pencil were determined the following procedure was used. Initially the pencil was specified by the $\underset{\sim}{T}_A$ of the object. Using the basal coefficients relative to a suitable base-ray, the axial SPC of the principal ray of the pencil were computed (see §75). The computer programmes employed for the determination of the aberrations of a pencil of rays were designed to accept the axial SPC of the principal ray of the pencil and use (74.1) to determine the axial SPC of the required rays of the pencil. If the aberrations were to be predicted, the axial SPC of the rays were converted to the corresponding basal SPC and the image heights of the rays determined from the basal series for $\underset{\sim}{H}$. If the rays were to be traced, the ray trace scheme of Ford[1] was used to determine the image height of the rays. In both cases, subtraction of the image height of the principal ray gave the required aberrations. If a meridional (or sagittal) fan was required instead of a full pencil, only those rays for which m (or n) was zero, were considered. All spot diagrams appearing in this thesis were constructed with $\delta = 5 \times 10^{-3}$.

---

1 Ford (1960).

The computer programmes were coded in Fortran and run initially on an IBM 1620 computer and later on an IBM 360/50. The computations employed in this thesis were performed by the 360. The coordinates for the spot diagrams were punched onto cards and plotted by an IBM 1620 plotter on-line to the 1620 computer at Mt. Stromlo Observatory.

(c) Note that although the aberrations of a pencil of rays are referred to the <u>same</u> ray, irrespective of whether the aberrations are traced or predicted, the point in the image plane which is the origin for the aberrations may be different in the two cases. This is because in general the predicted and traced aberrations of the principal ray do not agree. The differing origins do not matter when considering traced and predicted spot diagrams. However, for the meridional fans (§76) and the annular curves (§78) this corresponds to a translation of the predicted curve with respect to the traced curve. This effect is usually quite small and does not substantially alter the quality of the predictions.

XII. APPLICATIONS OF THE BASAL THEORY AND A COMPARISON OF ITS PREDICTIONS WITH THOSE OF RAY TRACING

75. Analysis of Principal Rays Using the Basal Coefficients

(a) It has already been mentioned that for certain classes of systems (of which the one at hand is a member) it is desirable to refer the aberrations of the members of a pencil of rays to the principal ray of the pencil. (§7,56,66). When this is done in the basal theory it is obvious that the basal coefficients will have to be used to obtain the coordinates and image heights of principal rays. It is pertinent to investigate the predictions of these quantities.

The equations required to obtain the coordinates of principal rays were developed in §64a,65. Since the base-rays are principal rays, the coefficients in (65.2) are obtained from Table 64/1. In order to compute these it is first necessary to compute the diaphragm coefficients from (63.3). The second order entrance pupil coefficients as well as the coefficients t_{sy11}, t_{sy16}, t_{sz13} and t_{sz20} for each of the three base-rays are presented in Table 75/1. The values of S_{Ay} for principal rays corresponding to various values of T_{Ay} were computed for each of the three base-rays (a), (b) and (c) by using (65.2) and the values of the coefficients given in Table 75/1. The results of these computations are presented in Table 75/2 for direct comparison with the corresponding entries in Table 71/1. Fig.5 graphically represents the predictions of the two base-rays (a) and (b) and of the axial theory. The predictions of the axial theory were obtained by means of M(33.6). The required axial b-coefficients are

$$C_b = 0.0769 \quad , \quad \bar{C}_b = 0.0815 \quad , \quad \bar{S}_{6b} = -0.0599 \quad ,$$

so that, for principal rays in the neighbourhood of the axis,

$$S_{Ay} = 0.0815\ T_{Ay}^{\ 3} - 0.0371\ T_{Ay}^{\ 5} + 0(7) \quad . \qquad (75.1)$$

From fig.5 it is evident that the axial predictions break down at about 35° to 40° half-field and that the base-rays (a) and (b) give quite good predictions over half-field angles between 25° to 65° and 55° to 72° respectively. The predictions of base-ray (c) are not shown, however it is to be noted that the second and higher order coefficients in (65.2) are small for this base-ray and that the coefficient p_{sy2} is quite close to the asymptotic value of 0.339 (see §71, in particular (71.1)).[1]

As a further test of the quality of the predictions of S_{Ay} the points of intersection with the diaphragm were determined by ray tracing for sets of rays which had been <u>predicted</u> to be principal. If $\rho_{tr} = (\underset{\sim}{Y}_{Ad}\cdot\underset{\sim}{Y}_{Ad})^{\frac{1}{2}}$, the smaller the value of ρ_{tr}/ρ_d the better is the prediction of S_{Ay}. The values of this ratio are tabulated in Table 77/1 for various principal rays and the smallness of the values is indicative of the generally satisfactory nature of the predictions.

(b) The image height of a principal ray is given by the series (66.1). The $\underset{\sim}{h}_p$-coefficients appear in Table 66/1 and may be computed for each of the base-rays (a), (b) and (c) by using the appropriate $\underset{\sim}{h}$- and entrance pupil coefficients from Tables 73/1 and 75/1. The results of these computations are shown in Table 75/3. The image heights of the various principal rays of Table 75/2 were computed using (66.1) and the appropriate base-rays and appear as the last column of Table 75/2. These predicted values may be compared with the corresponding traced values in Table 71/1. Alternatively, fig.6 compares graphically the predicted

---

[1] For a base-ray with $T_{By} = 12$, $p_{sy2} = 0.337$ and the other coefficients are of the order of 0.0001 to 0.00001.

TABLE 75/1 Entrance Pupil Coefficients

Base-Ray	(a)	(b)	(c)
p_{sy1}	0.70392	0.93857	1.74471
p_{sy2}	0.15517	0.25124	0.31949
s_{sy1}	-0.10407	-0.22408	-0.63091
s_{sy2}	0.20810	0.26068	0.28807
s_{sy5}	0.07487	0.03078	0.00380
s_{sy8}	0.05671	0.10077	0.21044
s_{sy9}	0.09032	0.14387	0.21424
s_{sy10}	0.04083	0.02876	0.00914
t_{sy11}	-0.01939	-0.00961	-0.00069
t_{sy16}	-0.00619	-0.01300	-0.00292
p_{sz3}	0.60466	0.63763	0.68243
p_{sz4}	0.06543	0.13291	0.22881
s_{sz3}	-0.12217	-0.18025	-0.24542
s_{sz4}	0.09033	0.14387	0.21424
s_{sz6}	0.04092	0.02750	0.00837
s_{sz7}	0.08167	0.05753	0.01829
t_{sz13}	-0.00619	-0.01300	-0.00292
t_{sz20}	0.03716	0.01398	0.00184

TABLE 75/2 Predicted Values of S_{Ay} and H_{yk}' for Principal Rays

Base-ray	T_{Ay}	S_{Ay}	H_{yk}'
	0.1766	0.0072	0.1586
	0.3620	0.0057	0.3435
	0.5660	0.0132	0.5146
	0.8047	0.0300	0.6756
S_{By} = 0.0719, T_{By} = 1.0989	1.0989	0.0719	0.8279
	1.4893	0.1428	0.9711
	2.0570	0.2732	1.1171
	2.9991	0.5049	1.4228
	0.8047	0.0251	0.7068
	1.0989	0.0659	0.8382
	1.4893	0.1384	0.9722
S_{By} = 0.2734, T_{By} = 2.0570	2.0570	0.2734	1.1039
	2.9991	0.5253	1.2322
	3.9620	0.7941	1.3721
	4.9582	1.0272	1.6850
	2.0570	0.2216	1.1435
	2.9991	0.5173	1.2325
	3.9620	0.8203	1.2956
S_{By} = 1.1345, T_{By} = 4.9582	4.9582	1.1345	1.3389
	7.0758	1.8221	1.3955
	11.943	3.3196	1.7107

TABLE 75/3 $\underset{\sim}{h}_P$- and $\underset{\sim}{v}_P$-coefficients

Base-ray	(a)	(b)	(c)
h_{Pyb}	0.44358	0.18134	0.03527
h_{Pzb}	0.75342	0.53665	0.27004
s_{Phy5}	-0.22847	-0.07393	-0.00701
s_{Phy10}	-0.14305	-0.09447	-0.03175
s_{Phz7}	-0.28195	-0.17273	-0.04735
t_{Phy11}	0.08414	0.02756	0.00137
t_{Phy16}	0.04867	0.04803	0.00814
t_{Phz13}	0.04867	0.04803	0.00814
t_{Phz20}	-0.12829	-0.04199	-0.00478
v_{Pb}	0.18576	0.07444	0.01409
w_{Pb}	0.31964	0.22619	0.11290
s_{Pv5}	-0.09780	-0.03092	-0.00283
s_{Pv10}	-0.04963	-0.02033	-0.00147
s_{Pw7}	-0.12182	-0.07376	-0.01993
t_{Pv11}	0.02816	0.00807	-0.00009
t_{Pv16}	0.01972	0.01530	0.00094
t_{Pw13}	0.02186	0.02083	0.00345
t_{Pw20}	-0.05545	-0.01793	-0.00201

image heights with the corresponding traced heights. In addition to the predictions made with base-rays (a) and (b) fig.6 shows the secondary axial predictions. If the tertiary terms are included in the axial predictions the general appearance of the curve is unaltered although the agreement is good up to about $T_{Ay} = 0.8$. The range of field angles over which the predictions are satisfactory is in all cases much the same as that for the predictions of the corresponding coordinates (fig.5), perhaps a little shorter.

(c) It is possible to predetermine what the $\underline{h}_p$-coefficients should be for any base-ray in order that the mapping of the hemisphere onto a plane be equidistant. For an equidistant mapping $H_{Ayk}' \propto \vartheta_{A1}$ where ϑ_{A1} is the radian measure of the angle ψ_1. For a system of unit focal length with the object at infinity paraxial optics demands that the constant of proportionality be unity. Thus, the equidistant mapping is

$$H_{Ayk}' = \vartheta_{A1} \quad . \tag{75.2}$$

Furthermore, if $O\bar{T}$ coordinates are used, $\vartheta_{A1} = \text{atan } T_{Ay}$. If atan T_{Ay} is expanded as a Taylor series in T_{Ay}, (75.2) gives

$$H_{Ayk}' = T_{Ay} - \tfrac{1}{3}T_{Ay}^{\ 3} + \frac{1}{5}\,T_{Ay}^{\ 5} - \frac{1}{7}\,T_{Ay}^{\ 7} + 0(9) \quad . \tag{75.3}$$

Thus, if distortion is the only aberration present, the coefficients of distortion must be

$$\bar{C}_{ak}' = -0.33333 \quad , \quad \bar{S}_{6ak}' = 0.20000 \quad , \quad \bar{T}_{10ak}' = -0.14286 \quad ,$$

which are in fair agreement with the actual coefficients in Table 70/4. However, since there are other aberrations present the location of the centre of the core (§79a) of the spot diagram corresponding to the pencil $\underline{T}_A$ must be given by (75.3). Granted that comatic aberrations are small, the centre of the core is the point of intersection of the principal ray with the image plane. If $\underset{\sim}{S}_A$ is eliminated from the series for $\underset{\sim}{H}_A$ by substituting the series M(33.6) in place of S_{Ay} (cf. §66) it is found that the image height of principal rays is given by

$$H_{Ayk}' = T_{Ay} - 0.3379\, T_{Ay}^{3} + 0.1994\, T_{Ay}^{5} - 0.1411\, T_{Ay}^{7} \quad , \qquad (75.4)$$

and the agreement with (75.3) is improved. It is evident that to a good approximation the projection of the hemisphere is equidistant in the neighbourhood of the axis of the system.

Now expand atan T_{Ay} about T_{By} . It is found that

$$H_{Ayk}' = \vartheta_{B1} + \alpha_{B1}^{2} T_y [1 - V_{B1} \alpha_{B1}^{2} T_y + \tfrac{1}{3} \alpha_{B1}^{4} (3V_{B1}^{2} - 1) T_y^{2} + 0(3)] \ . \qquad (75.5)$$

The values of the coefficients appearing in (75.5) were computed for each of the three base-rays (a), (b) and (c) and are given in Table 75/4. These values are to be compared with the corresponding values of the computed image height coefficients (Table 75/3). If the projection is exactly equidistant, $\vartheta_{B1} = H_{Byk}'$. From Tables 71/1 and 75/4 it is evident that this is not the case, although the differences are small. For base-ray (a) the difference is of aberration magnitude and for base-ray (c), 3% of the focal length. The actual first order coefficients are in all cases somewhat smaller than those in Table 75/4, indicating that in the outer regions of the field the constant of proportionality in (75.2) is somewhat less than unity. For the purpose for which the system was designed (cloud surveys) the mapping is no doubt equidistant to a sufficiently good approximation. For any purpose where precision is important, aerial mapping for instance, a knowledge of the basal coefficients would enable the necessary corrections to be made when the photographic plates are analysed.

TABLE 75/4 The h_{Py}-coefficients for the Equidistant Projection

Base-ray	(a)	(b)	(c)
ϑ_B	0.83249	1.1183	1.3717
h_{Phyb}	0.45298	0.19116	0.03909
s_{Phy5}	-0.22549	-0.07517	-0.00758
t_{Phy11}	0.08126	0.02723	0.00145

(d) It is of interest to note that the $\underset{\sim}{h}_P$-coefficients are almost equal to the corresponding $\underset{\sim}{h}$-coefficients, especially those of low order and for base-ray (a). In this case the principal ray of any pencil and the ray through the basal point of the entrance pupil have, to a close approximation, the same image height. This does not imply that the principal rays in the neighbourhood of the base-ray may be represented by rays through the basal point of the entrance pupil. This is so since the $\underset{\sim}{v}_P$-coefficients are in some instances quite different to the corresponding $\underset{\sim}{v}$-coefficients. The ratio of any $\underset{\sim}{h}$-coefficient and the corresponding $\underset{\sim}{v}$-coefficient (§73) is approximately independent of the base-ray and of the particular coefficient. However, this is not the case with the $\underset{\sim}{h}_P$- and $\underset{\sim}{v}_P$-coefficients.

76. Meridional and Sagittal Fans

(a) From the definitions of §3 it is evident that meridional and sagittal fans are two-parameter families of rays. For the meridional fan the parameters are S_y and T_y (or the axial coordinates S_{Ay} and T_{Ay}) and for sagittal fans the parameters are S_z and T_z. In each case the remaining two SPC are zero. The definition of sagittal fan will be generalised as follows: a fan of rays from a point object is said to be a <u>sagittal fan</u> if, in the object space of Ⓚ, each member of the fan lies in the plane normal to Ⓜ and containing the principal ray from the object. In this case S_y and T_y are not zero but are constant for the fan. It is often possible to obtain useful quantitative information about the aberrations of a system from a consideration of meridional and sagittal fans alone.

Fig.7,8 and 9 compare the traced aberrations of several meridional fans with the second and third order basal predictions obtained by using base-rays (a), (b) and (c) respectively. The scale along both axes is 0.01 focal lengths. ϵ_y is the aberration referred to the principal ray and ρ is the $n\delta$ of §74b, in particular (74.1). The following general features of the traced curves (and to a lesser extent of the third order predictions) may be noted: for low field angles the traced curves are virtually symmetric with respect to rotations through 180^o about the origin and there is a fairly wide range of values of ρ for which ϵ_y is quite small. These features are characteristic of the symmetric nature of the aberrations of pencils at low field angles and of the well defined core in the spot diagrams for these pencils (§79). As the field angle is increased this symmetry about the origin vanishes and there is no longer any extended region where ϵ_y is predominantly small. This is due to the increased effects of comatic aberrations.

These basic features are reproduced by the third order predictions for pencils about the base-ray or inclined slightly to the base-ray. However, it is a general feature that the asymmetry is overestimated. The actual aberration curves are clearly cubic or of higher order in S_y (or in ρ). It is therefore little wonder that the second order predictions are so poor, and gratifying that the agreement for the third order predictions is as good as it is. The inclusion of the fourth and higher orders will introduce generally small corrections to the basic cubic curve. If the value of T_{Ay} for the pencil is less than T_{By}, that portion of the predicted curve corresponding to positive values of ρ generally agrees better with the traced curve than does the part of the curve for negative ρ. The reverse is true if $T_{Ay} > T_{By}$. This may be explained as follows. For principal rays S_{Ay} is an increasing function of T_{Ay}. If, for a given pencil and base-ray, $T_{Ay} < T_{By}$, the coordinate of the principal ray of the fan is less than S_{By}. From (74.1) it is clear that the basal coordinates S_y corresponding to positive ρ will be smaller than the values corresponding to negative ρ. Consequently the basal predictions are likely to be more reliable for positive ρ than for negative ρ. The case $T_{Ay} > T_{By}$ submits to a similar explanation - for $\rho < 0$, S_y is smaller than for $\rho > 0$ and consequently the predictions will be more reliable for negative ρ.

(b) Various sagittal fans were also considered for each base-ray. Fig.10 gives the curves for ϵ_y against ϵ_z for the sagittal fans of half-field angles of 45° and 64° and fig.11 for sagittal fans in the neighbourhood of base-ray (c). (Note: The scale along the ϵ_y-axis for the 45° fan is one tenth of that along the ϵ_z-axis but for all other curves the two scales are the same.) The basal prediction for the 45° pencil was made with base-ray (a) and for the 64° pencil with base-ray (b). The predictions for fig.11 were made with base-ray (c). In the 45° and 64° fan T_y is quite small whereas for the fans of fig.11 it is at times large. Generally, the agreement is rather poor. This apparently arises from a failure of the basal series for ϵ_y to approximate the correct values of ϵ_y for sagittal fans. The predicted values of ϵ_z agree quite well with those obtained from ray traces. The predictions of ϵ_y for large field angles is much better than for low field angles and, due to increased comatic aberrations, the overall magnitude of ϵ_y is larger for large field angles. When considering the poor agreement of the predicted sagittal curves for low field angles the fact that ϵ_y is very small should be bornein mind.

Consider a sagittal fan such that $S_y = T_y = T_z = 0$. Then

$$\epsilon_y = s_{hy8} S_z^2 + O(4) \quad , \quad \epsilon_z = h_{za} S_z + t_{hz17} S_z^3 + O(4) \quad . \tag{76.1}$$

When ϵ_y is predicted to the third order there is only a single term in the series for ϵ_y. Similarly, when predicting ϵ_z to the <u>second</u> order there is again only a single term in the corresponding series. It has already been seen that the third order predictions of ϵ_y are in poor agreement with the traced values. In addition, the second order predictions of ϵ_z are also in poor agreement with the traced values of ϵ_z. This can be seen from fig.10 where the bars on the curves correspond to equal values of S_z. It is now seen that in each of these cases the respective series in S_z contains only a single term. If S_y and T_y are non-zero the situation is unchanged. ϵ_y can depend on S_z only through the term S_z^2, and ϵ_z (to the second order) must be linear in S_z. Evidently the dominant terms in the series for $\underset{\sim}{\epsilon}$ represent the aberration over only a very small neighbourhood of the base-ray.

77. Vignetting

(a) When the spot diagrams, meridional and sagittal fans and annular curves appearing in this thesis were computed, each ray was examined for vignetting. Although it is assumed that vignetting by the surfaces is not important, the possibility that this might occur was not lost sight of during ray tracing. However, for the pencils considered, vignetting was entirely by the diaphragm. When the basal coefficients were applied, vignetting was determined according to the principles of §63. The point of intersection of the rays with Ⓓ was found by means of (63.4) and examined in the usual manner by (63.5). It was found that (for this particular system) it is quite sufficient to take (63.4) correct to the second order. If the pencil was about the base-ray, the first order terms of (63.4) are sufficient.

(b) For any pencil the area A_E of the entrance pupil is determined from (64.10). Using the values of the entrance pupil coefficients given in Table 75/1, A_E is given to the third[1] order in ρ_d and $\underset{\sim}{T}$ by the formulae

$$A_E = 2.4107 \times 10^{-2}(0.4256 + 0.1545\ T_y) \quad ,$$
$$A_E = 2.4107 \times 10^{-2}(0.5065 + 0.1918\ T_y) \quad ,$$
$$A_E = 2.4107 \times 10^{-2}(1.1905 + 0.2111\ T_y) \quad , \qquad (77.1)$$

for the base-rays (a), (b) and (c) respectively. It is immediately obvious that A_E increases with increasing field angle. If vignetting is determined to the first order only, the area of the entrance pupil will be underestimated for pencils with $T_y > 0$ and overestimated for pencils with $T_y < 0$. The cross-sectional area of any pencil can be determined by using the appropriate member of (77.1) with (64. 9). If the grid mesh (§74b) is δ the cross-sectional area associated with each ray is δ^2.

---

[1] Note that this entails only the <u>second order</u> entrance pupil coefficients and hence is equivalent to determining vignetting to the second order.

Thus, if vignetting is taken into account and spot diagrams determined in the manner of §75b, the number of rays transmitted by the system and constituting the spot diagram will be

$$N = A/\delta^2 = 4 \times 10^4 A \quad ,$$

taking $\delta = 0.005$.

Values of A and N were determined with each of the base-rays (a), (b) and (c) for various pencils and the results are presented in Table 77/1. The first column specifies the base-ray, ψ_1 is the half-field angle of the pencil, T_{Ay} and S_{Ay} the axial SPC of the principal ray of the pencil where S_{Ay} was predicted using (65.2), A is the cross-sectional area of the pencil expressed in hundredths of a square unit and N is the number of rays that are predicted to be transmitted by Ⓚ. N_{tr} is the number of rays that were actually transmitted when the traced spot diagram was constructed and ρ_{tr} is the distance from the centre of Ⓓ (as determined from the traced spot diagram) of the point of intersection of the predicted principal ray with Ⓓ. ρ_{tr}/ρ_d is the measure of the quality of the prediction of the value of S_{Ay} employed in §75a. The area A decreases slowly with the field angle. In fact it is readily verified that A is to a good approximation given by

$$A = A_0 \cos^r \psi_d \quad , \tag{77.2}$$

where ψ_d is the angle of inclination of the principal ray at the diaphragm and A_0 is a constant. For small half fields $r \approx 1.1$, and decreases to values around 0.8 for large half-fields. The result (77.2) is remarkable since it is equivalent to A being proportional to the cross-sectional area of the pencil at the diaphragm. It is evident from the close agreement of N and N_{tr} that for the purpose of obtaining an assessment of the aberrations of this system it is sufficient to consider vignetting

TABLE 77/1 Predicted Principal Rays and Cross-Sections of Various Pencils

	ψ_1	T_{Ay}	S_{Ay}	A	N	N_{tr}	ρ_{tr}/ρ_d
	30°	0.5774	0.0139	0.7249	290	295	0.0041
	35°	0.7002	0.0230	0.7202	288	290	0.0009
T_{By} = 1.0989	40°	0.8391	0.0369	0.7132	285	287	0.0011
	45°	1.0000	0.0573	0.7011	280	281	0.0001
S_{By} = 0.0719	50°	1.1918	0.0869	0.6832	273	273	0.0006
	55°	1.4281	0.1305	0.6605	264	265	0.0014
	60°	1.7321	0.1953	0.6325	253	257	0.0020
	55°	1.4281	0.1296	0.6598	264	259	0.0118
T_{By} = 2.0570	60°	1.7321	0.1952	0.6458	258	257	0.0023
	64°	2.0503	0.2718	0.6313	253	252	0.0011
S_{By} = 0.2734	67°	2.3559	0.3511	0.6179	247	248	0.0011
	70°	2.7475	0.4586	0.6033	241	239	0.0029
	74°	3.4874	0.6746	0.5907	236	239	0.0100
T_{By} = 4.9582	76°	4.0108	0.8355	0.5818	233	233	0.0030
	78.6°	4.9594	1.1348	0.5722	229	228	0.0002
S_{By} = 1.1345	80°	5.6713	1.3642	0.5663	227	222	0.0013
	82°	7.1154	1.8350	0.5579	223	217	0.0065

TABLE 77/2 Predicted Principal Rays and Cross-Sections of Various Pencils which are Skew to the Base-Ray

	T_{Ay}	T_{Az}	S_{Ay}	S_{Az}	A	N	N_{tr}	ρ_{tr}/ρ_d
T_{By} = 2.0570	1.7321	0.20	0.1964	0.0229	0.6425	257	254	0.0054
	1.7321	0.40	0.2004	0.0463	0.6331	253	264	0.0021
S_{By} = 0.2734	2.0503	0.20	0.2729	0.0268	0.6307	252	250	0.0033
	2.0503	0.40	0.2764	0.0541	0.6237	250	251	0.0039

to the second order.[2]

Table 77/2 is the analogue of Table 77/1 for four skew pencils. (64.8) cannot differentiate between skew and meridional pencils and for small values of T_z only a small variation of A with T_z is introduced by the denominator of (64.11). However, if the terms $A^{(4)}$ are included in A_E, a definite dependence on T_z is introduced into A. Note the large increase in N_{tr} as T_z is increased for the $T_{Ay} = 1.7321$ skew pencils. It would be expected that since the field angle is increased by increasing T_z, N_{tr} would decrease. The cause of this discrepancy is unknown.

78. Aberrations Associated with Annular Apertures

The ability to predetermine the shape of the entrance pupil when investigating the performance of a system means greater use can be made of analyses of the aberrations associated with annular apertures. If the diaphragm is replaced by an annular aperture of internal radius ρ and external radius $\rho+d\rho$, where $d\rho$ is small, only those rays which intersect Ⓓ at a distance ρ from its centre are actually transmitted by the system and the image patch is a closed curve called the aperture curve. By combining the curves corresponding to various values of ρ the complete image can be built up.[1] The dependence of the aberrations on the aperture may be studied in detail and information not readily obtained from spot diagrams is available along with an overall picture of the complete image.

---

2 It will be seen in §79d that if vignetting is determined to the first order, a remarkable effect actually improves in appearance the quality of the predictions of the aberrations.

1 Herzberger (1957), part IV.

The shape of the entrance pupil is given by the series (64.2) and as a first approximation second and higher powers of ρ, given by (64.3), may be neglected. This is certainly the case for pencils of small aperture and it appears as though no serious errors are made for the system under investigation if this is done. For a meridional pencil

$$S_y = p_{sy2}T_y + s_{sy5}T_y^2 + (p_{sy1}+s_{sy2}T_y+t_{sy5}T_y^2)\rho \cos\theta + O(3)$$

$$S_z = (p_{sz3}+s_{sz6}T_y+t_{sz12}T_y^2)\rho \sin\theta + O(3) \quad . \qquad (78.1)$$

ρ and θ are as in (64.3) and O(3) denotes terms of the third and higher degree in T_y. Evidently the entrance pupil is to a good approximation an ellipse centred on the point of intersection of the principal ray with the plane of the entrance pupil. When considering off-axis imagery it is not sufficient to consider the entrance pupil as a circle.[2] In fact, if polar coordinates $(\bar{\rho},\theta)$ are set up in the entrance pupil and image curves $\bar{\rho}$ = constant are constructed for various values of $\bar{\rho}$, stopping down the diaphragm will cause parts of several different curves $\bar{\rho}$ = constant to be vignetted. This is probably the reason why the aperture curves appear to have had little use. For large field angles, the pencil of rays determined by $\bar{\rho}$ = constant will in actual fact be nothing more than a collection of rays in the neighbourhood of the sagittal plane of the true pencil and the corresponding image curve will not represent the aberrations associated with any particular aperture.

By prescribing values for ρ, the corresponding aperture curve can be obtained by calculating $\underset{\sim}{S}$ according to (78.1) and substituting these values into the series for $\underset{\sim}{H}$. The origin of the aberrations is

---

2 cf. Herzberger (1957) fig.2 .

again taken to be the point of intersection of the principal ray with the image plane. By this means a set of aperture curves of the 40°, 45°, 64° and 67° pencils were produced. The predictions were all to the <u>third order</u>. For each pencil, annular apertures corresponding to 1/5, 2/5, 3/5, 4/5 full aperture were chosen (i.e., apertures of f/50, f/25, f/16.7 and f/12.5). As would be expected, some of the rays at full aperture were vignetted, due of course to the inadequacies of first order vignetting, and the corresponding aperture curves are not shown.

Fig.12 shows the aperture curves of the 40° and 45° pencil and the number by the side of each curve indicates the aperture. The predicted curves were obtained using base-ray (a). The f/50 curves are not shown but because of the presence of longitudinal spherical aberration these lie just <u>outside</u> the curves for an aperture of f/25. The f/50 traced and predicted curves agree very well for the 45° pencil but are displaced slightly with respect to one another for the 40° pencil. The symmetric appearance of all the curves is due to the strong spherical aberration present at low field angles. The predicted curves for the 40° pencil show evidence of a uniform displacement in the negative ϵ_y-direction with respect to the corresponding traced curves. This is in part due to the difference between the predicted and traced image heights of the principal ray.

Aperture curves of the 64° and 67° pencil (fig.13) were predicted using base-ray (b). Again the f/50 curves are not shown but exhibit much less asymmetry than the curves for higher apertures and their predicted and traced curves agree very well for the 64° pencil. In the case of the 67° pencil, the two curves are again displaced with respect to one another and this is due entirely to the difference between the traced and predicted image heights of the principal rays. For each of the three apertures shown in fig. 13 the curves exhibit asymmetry, especially the f/16.7 and f/12.5 curves. The agreement for the 64^{o} pencil is good, but this is not the case for the higher aperture curves of the 67^{o} pencil. It is to be noted that in all cases there are regions where the traced and predicted curves are in close agreement and correspondingly there are regions where just the opposite is the case. These regions correspond to small and large values for $\underline{S}$ respectively.

79. Spot Diagrams in the Ideal Image Plane

(a) A spot diagram pictorially represents the points of intersection with the image plane of a pencil of rays whose members intersect a plane normal to the principal ray of the pencil in uniformly distributed points. Consequently the number of points on the spot diagram contained in a small circle centred on a given point in the image plane is proportional to the intensity at that point and hence a spot diagram not only represents the general shape of the image patch but also the intensity at any point in the image. A typical spot diagram consists of two parts: the core and the flare.[1] The core is that part of the diagram where the intensity or point density is the highest and the flare is the weakly illuminated region surrounding the core. Usually these two regions are readily identified, especially in image planes near the plane of best focus. The core principally determines the resolving power of the lens and the flare determines the contrast produced by the lens. In particular, the core determines the high-contrast resolving power and the flare determines the low contrast resolving power.[1] The transfer function can be determined

---

1 Stavroudis and Sutton (1965) §3.

from the coordinates of the points on the spot diagram.[2] Spot diagrams represent the quality of the image and take no account of the distortion. Thus it is irrelevant where the origin for the aberrations is taken and a difference in origin for the predicted and traced diagrams is of no consequence (cf. §75b,78,81b).

(b) For various field angles the spot diagrams in the ideal image plane were predicted to the third order using the three base-rays and are compared in fig.14,15 and 16 with the corresponding spot diagram determined by ray tracing. In all cases the right hand diagram was predicted and the corresponding left hand diagram was traced. The grid mesh is $\delta = 0.005$ and on the diagrams the scale is 1cm. = 0.002 focal lengths. The ϵ_y-axis is vertical.

The general character of the images as deduced from the traced diagram is as follows: at low field angles the image is symmetric (spherical aberration) with a small, well defined core and large, evenly illuminated flare. As the field angle is increased the core grows and the flare shrinks. This is due to changes in the longitudinal spherical aberration. The image remains symmetric up to nearly 60° half-field where coma is barely visible in the flare. The flare is asymmetric beyond 60° and becomes highly comatic at about 70° half-field. It shrinks, though remaining comatic, and is almost absorbed into the greatly expanded core at 80° half-field. The core itself grows and becomes noticeably comatic at about 67° half-field. At low field angles the principal ray intersects the image plane in the centre of the core, whereas at large field angles

---

2 Kubota and Miyamoto (1963) §3.

the point of intersection is at the top edge of the spot diagram (as they appear in the figures).

Traced and predicted spot diagrams for the two skew pencils T_A = (1.7321,0.2) and T_A = (2.0503,0.4) are shown in fig.17 where the same conventions and scale apply as before. Again the predictions are third order. Since the system is symmetric, all pencils are meridional[3] and in this case their meridional[3] planes are inclined to Ⓜ. The spot diagrams must therefore be symmetric about a line in the image plane both through the point of intersection of the principal ray and the axis of the system. This is the case for the traced diagrams. The apparent asymmetry in the detailed structure of the diagrams arises from the manner in which they were determined - the selected rays were not symmetrically distributed with respect to the relevant meridional[3] plane.

(c) Generally, the entire third order predicted spot diagram for a pencil about the base-ray is in excellent agreement with the corresponding traced diagram. As the magnitude of the basal coordinate $\underset{\sim}{T}$ is increased, the disagreement between the flares of the predicted and traced spot diagrams becomes pronounced. However, the predicted and traced cores are almost invariably in excellent agreement. In particular, compare the predicted and traced spot diagrams for the pencil at a half-field of 70°. Moreover, usually only a portion of the flare is in poor agreement and the rays constituting these portions are associated with large values of $\underset{\sim}{S}$.

The departure from symmetry about some line in the image plane

---

3 This refers to the usual sense of "meridional" in the context of the axial theory.

is pronounced for the predicted spot diagrams of the skew pencils (fig.17). This can be due to two factors: first, in the basal theory the image produced by a skew pencil is intrinsically asymmetric (§50a). Second, if vignetting is determined to the second order for skew pencils, the cross section of the pencil is predicted to be virtually independent of S_z (§77b) and thus rays which are in fact transmitted are predicted to be vignetted, or conversely. In the case at hand, the second possibility may be discounted and it appears as though the inclusion of the first three orders is not sufficient to correct the intrinsic asymmetry of the predicted image. At first sight it appears as though the predictions for the skew pencils were outstandingly poor. However, the core is reproduced satisfactorily, especially with respect to its size and so are some parts of the flare. In the predictions the top left hand corner of the flare appears to have been displaced into the diagonal regions of the spot diagram. These displaced regions can again be correlated with large values of S_y or S_z.

It is instructive to compare fig.12,13 for the aperture curves with the corresponding spot diagrams. These show that (i) the flare corresponds to apertures greater than about f/20 for low half-fields and f/16 for higher half-fields,[4] and (ii) the disagreement between the

---

4 The result (i) illustrates two things. First, by stopping the system down to about f/16 the intensity of the core is virtually unaltered whereas the extent of the flare is greatly reduced. Thus, without increasing the exposure required to expose the core correctly, the resolution and contrast is increased. Second, although a system may be nominally of aperture ν, say, its effective aperture, as determined from the exposure required to produce a correctly exposed negative, may be much less than ν.

predicted and traced aperture curves increases sharply with aperture from about f/16, consistent with the observed discrepancies in the flares. Spot diagrams were also determined for the 45° and 55° pencils with the system working at f/14 and are presented in fig.18 which is on the same scale as fig.14-17. The core is apparently unaffected but the flare is greatly reduced. However, the discrepancy between the predicted and traced flares still remains.

(d) It is in general advisable to determine vignetting to the second order when the basal coefficients are used. However, for this system better agreement between the appearances of the predicted and traced spot diagrams is paradoxically obtained if vignetting is determined only to the first order! For example, fig.19 compares the traced spot diagram for the pencil at 55° half-field with the corresponding diagrams predicted with the coefficients of base-ray (a) when vignetting is determined (i) to the first order (centre diagram) and (ii) to the second order (right hand diagram). As a second example (not shown), the prediction of the diagram for the pencil at 40° half-field is also improved if vignetting is determined only to the first order. These surprising results may be partially explained in terms of the entrance pupil coefficients.

A necessary condition for the spot diagram corresponding to first order vignetting to be in better agreement with the traced diagram than is the diagram corresponding to second order vignetting can be determined as follows: if $T_y > 0$, those rays vignetted to the first order in excess of those vignetted to the second order must have their aberrations poorly predicted, that is, their values of S_y must lie in a neighbourhood of the maximum value for S_y consistent with vignetting. Similarly, if $T_y < 0$, the additional rays transmitted when vignetting is determined only to the first order, must have S_y in a neighbourhood of its minimum. Assume that for the meridional pencil $(T_y,0)$ the entrance pupil corresponding to a diaphragm of radius ρ is given by (78.1). Write the second order terms of $\underset{\sim}{S}$ as

$$s_y^{(2)} = s_{sy2}\rho T_y \cos\theta + s_{sy5}T_y^2 \quad , \quad s_z^{(2)} = s_{sz6}\rho T_y \sin\theta \quad .$$

Then the required condition is

$$s_y^{(2)} > 0 \quad , \tag{79.1}$$

both for $T_y > 0$ with S_y near its maximum and for $T_y < 0$ with S_y near its minimum. ($s_z^{(2)}$ is of no concern and need not be considered.) (79.1) is an inequality in $\rho \cos\theta$ and without loss of generality it is sufficient to consider only meridional rays. Then $-\rho_d \leq \rho \leq \rho_d$ and $\theta = 0$. Since interest is centred on rays which are on the verge of being vignetted, $s_y^{(1)}$ is certainly the numerically dominant term of S_y. Consequently, if p_{sy1} is positive S_y will be a maximum for ρ near ρ_d and a minimum for ρ near $-\rho_d$. Otherwise, if p_{sy1} is negative, the maximum is near $-\rho_d$ and the minimum near ρ_d. The condition (79.1) then becomes:

> <u>The inequality</u> $s_{sy2}\rho T_y + s_{sy5}T_y^2 > 0$ <u>must be satisfied by</u> $\rho = \rho_d p_{sy1}/|p_{sy1}|$ <u>if</u> $T_y > 0$, <u>or by</u> $\rho = -\rho_d p_{sy1}/|p_{sy1}|$ <u>if</u> $T_y < 0$. (79.2)

Note that if s_{sy5} is positive and p_{sy1} and s_{sy2} are of the same sign then the above condition is automatically satisfied. That this is the case for the system under consideration is evident from Table 75/1. In other cases the condition may or may not be satisfied for all or some values of T_y. The condition is necessary but <u>not</u> sufficient - it may well be that the included or excluded rays "over correct" the predicted spot diagram. Thus it is partly fortuitous that the correction is so good in fig.19.

80. On the Effects of a Translation of the Image Plane

(a) An indication of the use of the basal coefficients for determining the effects on the image of a longitudinal translation of the image plane through a distance x was given in §51a. The present section is devoted to a numerical consideration of these effects. First of all, from the values of the h- and v-coefficients (Tables 73/1,2) it is evident that in general the first order coefficients h_a are sensitive to quite small translations of the image plane ($x \approx 0.01$, say) whereas the higher order coefficients are unaffected by such changes.

Consider the meridional fan of the pencil about base-ray (a). $\epsilon_y^{(3)}$ dominates $\epsilon_y^{(1)}$ when

$$28.87\ S_y^3 > 0.019\ S_y \quad , \quad \text{i.e., } |S_y| > 0.025 \quad .$$

$\epsilon_y^{(2)}$ is negligible (see §74a). In the region where the third order aberration dominates, the total aberration is large. Hence it is necessary to reduce $\epsilon_y^{(3)}$. If the third order <u>coefficients</u> are reduced by performing a translation, the first order aberration is increased enormously. Rather, it is necessary to balance the large $\epsilon_y^{(3)}$ by introducing first order aberrations of the opposite sign. h_{ya} must be increased and since v_a is negative the image plane must be brought nearer to the lens. The effect of this will be to reduce the overall dimensions of the image but also to increase the aberrations in regions where $|\epsilon_y^{(1)}| > |\epsilon_y^{(3)}|$ (i.e., the core). To double, say, the range of values of S_y such that $\epsilon_y^{(1)}$ is numerically greater than $\epsilon_y^{(3)}$ it is necessary to increase h_{ya} four-fold and, from (53.2), x is given by $4h_{ya} = h_{ya} + xv_a$. Hence $x = -0.071$ and the second and third order coefficients are changed by negligible amounts. These considerations were based on the meridional aberrations alone and since the aberration is principally spherical aberration, substantially the same conclusions are arrived at by considering the sagittal aberrations.

(b) The pencil about base-ray (c) cannot be treated in this simple manner since the aberration is comatic. An alternative process is discussed below and illustrated by applications to the pencils with half-fields of 45° and 78.6° respectively. Since the aberration is referred to the principal ray, $\underset{\sim}{\epsilon}$ is given by (7.8). In the translated image plane the image height is $\underset{\sim}{H}(x)$ where

$$\underset{\sim}{H}(x) = \underset{\sim}{H} + x\underset{\sim}{V}$$

and thus

$$\underset{\sim}{\epsilon}(x) = \underset{\sim}{\epsilon} + x(\underset{\sim}{V}-\underset{\sim}{V}_P) \quad , \tag{80.1}$$

(cf. (66.2)). For the meridional fan, $\epsilon_y(x)$ is zero when

$$x = -\epsilon_y/(V-V_P) \quad , \tag{80.2}$$

and for the sagittal fan, $\epsilon_z(x)$ is zero when

$$x = -\epsilon_z/W \quad . \tag{80.3}$$

(80.2,3) express x as series in $\underset{\sim}{S}$ and $\underset{\sim}{T}$ for the appropriate fan. In the limit of zero aperture x will be a function of T_y for both the meridional and sagittal fans and will be the generalisation of $\underset{\sim}{x}_M$ to any meridional principal ray. The zero order terms in these series are $\underset{\sim}{x}_M$. In both (80.2,3) x behaves as a quadratic in ρ for finite apertures since the dominant term of $\underset{\sim}{\epsilon}$ becomes cubic and that of $\underset{\sim}{V}-\underset{\sim}{V}_P$ remains linear. (Note: ρ is the ρ of §76.) If the aberration is symmetric, x is an even function of ρ. It is necessary that the values of x determined from (80.2,3) should be in good agreement with the corresponding values determined by ray tracing. That this is the case is shown by fig.20 which compares the traced and predicted values of x for the 45° pencil (base-ray (a)) and the 78.6° pencil (base-ray (c)). One unit is 0.01 focal lengths and the curves labelled M and S correspond to (80.2) and (80.3) respectively.

As the first example consider the meridional pencil of 45^o half-field. Since the aberrations are symmetric there is no need to consider the sagittal fan. Under the assumption that the core arises from rays for which $\underset{\sim}{\epsilon}$ is a slowly varying function of $\underset{\sim}{S}$, inspection of fig.7 shows that the core[1] corresponds to $-0.02 < \rho < 0.018$ and has a radius of 0.0002. The flare is large and arises from $|\rho| > 0.02$ and the aberrations in this region must be reduced. From fig.20 the aberration for $\rho = 0.032$ is reduced to zero by $x = -0.06$ and this is taken as the position of the translated image plane. In this plane the aberration of any ray may be determined from the meridional curve and fig.20: let x_0 be the position of the new plane, then (80.1,2) gives

$$\epsilon_y(x_0) = \epsilon_y(1-x_0/x) \quad , \quad \epsilon_z(x_0) = \epsilon_z(1-x_0/x) \qquad (x \neq 0) \quad , \qquad (80.4)$$

where $\underset{\sim}{\epsilon}$ is the aberration in the undisplaced image plane and x is the position of the image plane in which the aberration of the ray in question is reduced to zero. (<u>Note</u>: x will in general be different for ϵ_y and ϵ_z.)

The boundary of the flare will still correspond to the extreme rays $|\rho| = 0.04$ and (80.4) gives the aberration of the ray for which $\rho = 0.04$ as -0.0019. The boundary of the core in the translated plane corresponds to the rays $|\rho| = 0.01$, for which (80.4) gives $\epsilon_y = 0.0011$. The radius of the flare is now about twice that of the core. That this is indeed close to the mark is evident from fig.21 where the predicted and traced spot diagrams in the image planes x = 0.0, -0.04, -0.06 and -0.08 are compared for the 45^o pencil. The highest resolution will evidently be obtained near the ideal image plane but the contrast will be poor.

---

[1] The rim of the core is taken to coincide with the local maximum and minimum of ϵ_y.

In the plane $x = -0.07$ (approx.) the contrast will be a maximum. In the previous sub-section $x = -0.07$ corresponded to the plane in which $\epsilon_y^{(1)}$ was greater than or of the order of $\epsilon_y^{(3)}$ for all ρ. Presumably this corresponds to the flare being absorbed into the core. For the 45° pencil this occurs for ρ between -0.06 and -0.08. The image plane in which the flare is just absorbed into the core is a generalisation of the plane of least confusion for spherical aberration.

(c) Inspection of the meridional curve of the 78.6° pencil (fig.9) shows that $\rho > 0.01$ corresponds to a weakly illuminated flare. Two cores corresponding to $\epsilon_y = 0.0_33$ and -0.0023 arise around $\rho = 0.0065$ and -0.022 respectively. It is desired to reduce the relative aberration of these cores. A reduction of the aberrations of a ray will occur in the translated image plane provided $0 < x_0/x < 2$ (see (80.4)). Thus, if $x_0 > 0$ the aberrations in most of the range $\rho < 0.01$ will be reduced and the corresponding meridional curve will be considerably flattened over this range. It appears as though $x_0 = 0.04$ would be a reasonable choice and the aberrations for $\rho < 0.0075$ will be reduced with this choice. The cores are then at $\rho = 0.002$, $\epsilon_y = 0.0_417$ and $\rho = -0.01$, $\epsilon_y = -0.0_351$. The flare will be extended further into the negative ϵ_y direction but is weakly illuminated.

In order to analyse the sagittal fan it is necessary to know ϵ_y and ϵ_z as functions of ρ. Determined with base-ray (c), the values of ϵ_z corresponding to various values of ρ are:

ρ	ϵ_z	ρ	ϵ_z
0.0	0.0	0.030	0.00226
0.010	0.00104	0.035	0.00223
0.020	0.00187	0.045	0.00158
0.025	0.00213	0.055	-0.0_436

The corresponding values of ϵ_y may be obtained from the sagittal curve (fig.11). For sagittal fans, the coefficients of the only non-zero terms in $V-V_P$ and H_y are s_{v8} and s_{hy8} respectively. It is evident from Tables 73/1,2 that ϵ_y is unaffected by a translation of the image plane since s_{hy8} does not vary by any sensible amount. If $x_0 = 0.04$, ϵ_z is reduced for $|\rho| < 0.048$ and has a maximum of 0.0013 at $|\rho| = 0.025$. The corresponding value of ϵ_y remains small, viz. -0.0_35 . For $|\rho| < 0.025$, ϵ_z varies slowly with ρ but increases more rapidly for $|\rho| > 0.025$. The intensity is hence greatest in a narrow range of values of ϵ_y and the corresponding values of ϵ_z are in the range -0.0013 to 0.0013, or roughly an order of magnitude greater than the range for ϵ_y. Thus there will be a bar of light in the ϵ_z direction with local cores at its ends and a flare in the negative ϵ_y direction. Since the cores of the meridional fan occur in the same narrow range of values of ϵ_y, it is likely that the complete image has the form of a meridional focal line with a flare in the negative ϵ_y direction.

In a similar fashion ϵ_z can be reduced to zero and the behaviour of ϵ_y examined. Since ϵ_z is zero at $\rho = 0$ and in the outer regions of the aperture, it is sufficient to select a value of x that reduces the maximum of ϵ_z. A suitable value is $x = 0.08$, reducing the aberration for $|\rho| = 0.034$ to zero and increasing the aberrations for $|\rho| > 0.045$. (The last region corresponds to a flare since ϵ_z varies rapidly with ρ. Consistent with the observations on the 45^o pencil, the flare is increased when the core is reduced.) In the translated plane the maximum of ϵ_z is 0.0_358 and corresponds to a strong core. There is a weak flare extending to $\epsilon_z = \pm\ 0.003$ with ϵ_y large and negative. The form of the meridional curve is as follows: for $\rho < -0.026$ and $\rho > 0.003$ the aberrations are worse with an overall range from -0.0077 to 0.0023 and no turning points. Thus the complete image is probably a strong sagittal focal line with a comatic flare in the negative ϵ_y direction.

The conclusions arrived at above should be compared with fig.22 where the predicted and traced spot diagrams in the image planes x = 0.08, 0.04, 0.0, and -0.04 for the 78.6° pencil are compared. It is seen that the aberrations of this pencil are comatic with astigmatism in the core. It is not until x is negative that a plane of least confusion is obtained. At x = -0.04 the flare is absorbed into the core but the latter is now so large that both resolution and contrast will be seriously impaired.

81. The Effects of the Wrong Choice of Base-Ray

In §68 reasons were given as to why a principal ray should be chosen as base-ray. These are now illustrated by specific examples. As an alternative base-ray to (a) take the ray (a') whose coordinates are $S_{By} = 0.04$, $T_{By} = 1.0989$ and which is not a principal ray. It is of interest to compare the h-coefficients for base-ray (a) with the corresponding coefficients for base-ray (a'), given in Table 81/1. The prominent feature is that the coefficients of second order, zeroth degree coma are quite large and hence $\epsilon_y^{(2)}$ cannot be neglected in spite of the fact that the image is known to be highly symmetric. The reason is that base-ray (a') does not go through the centre of the entrance pupil and hence the coefficients may not be interpreted as in §52,53.

The spot diagrams for the 45° pencil predicted with base-rays (a) and (a') are compared with the corresponding traced diagram in fig.23. Since the basal coordinate S_y (with respect to base-ray (a)) of base-ray (a') is negative, the prediction of some portions of the 45° spot diagram

which correspond to $n\delta < 0$ (§74b) will be better for base-ray (a') than for base-ray (a). That this is so can be seen only on close examination of fig.23 but the overall prediction is far better for base-ray (a). The meridional curves for the 35°, 47.7° and 55° pencils are shown in fig.24. The aberrations are referred to the principal ray, which is taken as that ray for which Y_d is a minimum. The overall agreement is far superior with base-ray (a) than with base-ray (a'). For the 47.7° pencil the effects of a bad prediction of the image height of the principal ray may be seen. In this case $T_y = 0$ and base-ray (a') is the ray with $\rho = -0.0214$. Its aberration is of course predicted exactly with base-ray (a') and hence, in order to correct for the inaccuracies in the prediction of the image height of the principal ray, the curve predicted with (a') must be translated in the negative ϵ_y direction so as to agree exactly with the traced curve at $\rho = -0.0214$.

TABLE 81/1 The *h*-coefficients for Base-Ray (a')

h_{ya}	= -0.06643	h_{za}	= 0.00237
h_{yb}	= 0.45190	h_{zb}	= 0.75353
s_{hy1}	= 3.03020	s_{hz3}	= 2.14741
s_{hy2}	= -0.79878	s_{hz4}	= -0.14078
s_{hy5}	= -0.17149	s_{hz6}	= -0.24749
s_{hy8}	= 1.39715	s_{hz7}	= -0.26547
s_{hy9}	= -0.16433		
s_{hy10}	= -0.13425		
t_{hy1}	= -40.36110	t_{hz3}	= -41.06781
t_{hy2}	= 15.07359	t_{hz4}	= 2.29819
t_{hy5}	= -2.28246	t_{hz6}	= 10.49456
t_{hy8}	= -53.63008	t_{hz7}	= -0.75556
t_{hy9}	= 4.92226	t_{hz12}	= -0.81579
t_{hy10}	= -0.31421	t_{hz13}	= 0.11522
t_{hy11}	= 0.21461	t_{hz17}	= -54.33189
t_{hy14}	= 7.32878	t_{hz18}	= 9.15856
t_{hy15}	= -0.88567	t_{hz19}	= -0.58480
t_{hy16}	= 0.10302	t_{hz20}	= -0.11415

82. Higher Order Coefficients and the Basal Predictions

(a) This chapter has been in part devoted to an analysis of the quality of the agreement between the aberrations predicted using the basal coefficients on the one hand and determined by ray tracing on the other. Predictions were made to the third order in the basal variables with the exception of some second order predictions made with meridional and sagittal fans. The second order predictions were quite useless. The fact that the third order predictions of particular aberrations may not be in good agreement with the traced aberrations does not imply the existence of some deficiencies or errors in the basal theory, rather the presence of certain higher order aberrations in amounts that cannot be neglected is implied. Similarly, good agreement indicates the absence of certain higher order aberrations but not necessarily of the individual higher order aberration-types since two or more aberration-types may annul each other.

To an experienced designer these discrepancies are a valuable pointer to the type and magnitude of the higher order aberrations. Spot diagrams provide scarcely more than a qualitative indication of the more important of these aberrations, whereas the meridional, sagittal and aperture curves can provide some quantitative information. The problem of analysing these discrepancies is, for several reasons, considerably more complicated in the basal theory than in the axial theory. First, the entrance pupil cannot be treated as a fixed circular aperture and the range of values of S_y is no longer symmetrical with respect to the origin,[1]

---

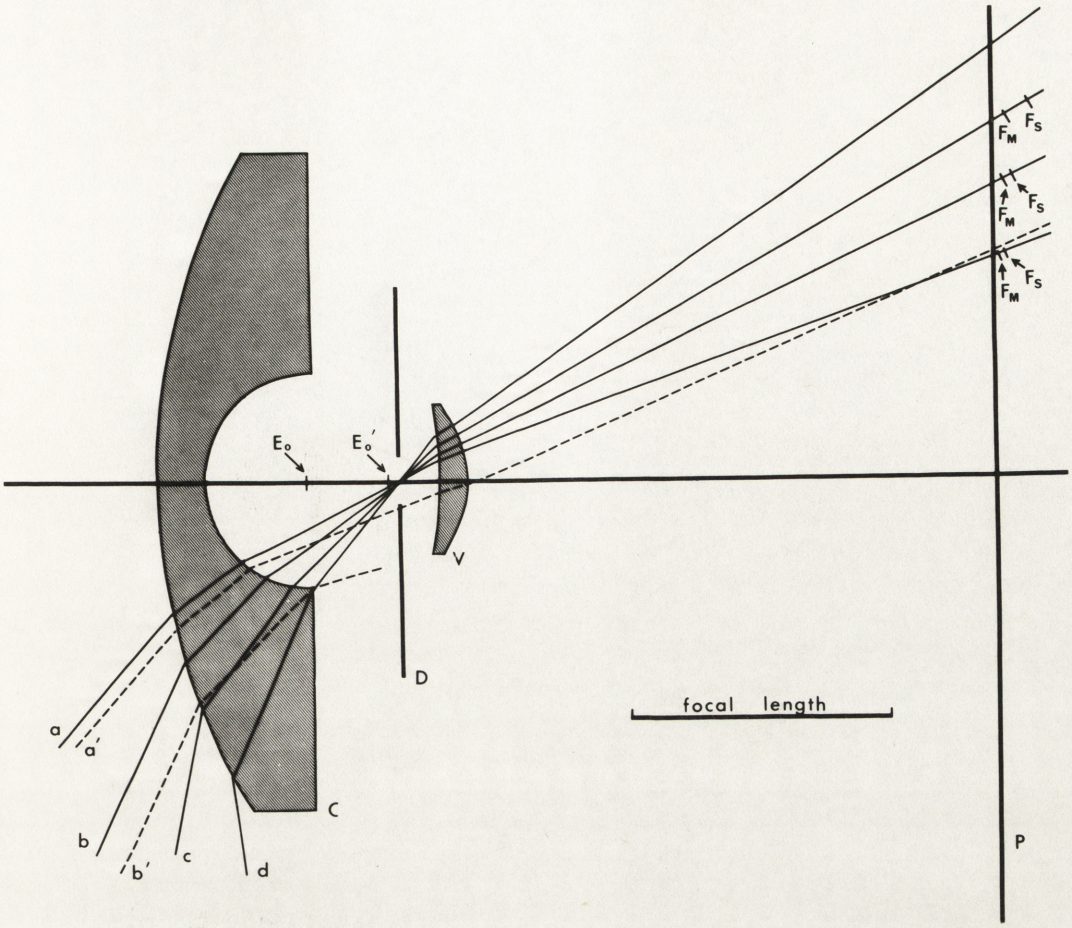
[1] It would have been better to have re-expressed $\underset{\sim}{H}$, the aberration, etc., in terms of $\bar{\underset{\sim}{S}} = \underset{\sim}{S} - \underset{\sim}{S}_p$ and $\underset{\sim}{T}$ where $\underset{\sim}{S}$ and $\underset{\sim}{T}$ are the basal SPC of a ray and $\underset{\sim}{S}_p$ is the value of $\underset{\sim}{S}$ for the principal ray of the pencil. The image height of the principal ray is then simply the distortion term of $\underset{\sim}{H}$, and $\bar{\underset{\sim}{S}}$ is symmetrically defined for all pencils.

in fact S_y may take exclusively positive or negative values for a pencil sufficiently inclined to the base-ray (see §82b below). Second, the number of coefficients and aberration-types is greatly increased.

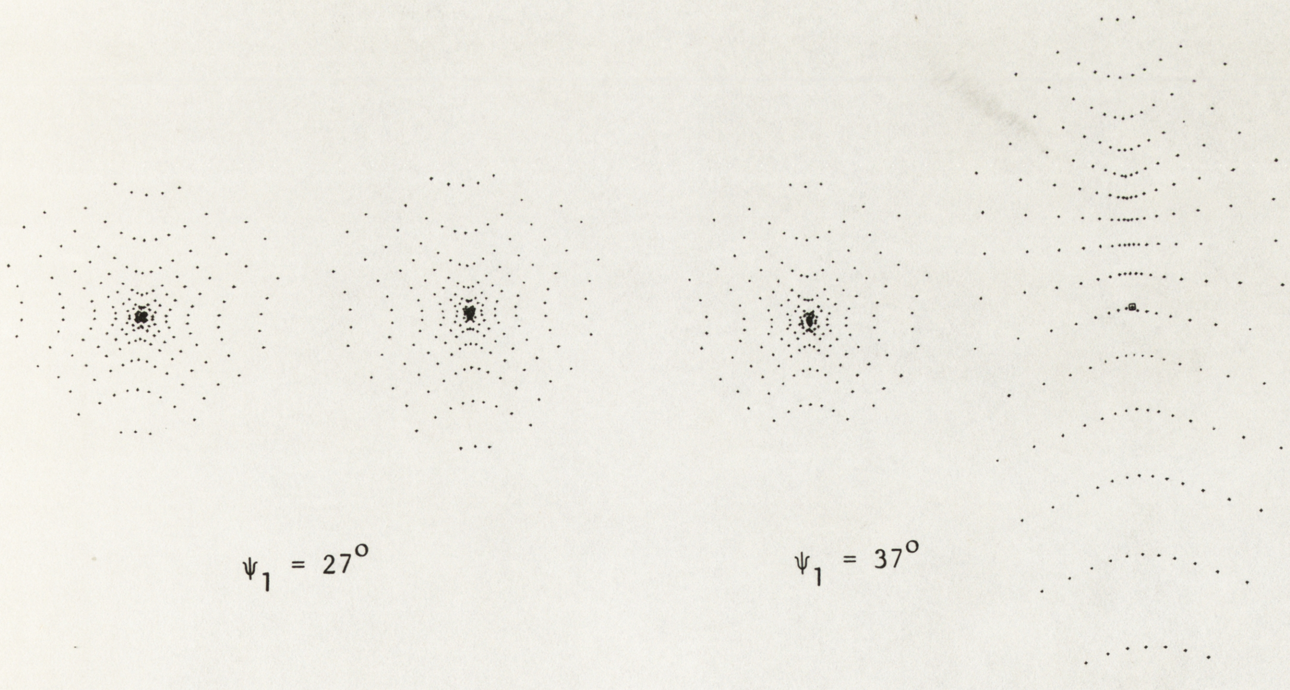
(b) Consider the pencils about the base-rays, in particular the pencil about base-ray (a). The discrepancies in the meridional fans are not symmetric and must therefore arise from a combination of terms proportional to S_y^4 and S_y^5. The coefficient of the former must be positive and of the latter, negative. Fitted to the observed discrepancies for the 47.7° pencil it is found that to a close approximation the higher order terms are $42\ S_y^4 - 470\ S_y^5$. For the sagittal fan ϵ_y can depend only on S_z^2 and to account for the discrepancies in the 47.7° fan the single higher order term $22\ S_z^4$ provides an excellent correction.

Not much emphasis can be placed on the numerical size of the variables $\underset{\sim}{S}$ and $\underset{\sim}{T}$ as a guide as to whether or not a particular term is negligible. For instance, for the 35° pencil $T_y = -0.4$ and $-0.1 < S_y < 0$ whereas for the 80° pencil $T_y = 0.71$ and $0.05 < S_y < 0.37$ and yet the predictions are as good in each case! If T_y had been 0.7 or $S_y \approx 0.4$ for base-ray (a) the predictions would have been useless.

APPENDIX A: FIGURES 1-24



<u>Fig.1</u> The sky-lens of Havlicék



<u>Fig.2</u> Spot diagrams predicted with axial theory

(Traced diagrams on the left of each pair and tertiary predictions on the right.)

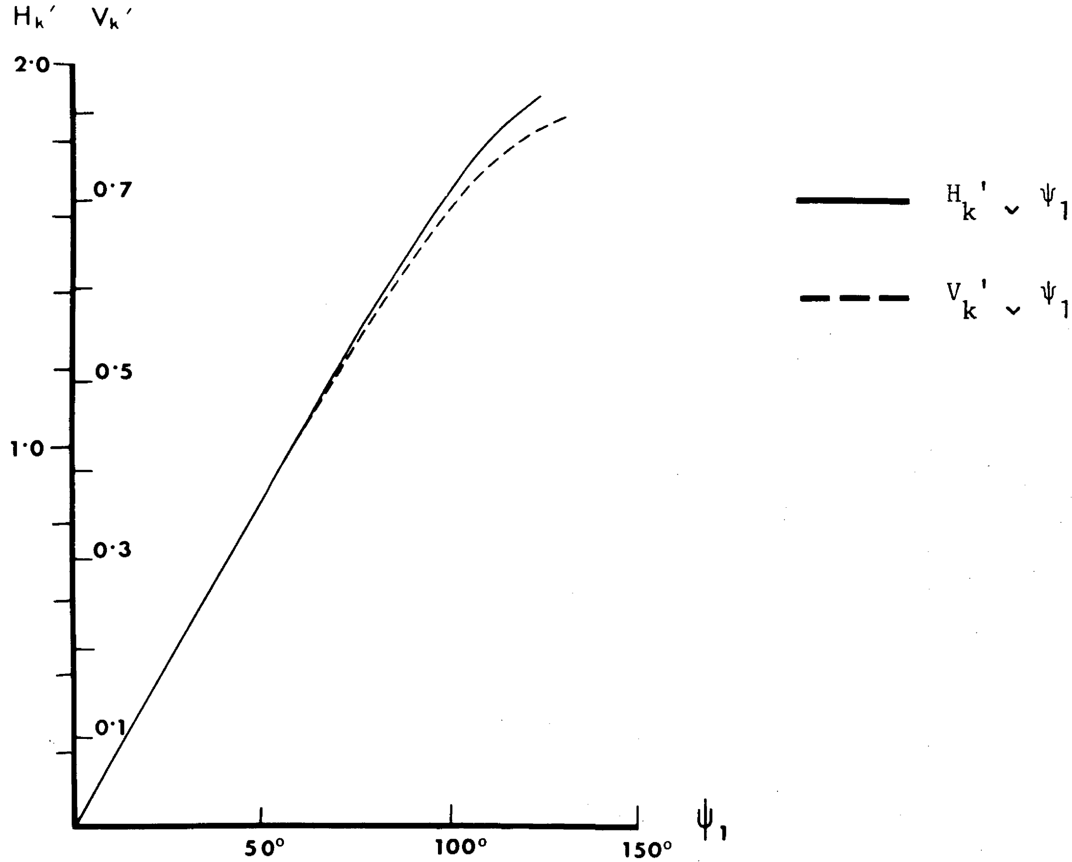


Fig.3 The nature of the projection of the hemisphere

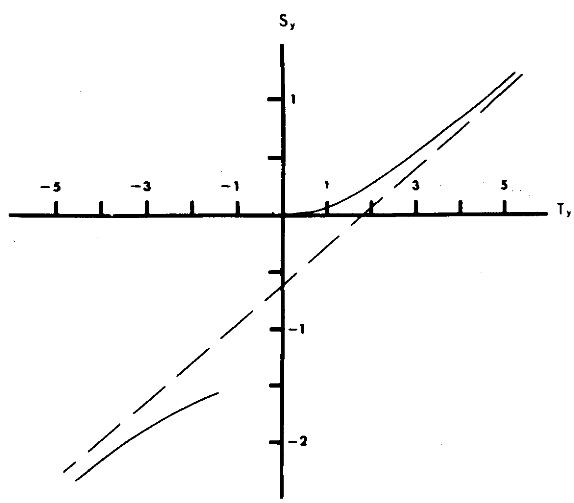


Fig.4 SPC of proper principal rays

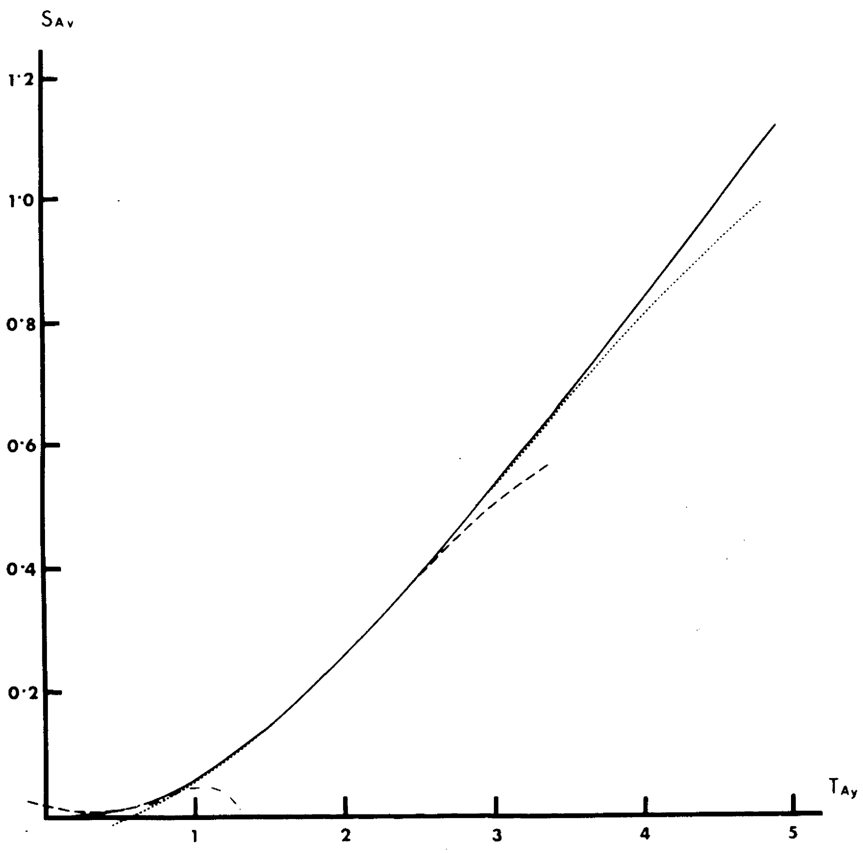


Fig.5 Predicted SPC of principal rays

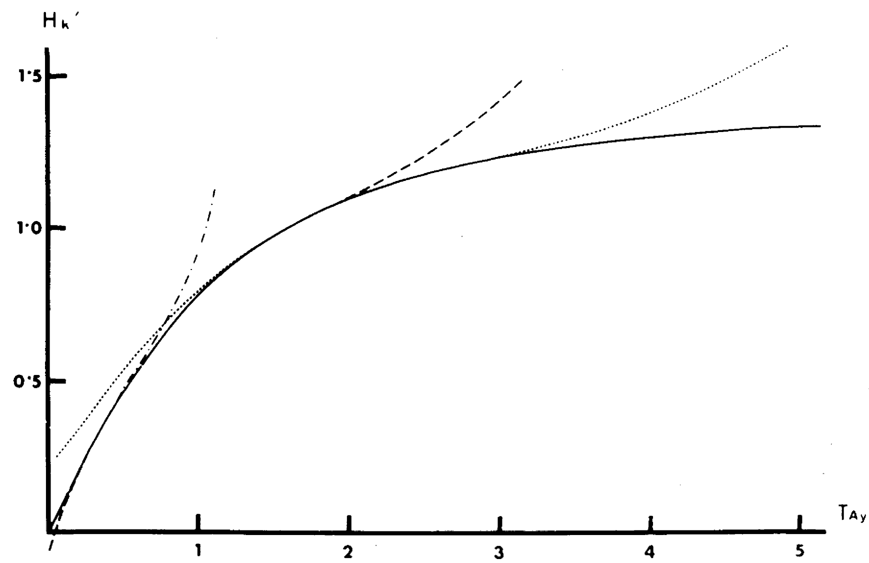
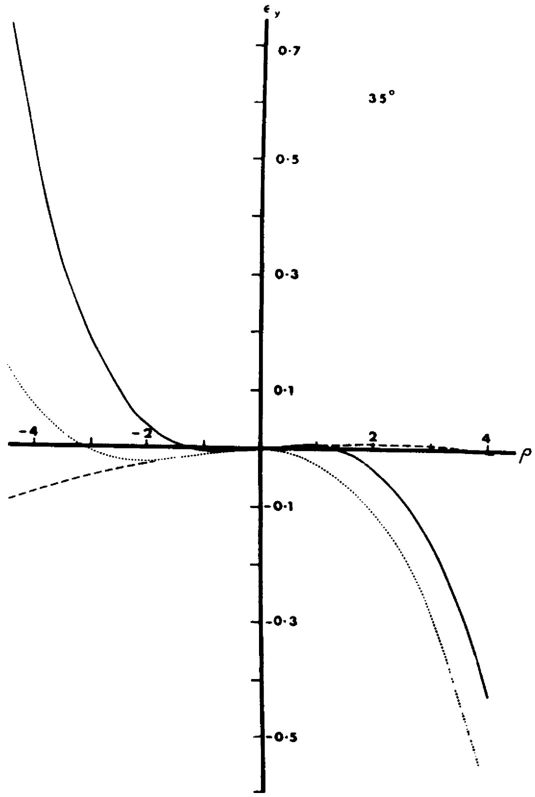
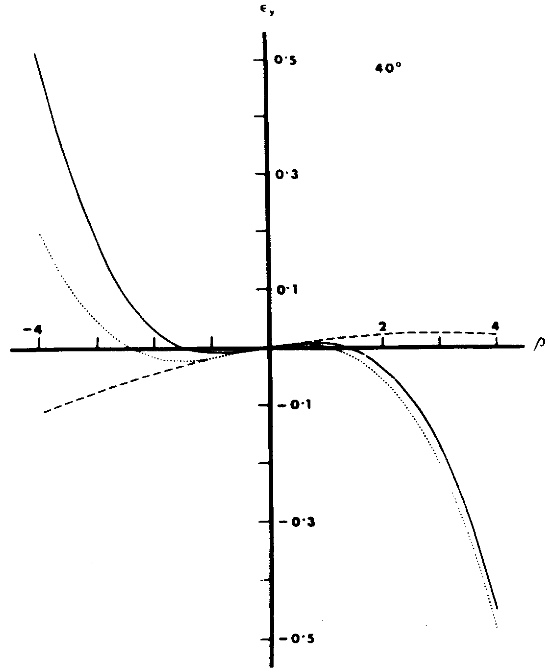
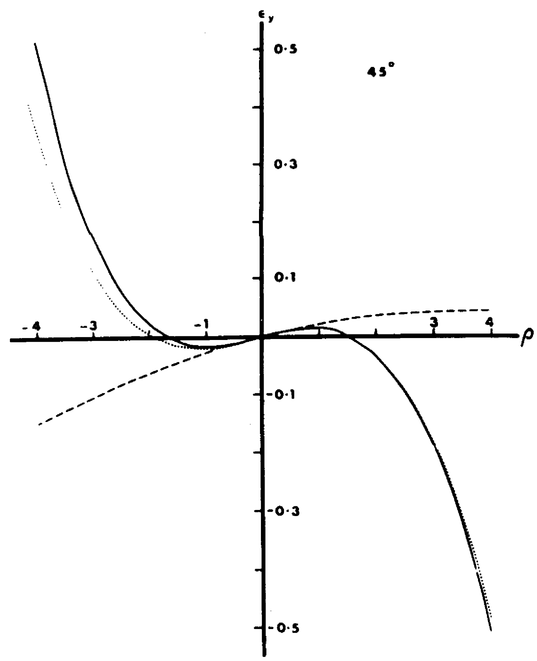


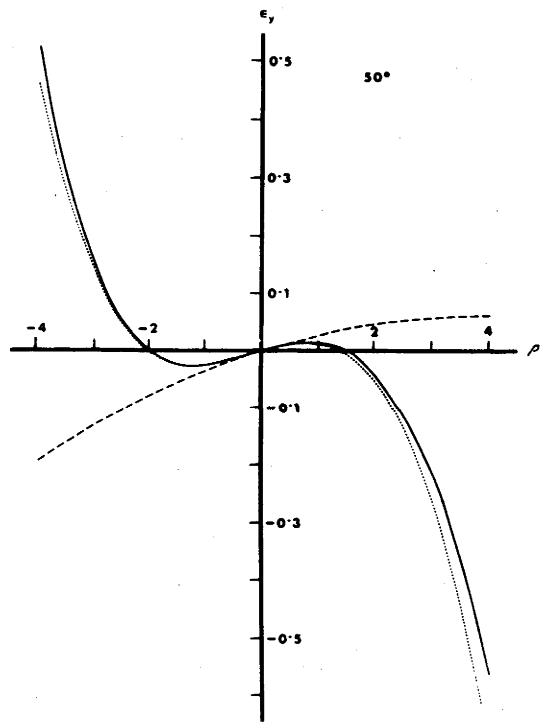
Fig.6 Predicted image heights of principal rays

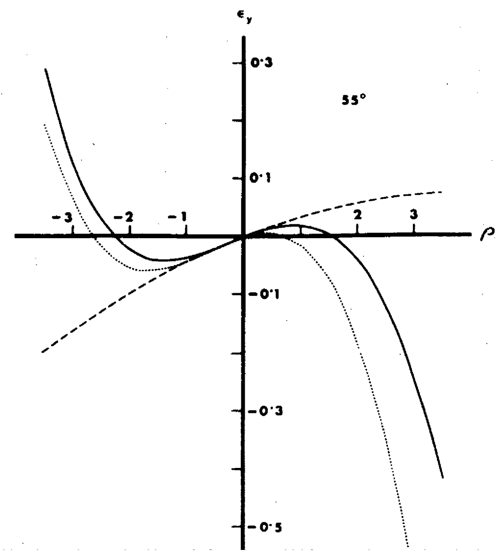
In both these figures ———— is traced, ·—·— is secondary axial prediction, — — — is third order base-ray (a), and ·········· is third order base-ray (b).





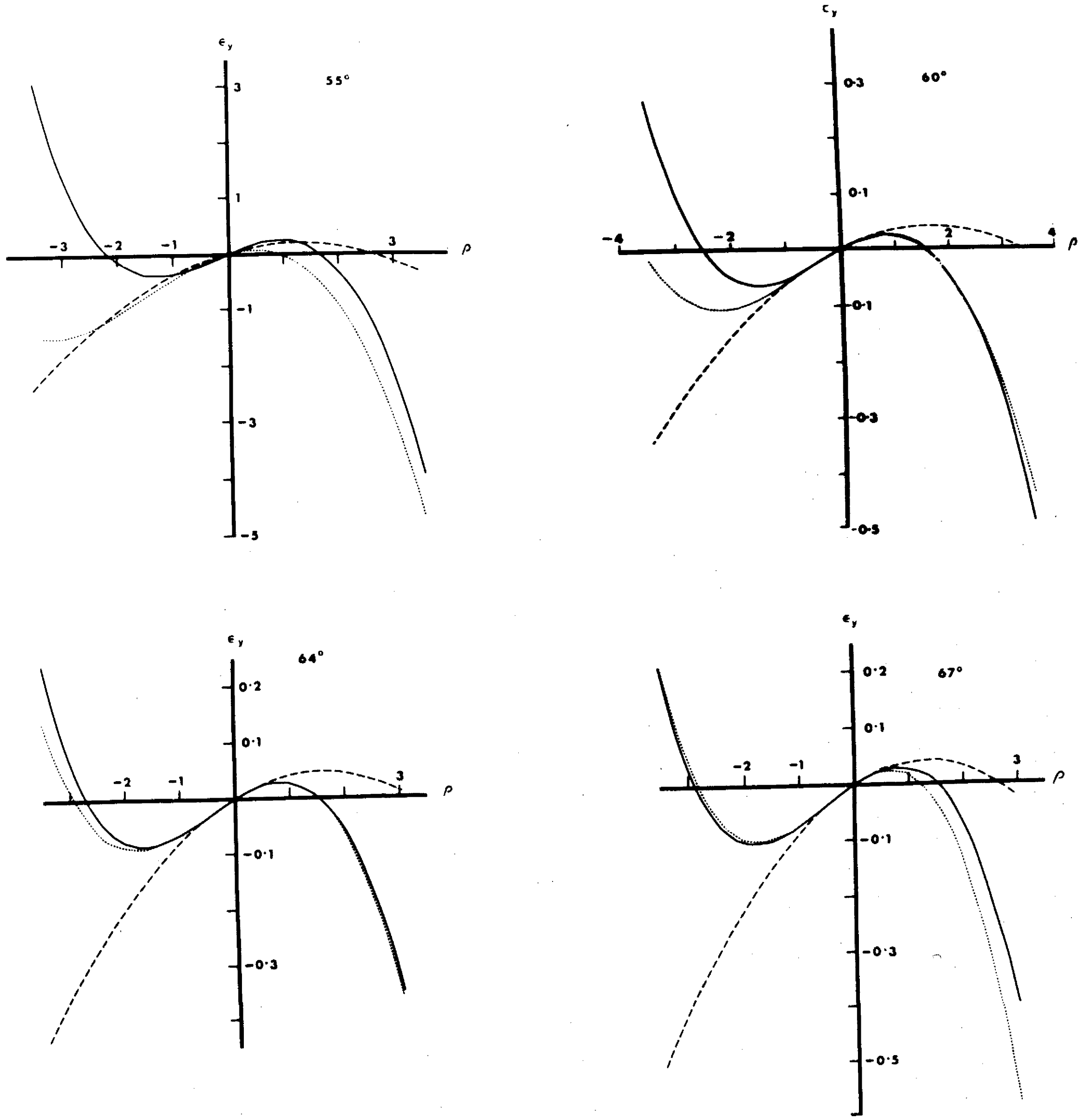






<u>Fig.7</u> Meridional fans, base-ray (a)

——— traced
— — — second order basal
.... third order basal



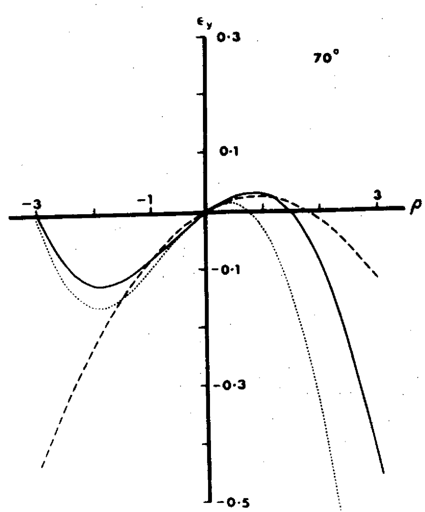
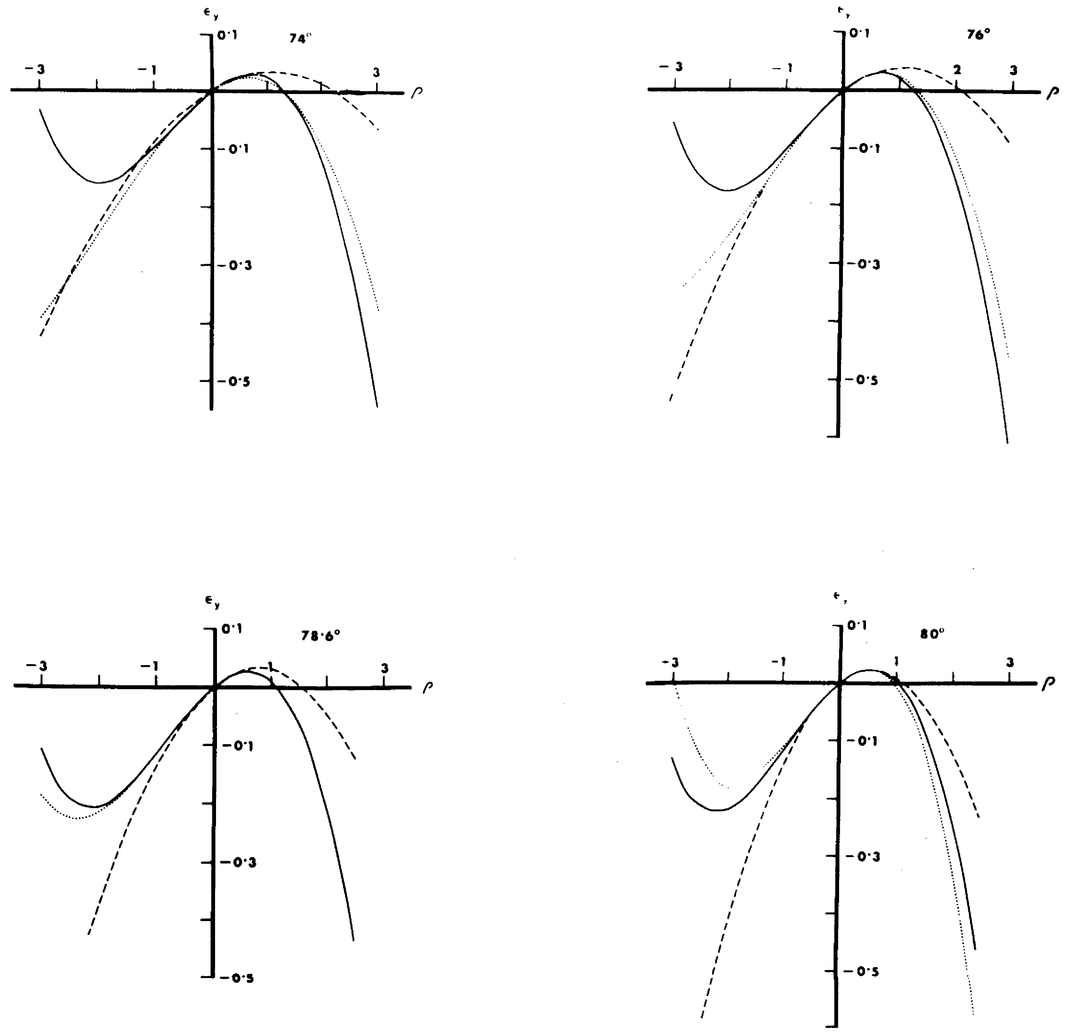


Fig.8 Meridional fans, base-ray (b)

traced

second order basal

third order basal



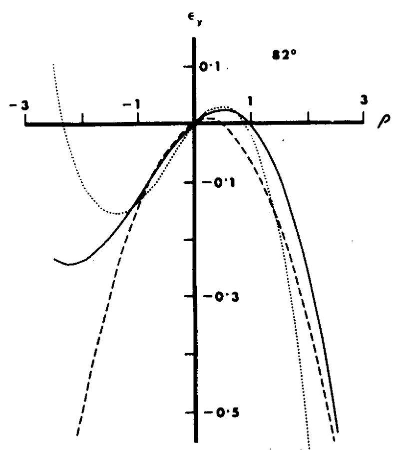


Fig.9 Meridional fans, base-ray (c)

——— traced

— — — second order basal

········· third order basal

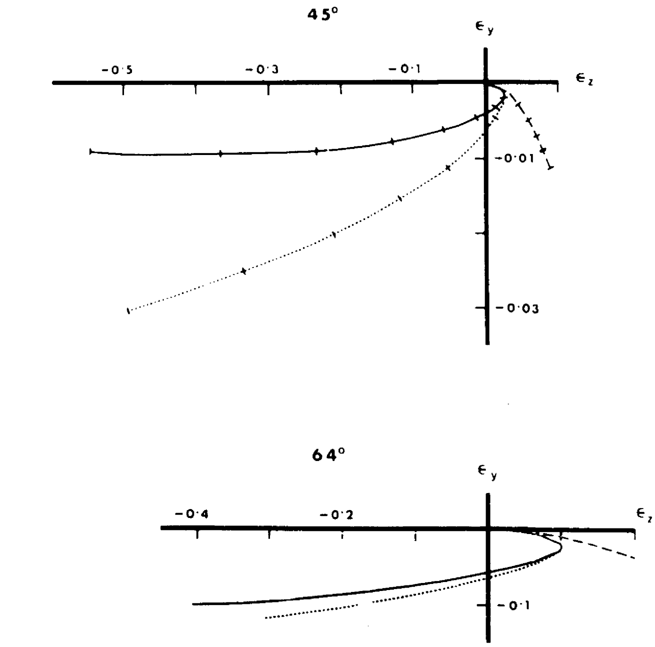


Fig.10 Sagittal fans, base-rays (a), (b).

traced
second order basal
third order basal

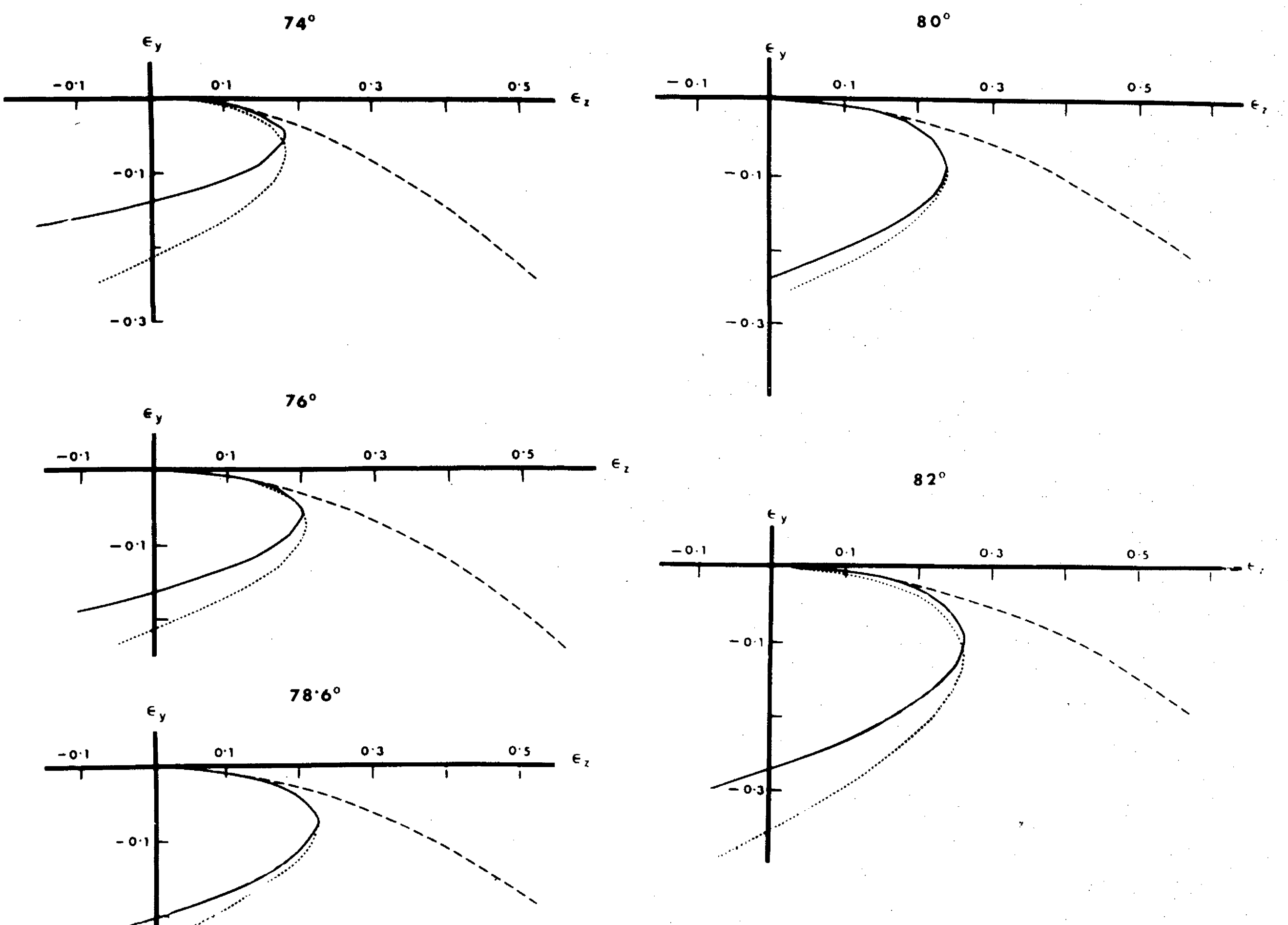
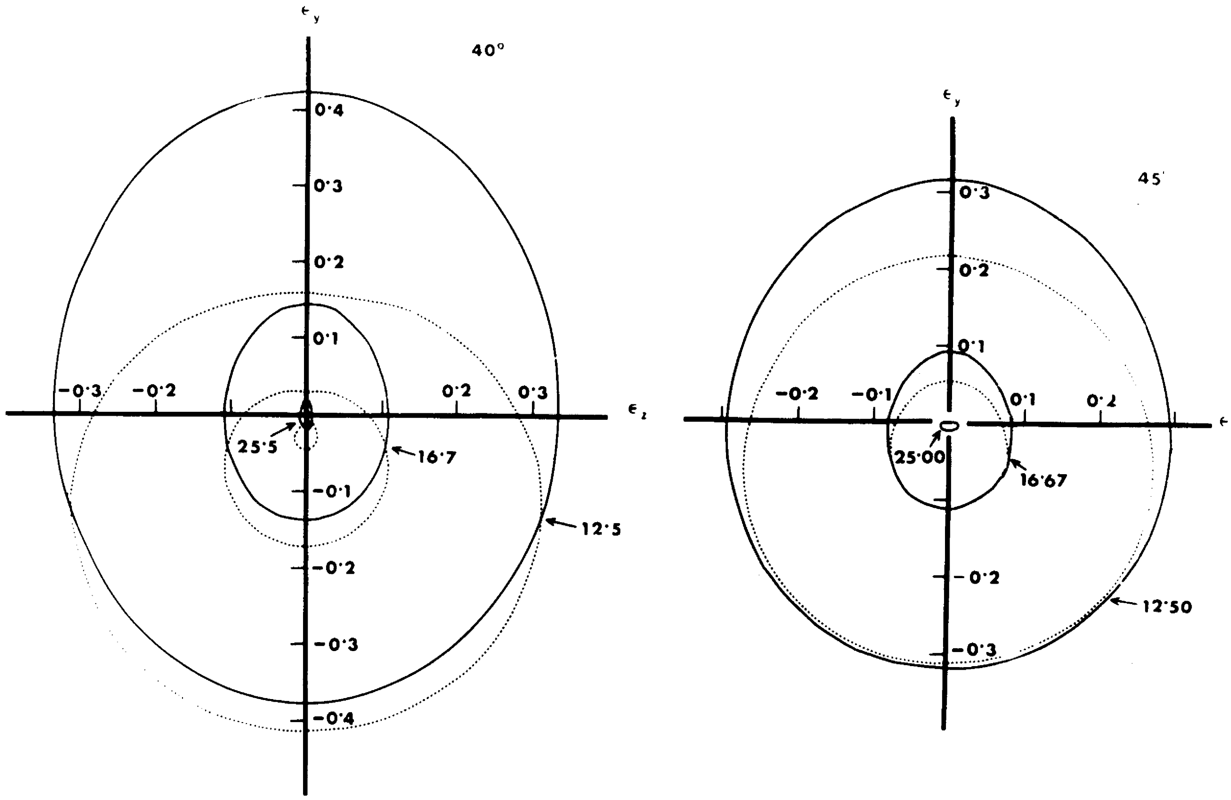
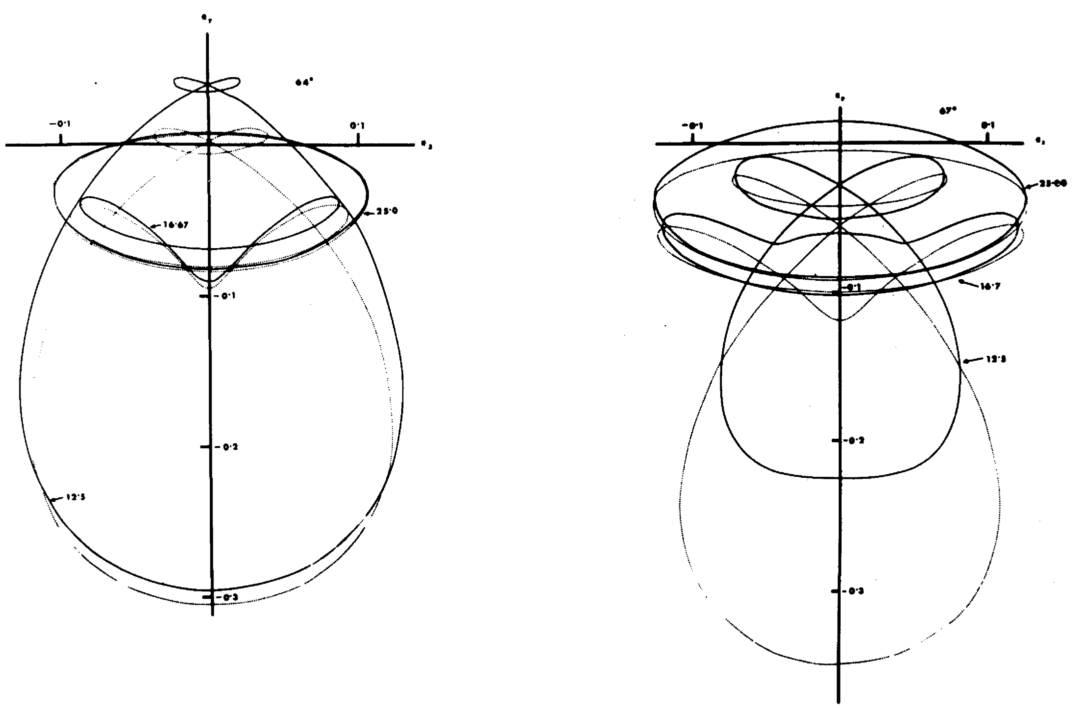


Fig.11

Sagittal fans, base-ray (c).



<u>Fig. 12</u> Aperture curves, base-ray (a)



<u>Fig. 13</u> Aperture curves, base-ray (b)

—— traced, ... third order basal predictions

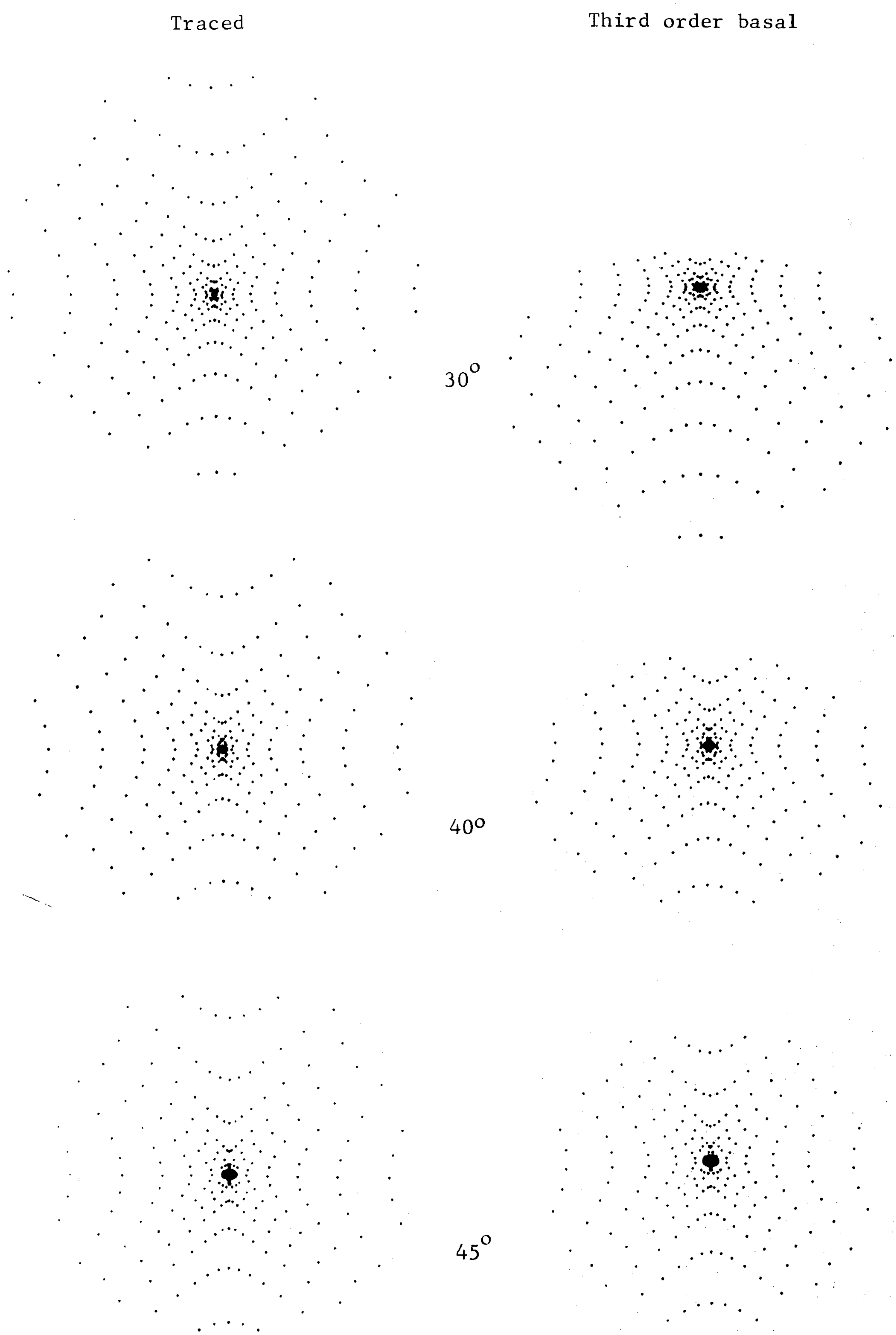


Fig. 14(i) Spot diagrams in ideal image plane. Base-ray (a).

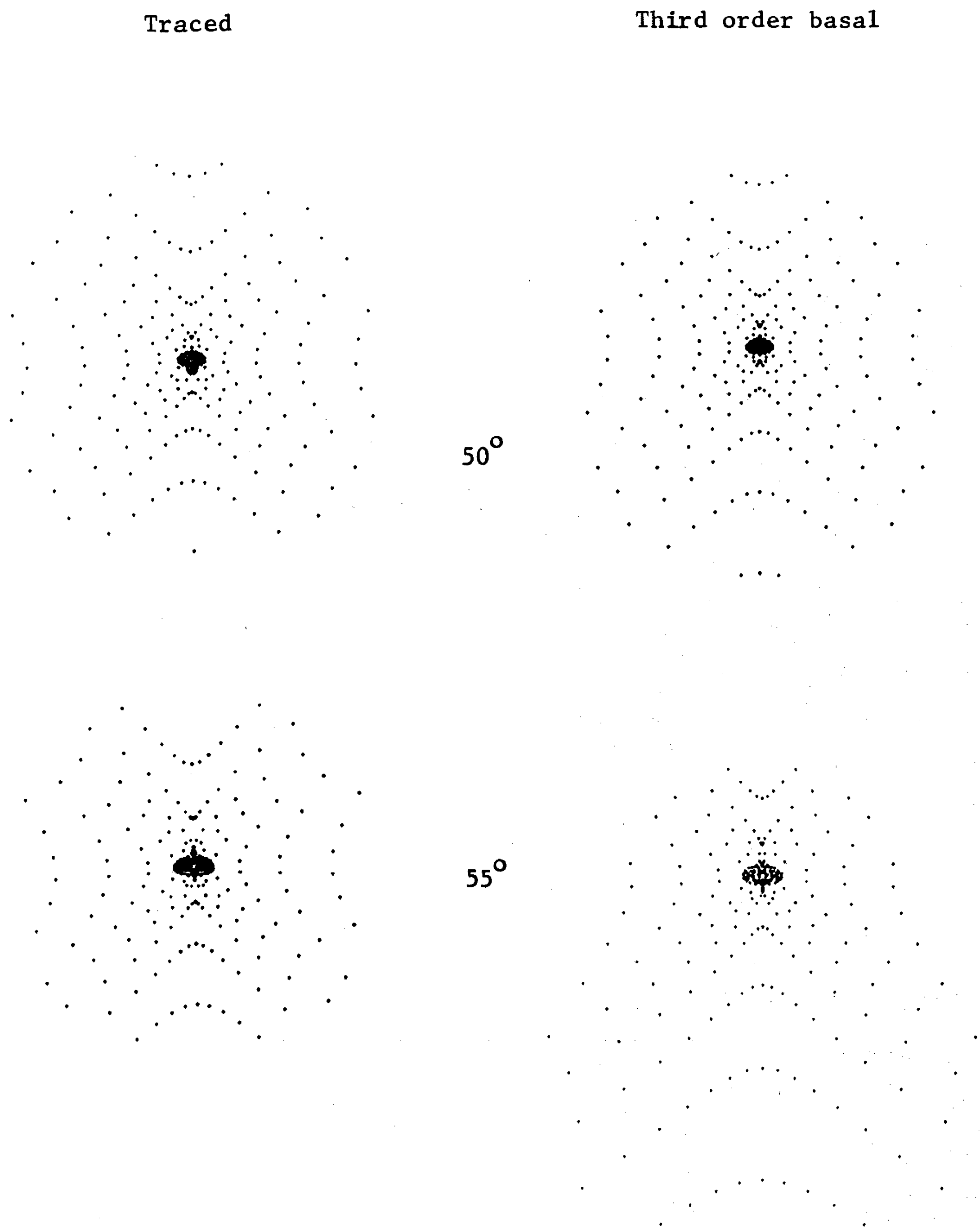
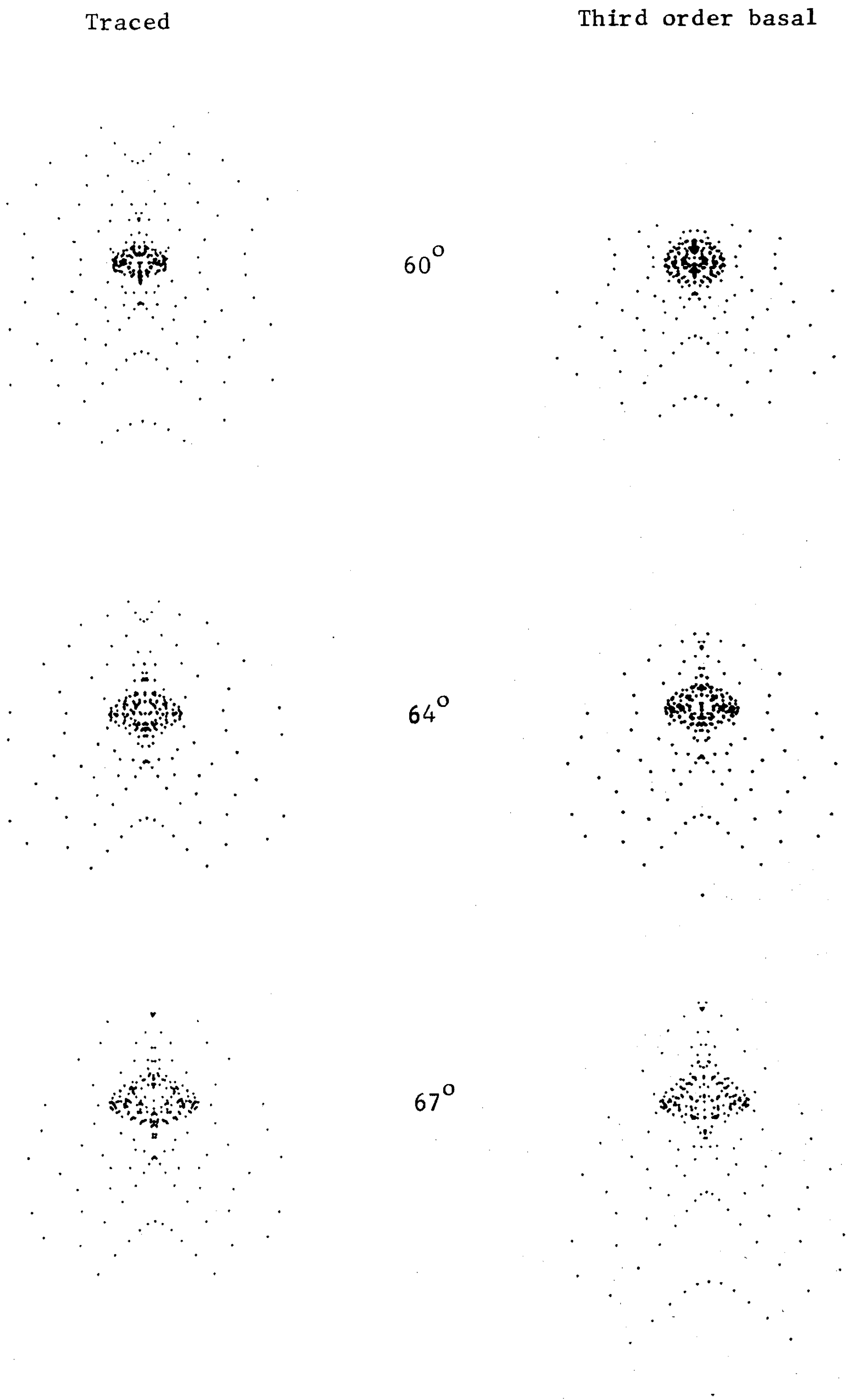
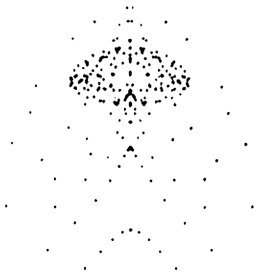


Fig. 14(ii) Spot diagrams in ideal image plane. Base-ray (a).



<u>Fig.15(i)</u> Spot diagrams in ideal image plane. Base-ray (b).

Traced | Third order basal



70°

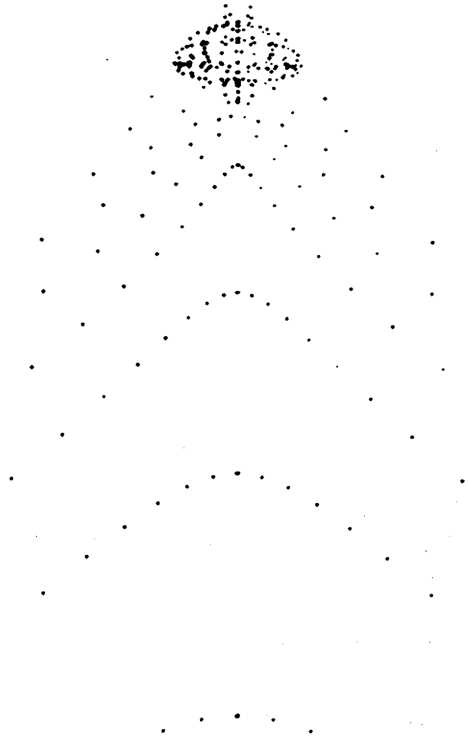
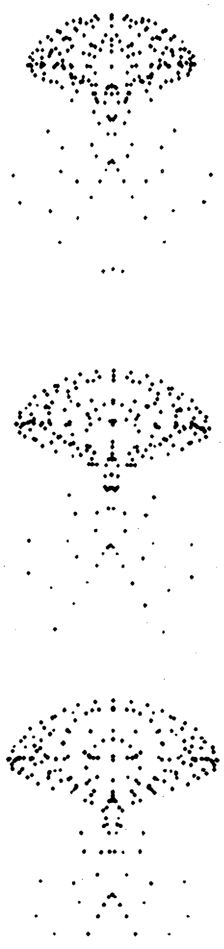


Fig.15(ii) Spot diagrams in ideal image plane. Base-ray (b).



76°

78.6°

80°

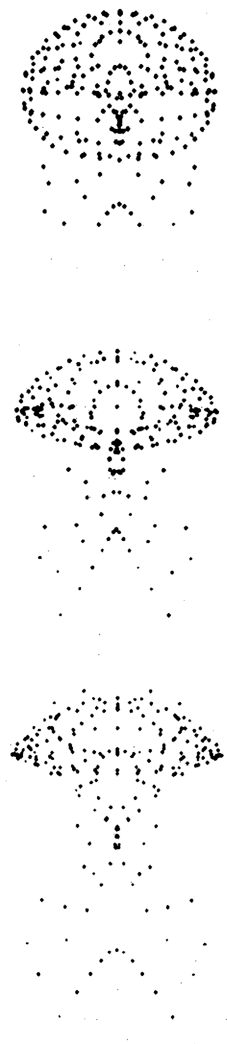


Fig.16 Spot diagrams in ideal image plane. Base-ray (c).

Traced | Basal prediction

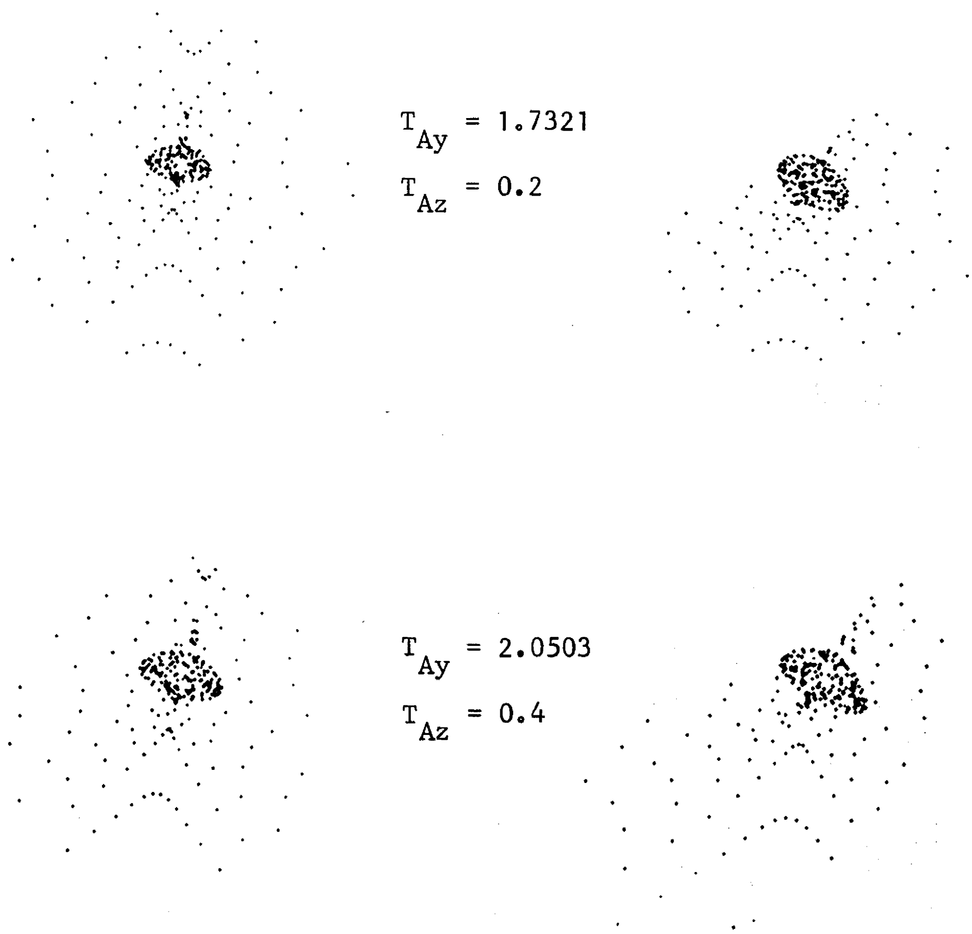


Fig.17 Spot diagrams for two skew pencils. Base-ray (b).

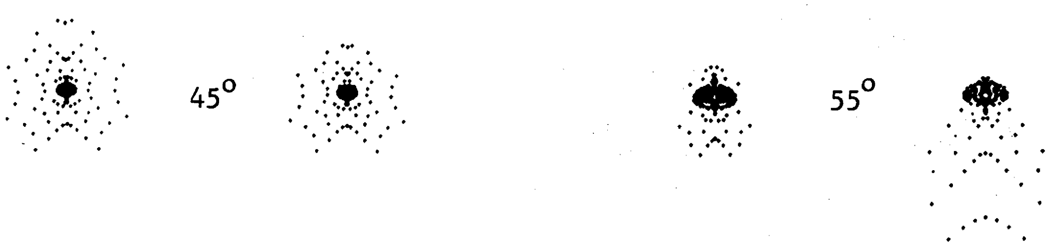


Fig.18 Spot diagrams for aperture of f/14. Base-ray (a).

(Third order basal prediction on the right of each pair.)

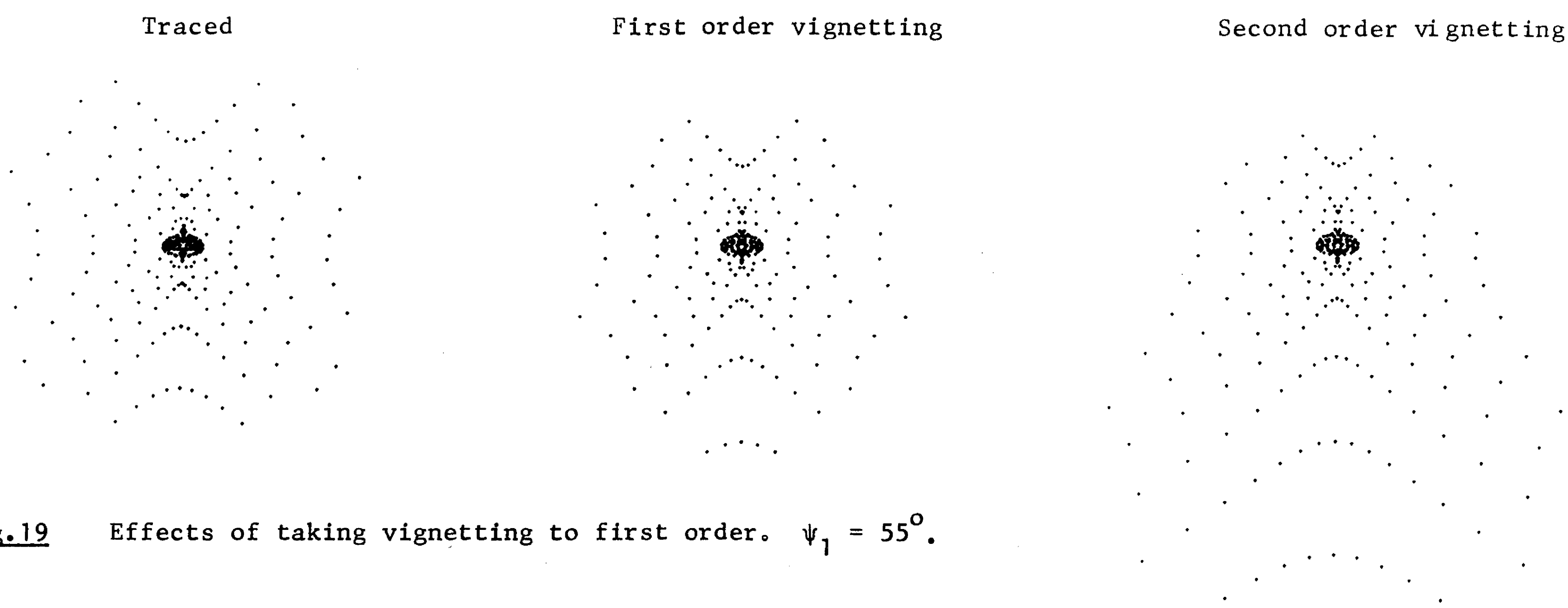


Fig.19 Effects of taking vignetting to first order. $\psi_1 = 55^o$.

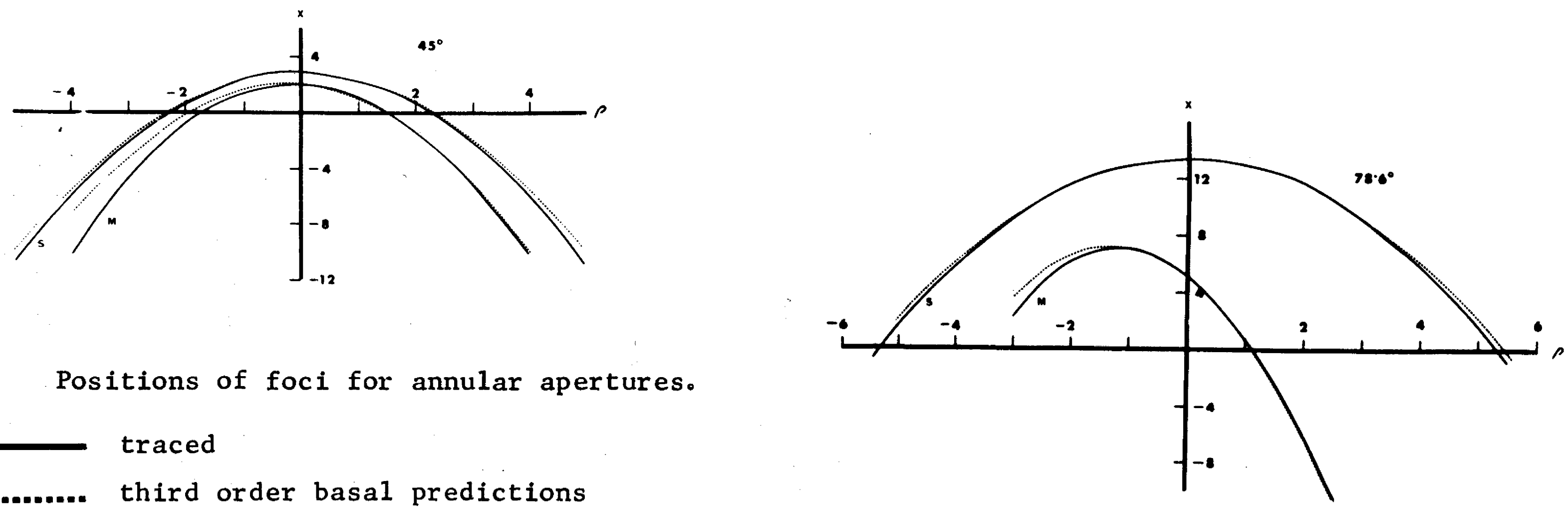


Fig.20 Positions of foci for annular apertures.

—— traced

.... third order basal predictions

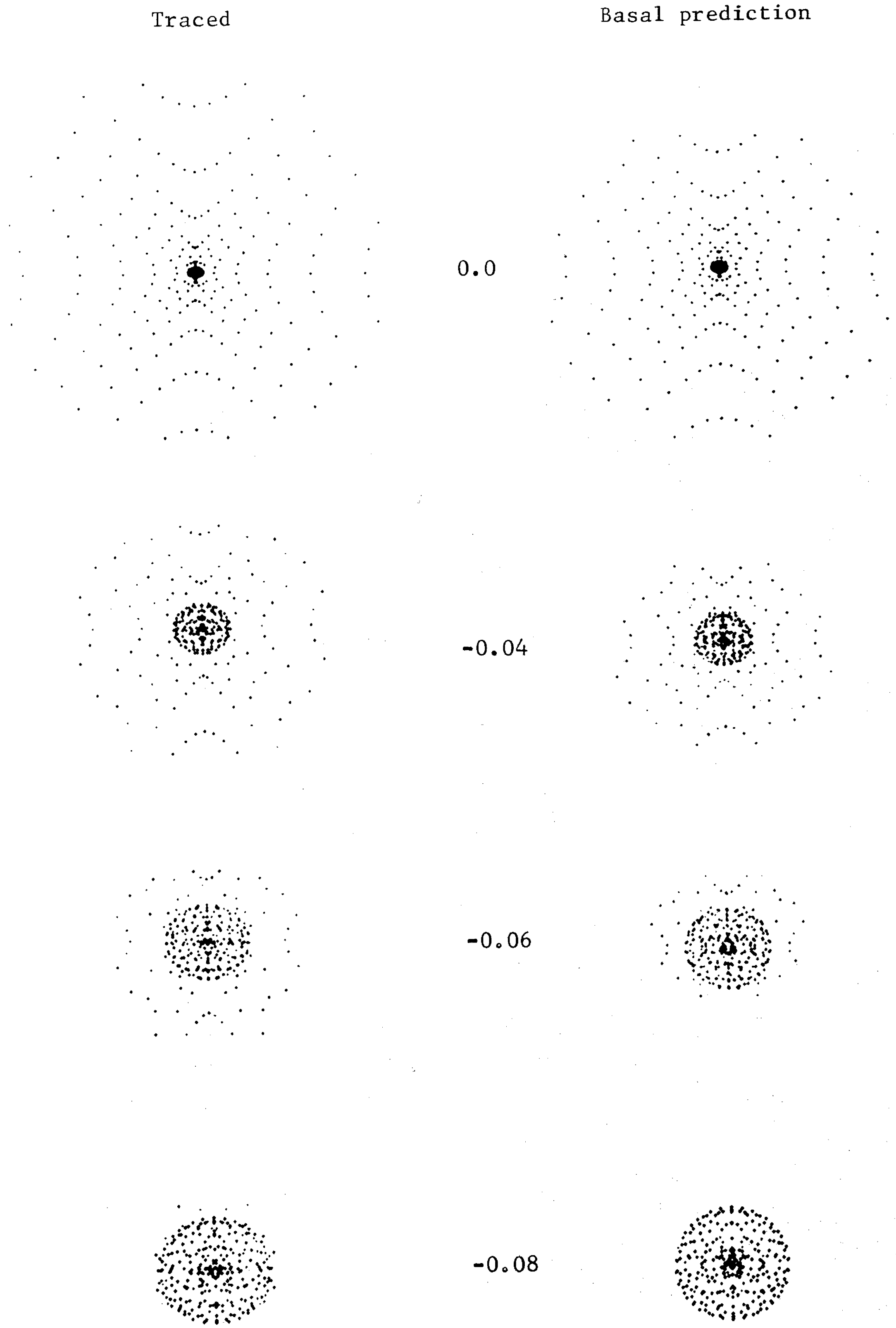
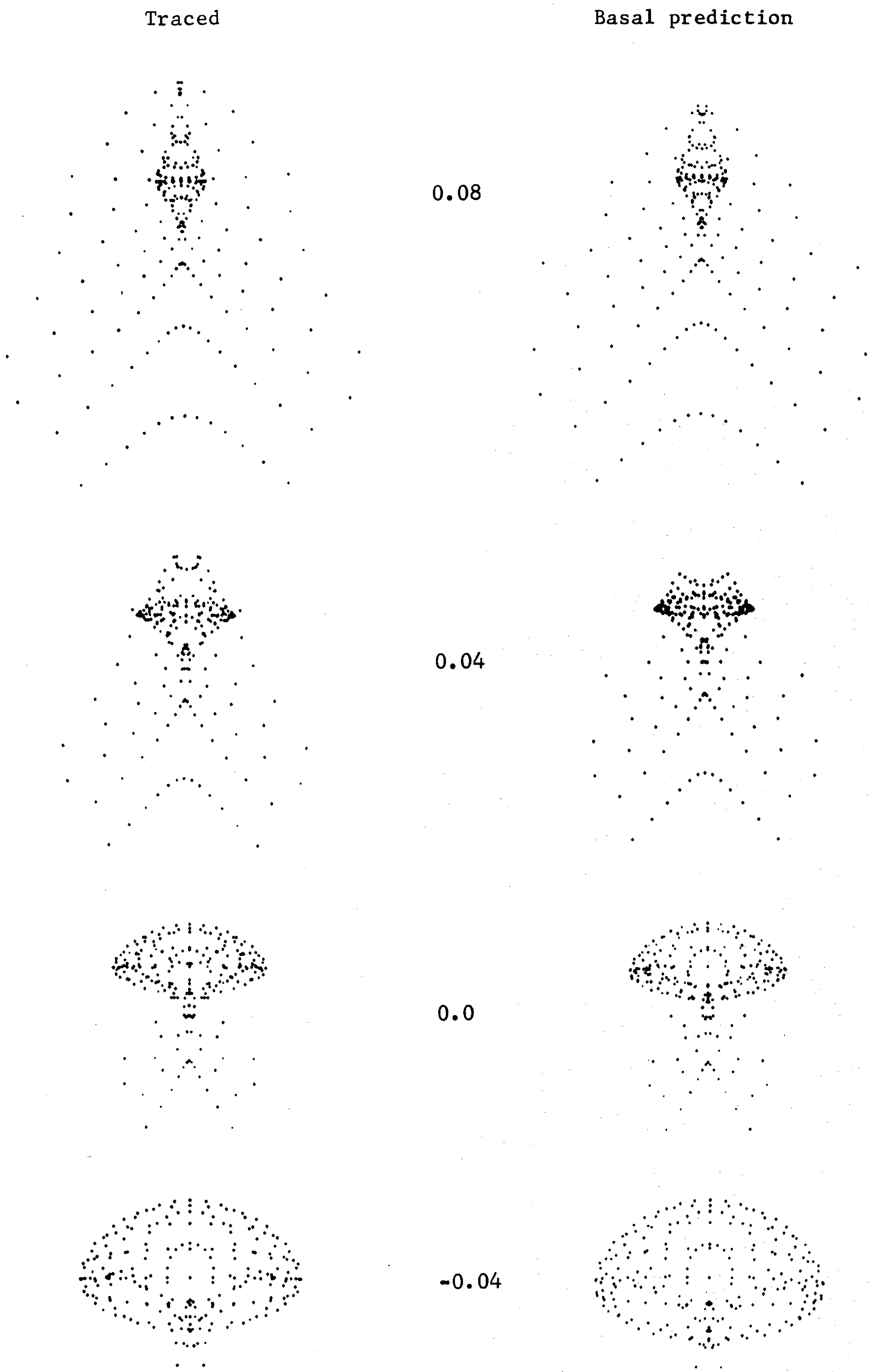


Fig.21 Spot diagrams in displaced image plane. $\psi_1 = 45^o$. Base-ray (a).



<u>Fig.22</u> Spot diagrams in displaced image plane. $\psi_1 = 78.6^{\circ}$. Base-ray (c).

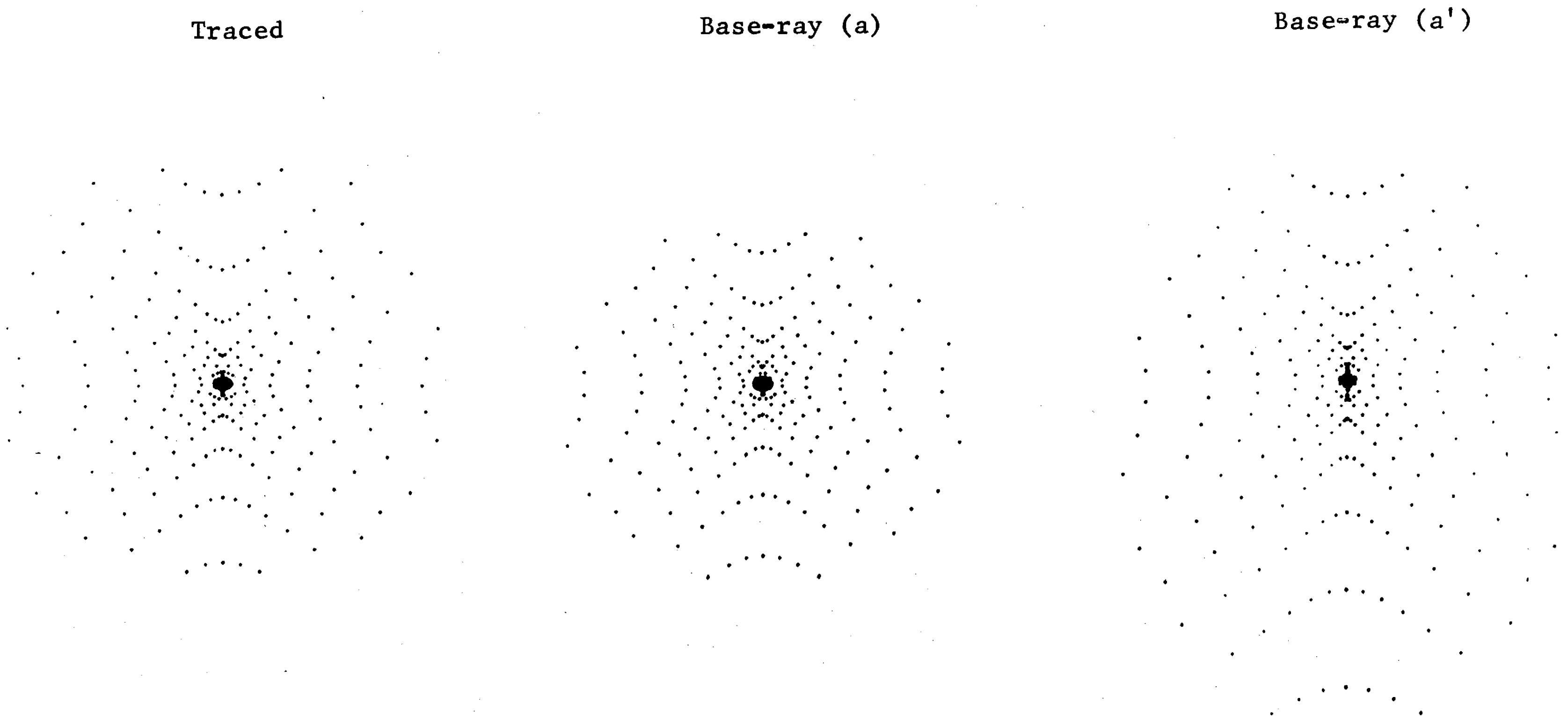


Fig.23 Comparison of spot diagrams predicted with base-rays (a) and (a').

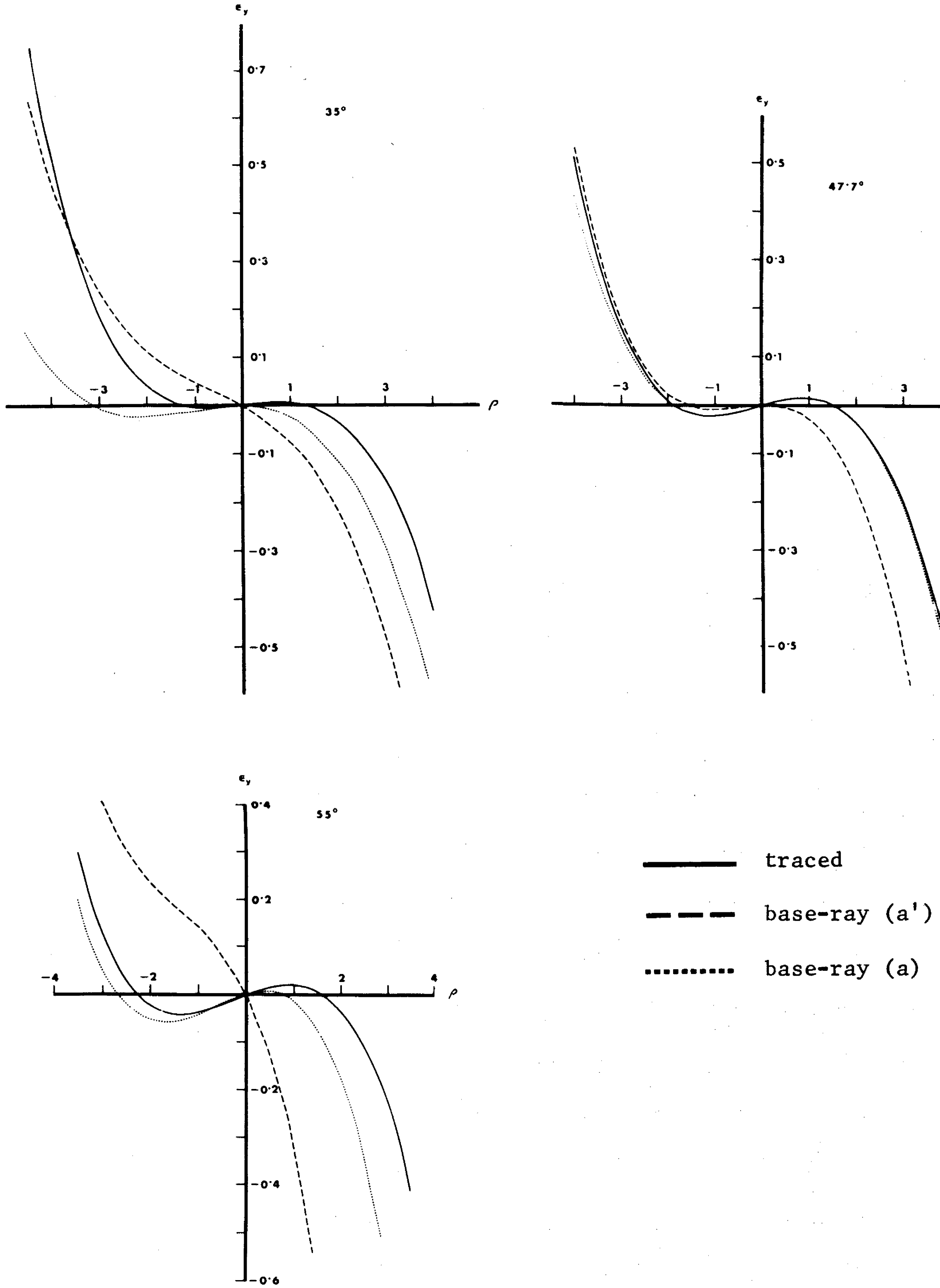


Fig.24 Comparison of meridional fans predicted with base-rays (a) and (a').

APPENDIX B: ON THE CHOICE OF COORDINATES

83. Introduction of Hamiltonian and Normal Coordinates

It was not at all obvious at first sight that translated coordinates were the most suitable coordinates for use in the basal theory. For instance, the physically meaningful $\underset{\sim}{V}_A$ did not vanish for the base-ray. This made it necessary to introduce the quantity $\underset{\sim}{V}$ which had physical meaning only through the addition of the zero order quantity $\underset{\sim}{V}_B$ (i.e., through (5.2)). The presence of this additive constant $\underset{\sim}{V}_B$ (and of other non-zero pseudo-parameters) is essentially the source of the algebraic complications inherent in the use of translated coordinates. Clearly, to ensure that $\underset{\sim}{V}_A$ vanishes for the base-ray, the x-axes of the coordinate systems must lie along the base-ray. This is the motivation for the introduction and analysis of hamiltonian coordinates. With hamiltonian coordinates two coordinate systems Ⓒ$_H$ and Ⓒ$_H$' are associated with each surface Ⓕ. Their common origin is Ⓟ$_B$ and their x-axes lie along the base-ray in the object and image spaces of Ⓕ respectively. In both cases the y-axis is in Ⓜ. Ⓒ$_H$ and Ⓒ$_H$' are the hamiltonian coordinate systems associated with the surface Ⓕ. Note that hamiltonian coordinates are equivalent to the coordinates usually used in discussions involving hamiltonian optics.

Another possible source of complication in the use of translated coordinates is that a ray incident along the x-axis is in general refracted away from the x-axis. This can be avoided only by the use of normal coordinates defined so that a single coordinate system Ⓒ$_N$ is associated with each surface and the origin of Ⓒ$_N$ is Ⓟ$_B$ with the x-axis along the normal to the surface at Ⓟ$_B$. As usual, the y-axis is in Ⓜ. Ⓒ$_N$ is the normal coordinate system associated with Ⓕ. The use of normal coordinates is of some advantage in the specification of aspheric surfaces in the basal theory (see §93). However, for the purposes of this appendix aspherical surfaces will be overlooked. Unless the coordinate systems associated with the image space of a surface and the object space of the next surface are related by a translation, transfer between these two surfaces is a relatively complicated matter (see §85). On the other hand, the association of two coordinate systems with each surface also introduces complications (see §84). In making the final assessment of the relative merits of the various coordinate systems, account was taken of both the theoretical and the practical aspects of the basal theory.

The general structure of the basal theory is the same in hamiltonian and normal coordinates as in translated coordinates; only the details differ. In particular the quasi-invariants are formally identical and (25.6) holds between the appropriate canonical and paracanonical variables. Consequently the principles of iteration are the same. The differences in detail particularly manifest themselves in the quadratic equation for R and the equations for $\Delta\underset{\sim}{G}$ as functions of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ and, in the case of normal coordinates, in the transfer equations.

84. Hamiltonian Coordinates

$Ⓒ_H$ and $Ⓒ_H'$ are related by a rotation of coordinates about the z-axis. If $\vec{y}$ is some vector in $Ⓒ_H$, the same vector specified in $Ⓒ_H'$ is $\vec{y}''$ where

$$\vec{y}'' = T(\phi)\vec{y} \quad , \tag{84.1}$$

and

$$T(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \tag{84.2}$$

with

$$\phi = \Delta I_B \quad .$$

The convention relating to the use of " is as follows: if Q is some quantity defined with respect to Ⓒ$_H$ for the ray Ⓡ before refraction at Ⓕ, then Q" is the quantity defined for the same ray Ⓡ and in the same way as Q but with respect to C_H' before refraction. Q" is not to be confused with Q' which is the analogous quantity defined after refraction. If and only if Q is a component of some three-vector $\vec{Q}$, will Q" be given by the same component of $\vec{Q}''$ according to (84.1). Rays are specified by their canonical variables $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ defined with respect to Ⓒ$_H$ as follows: $\underset{\sim}{Y}$ is the point of intersection of Ⓡ with the x-plane of $\underset{\sim}{C}_H$ and if the direction cosines of Ⓡ are α, β and γ in Ⓒ$_H$, $\underset{\sim}{V}$ is

$$\underset{\sim}{V} = \underset{\sim}{\beta}/\alpha \quad . \tag{84.3}$$

In the image space of Ⓕ, $\underset{\sim}{Y}'$ and $\underset{\sim}{V}'$ are defined analogously with respect to Ⓒ$_H$'. Note that $\underset{\sim}{V}$ and $\underset{\sim}{V}'$ vanish for the base-ray.

In hamiltonian coordinates the law of refraction (8.3) is

$$\vec{\beta}' - k\vec{\beta}'' = \bar{R}_A \alpha' \vec{n}' \quad , \tag{84.4}$$

where

$$\vec{n}' = \vec{n}'' = T(\phi)\vec{n} \quad , \quad \vec{\beta}'' = T(\phi)\vec{\beta} \quad .$$

Thus, analogous to (9.2,3),

$$\underset{\sim}{V}' - \underset{\sim}{V}'' = -\bar{R}_A \underset{\sim}{I}_A'' \quad , \tag{84.5}$$

where the definitions (9.3,5) for I_{Ax}, $\underset{\sim}{I}_A$ are carried over:

$$\underset{\sim}{I}_A = n_x \underset{\sim}{V} - \underset{\sim}{n} \quad .$$

(Thus $\underset{\sim}{I}_A'' = n_x''\underset{\sim}{V}'' - \underset{\sim}{n}''$ where $\underset{\sim}{V}'' = \underset{\sim}{\beta}''/\alpha''$, see above.) $\bar{R}_A$ will satisfy the equation (9.7,8) provided $\underset{\sim}{I}_A$, $\underset{\sim}{V}_A$ and n_x are replaced by $\underset{\sim}{I}_A''$, $\underset{\sim}{V}''$ and n_x'' respectively. $\vec{n}''$ can be expressed in terms of $\underset{\sim}{Y}''$ and $\underset{\sim}{V}''$

(84.5) expresses $\underset{\sim}{V}'$ in terms of $\underset{\sim}{V}''$ and $\underset{\sim}{Y}''$. This has two disadvantages. First, $\underset{\sim}{V}''$ does not vanish for the base-ray and second, $\underset{\sim}{V}'$ is required in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. The first disadvantage may be modified by redefining $\underset{\sim}{V}''$ as $\underset{\sim}{\beta}''/\alpha'' - \underset{\sim}{\beta}/\alpha$. This is in effect associating a single coordinate system $Ⓒ_H'$ with each surface. Consequently the disadvantages of the use of translated coordinates are introduced. Moreover, the $Ⓒ_H'$ associated with successive surfaces are related by rotations and the disadvantages of using normal coordinates are also introduced (see §85). Laying aside these facts for the moment, it is evident that $\underset{\sim}{V}''$ and $\underset{\sim}{Y}''$ can be expressed as a series in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. Thus $\bar{R}_A$, $\underset{\sim}{Y}'$ and $\underset{\sim}{V}'$ etc. could be first expanded in terms of $\underset{\sim}{Y}''$ and $\underset{\sim}{V}''$ and then these expansions converted to expansions in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. This clearly will involve a considerable amount of additional computation when computing the coefficients.

Both of these features can be removed by expressing the variables to which " is attached in terms of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ before any expansions are made. Since n_x and $\underset{\sim}{n}$ are the components of the vector $\vec{n}$, (84.1) may be applied in this case. It was seen in §9a that $(-\alpha I_{Ax}, -\alpha I_{Az}, \alpha I_{Ay})$ were the components of $\vec{n} \times \vec{\beta}$. This is still the case in hamiltonian coordinates. Thus (84.1) is applicable to αI_{Ax} , $\alpha \underset{\sim}{I}_A$. If this is done it is found that the quadratic equation for $\bar{R}_A$ in terms of the variables $\underset{\sim}{Y}''$, $\underset{\sim}{V}''$ becomes

$$\tau \bar{R}_A^2 + 2\sigma \bar{R}_A + \mu = 0 \quad ,$$

where

$$\mu = (k^2-1)(1+\underset{\sim}{V}.\underset{\sim}{V})$$

$$\sigma = -k^2[\underset{\sim}{I}_A.\underset{\sim}{V} \cos\phi - (I_{Ay}+WI_{Ax})\sin\phi] + (1+\underset{\sim}{V}.\underset{\sim}{V})(n_y \sin\phi + n_x \cos\phi) \quad .$$

$$\tau = k^2[I_{Ay}^2 + (I_{Az}\cos\phi - I_{Ax}\sin\phi)^2] - (1+\underset{\sim}{V}.\underset{\sim}{V})(n_y\sin\phi + n_x\cos\phi)^2 \quad .$$

This is considerably more complicated than (9.8) especially when it is noted that the first term in σ is cubic in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ and the first term in τ is of the fourth power in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. A similar process applied to (9.2) yields the following equations:

$$V' = -(\bar{R}_A I_{Ay} - V\cos\phi + \sin\phi)/(\cos\phi + V\sin\phi)$$

$$W' = -[\bar{R}_A(I_{Az}\cos\phi - I_{Ax}\sin\phi) - W]/(\cos\phi + V\sin\phi) \quad . \qquad (84.5)$$

These are to be compared with the relatively simple (10.6). The corresponding equations for Y' and Z' are also complicated by comparison with (10.7). The consequences of this are that the analogues of (26.1,2) in hamiltonian coordinates are considerably more complicated than in translated (or normal) coordinates. Note that if ϕ is zero the equations of translated coordinates are recovered (in which case $\underset{\sim}{V}$ is in reality $\underset{\sim}{V}_A$). This completes the investigation of hamiltonian coordinates.

85. Normal Coordinates

As far as refraction is concerned, the only difference between the treatment based on the use of normal coordinates and that based on the use of translated coordinates is that the normal $\vec{n}_B$ lies along the x-axis and some of the pseudo-parameters take special values:

$$\vec{n}_B = (1,0,0) \quad , \quad \alpha_B = \cos I_B \quad , \quad \beta_B = \sin I_B \quad .$$

If these are substituted into those equations, relevant to refraction, which were previously derived using translated coordinates, certain minor simplifications occur, especially in the solutions to the equations for X and R. However, little advantage accrues over the use of translated coordinates.

Now consider the transfer from the image space of Ⓕ to the object space of $Ⓕ_+$. The origin of the coordinates $Ⓒ_+$ is the point $(d', d'V_B', 0)$ in Ⓒ where $d' = d_B'\cos I_B'$. In $Ⓒ_+$ the direction cosines of the base-ray are

$$\vec{\beta}_{B+} = (\cos I_{B+} , \sin I_{B+} , 0)$$

and in ©,

$$\vec{\beta}_B{}' = (\cos I_B{}' , \sin I_B{}' , 0) \quad .$$

Since in general $I_{B+} \neq I_B{}'$, a rotation of coordinates is involved in transfer. This is where the differences between normal and translated coordinates arise and is the source of the added complications associated with the use of normal coordinates. The required transfer equations are most readily obtained from (5.7,8) by replacing (due to the change in origin) $\underset{\sim}{Y}$ by $\underset{\sim}{Y}'+d'\underset{\sim}{V}'$ and $\underset{\sim}{V}$ by $\underset{\sim}{V}'$. Thus:

$$Y_+ = (\cos I_B{}'/\cos I_{B+})\chi(Y'+d'V') \qquad Z_+ = \chi[Z'+d'W'+(\cos I_B{}'/\cos I_{B+})b(Z'V'-Y'W')]$$

$$V_+ = (\cos I_B{}'/\cos I_{B+})^2\chi V' \qquad W_+ = (\cos I_B{}'/\cos I_{B+})\chi W' \quad , \qquad (85.1)$$

$$\chi = [1+(\cos I_B{}'/\cos I_{B+})bV']^{-1} \quad . \qquad (85.2)$$

In the parabasal region these reduce to

$$y_+ = (\cos I_B{}'/\cos I_{B+})(y'+d'v') \qquad z_+ = z'+d'w'$$

$$v_+ = (\cos I_B{}'/\cos I_{B+})^2 v' \qquad w_+ = (\cos I_B{}'/\cos I_{B+})w' . \qquad (85.3)$$

The quasi-invariants are given by

$$G_y = N_y(y_M V - v_M Y) \quad , \quad G_z = N_z(z_S W - w_S Z) \quad ,$$

where

$$N_y = N \cos^3 I_B \quad , \quad N_z = N \cos I_B \quad .$$

From (85.1-3) it follows that the transfer increments to the quasi-invariants are <u>not</u> zero but have the values

$$\nabla G_y = N_y{}'(\cos I_B{}'/\cos I_{B+})b\chi V'(v_M{}'Y'-y_M{}'V') \quad ,$$

$$\nabla G_z = N_z{}'(\cos I_B{}'/\cos I_{B+})b\chi W'(w_S{}'Y'-z_S{}'V') \quad . \qquad (85.4)$$

(Note: In the parabasal region these surface increments are zero.) (25.1) will no longer hold but must be modified by including the sum over the transfer increments. Hence the δ_y and δ_v will include contributions due to transfer between surfaces. It is evident that the final aberration coefficients will not be simply the sum of the surface contributions but will include transfer contributions. These are not optical in nature since they arise from geometrical considerations rather than from a refraction or reflection. (Compare this with the alteration in the aberration coefficients due to a rotation of the image plane. In that case the final coefficients could not be expressed as a sum of contributions by the individual surfaces. In the present case the surface contributions are supplemented by the transfer contributions. In both cases the application of the methods of design and correction based on surface contributions is not possible.)

86. Principal Reasons for the Choice of Translated Coordinates

From the results of the preceding two sections it is evident that the complications inherent in the use of hamiltonian or normal coordinates over and above those already present in the use of translated coordinates arise in both cases from the necessity of having to perform rotations of coordinates - in the case of hamiltonian coordinates between the object and image space of the same surface, in the case of translated coordinates between the image space of one surface and the object space of the next. As compared to the corresponding computations employing translated coordinates, the computation of the surface contributions to the paracanonical coefficients of the quasi-invariants is considerably more complicated in the case of hamiltonian coordinates and a little simpler in the case of normal coordinates. However, with normal coordinates the transfer increments to these coefficients must also be computed. While the expansion of (85.4) in terms of $\underset{\sim}{Y}'$ and $\underset{\sim}{V}'$ is a simple affair, it is necessary to convert this into an expansion in terms of $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$ (or in terms of the paracanonical coordinates if they are defined in this case). This essentially involves the use of pseudo-expansions and iteration and as such is quite a lengthy process from the point of view of computation. From the practical aspect of the computation of the $\underset{\sim}{G}$-coefficients alone, it is apparently simpler to use translated coordinates.

If hamiltonian or normal coordinates are used, the directions of the x-axes of the coordinate systems in the object and image spaces of the system as a whole are determined by the base-ray and in fact will be different for different base-rays. This is not the case if translated coordinates are used since, although the same in the object and image space, the direction of the x-axes can be fixed independently of the choice of base-ray. This has implications on the relationship between the aberration coefficients and the $\underset{\sim}{G}$-coefficients in each of the three coordinates. In the object space it is immaterial whether or not the object surface is plane and normal to the x-axis since it is possible to use non-linear paracanonical coordinates. However, in the image space the aberration or image height, as the case may be, is expressed simply in terms of the quasi-invariants only if the image plane is normal to the x-axis (see §51b). In general this will not be the case with hamiltonian or normal coordinates whilst for translated coordinates it may always be arranged to be the case (simply by choosing a pseudo-axis normal to the image plane). Thus the power of the surface contribution for the design and correction of systems is available only if translated coordinates are used.

APPENDIX C: RAY TRACING THROUGH SYSTEMS OF ARBITRARY SURFACES

87. Comments on Ray Tracing. Specification of Surfaces.

(a) Before the basal coefficients of any system can be computed a base-ray must be accurately traced through the system. If the system is symmetric and the surfaces are exclusively spherical, the ray trace scheme of Ford[1] is ideally suited for this purpose since many of the pseudo-parameters required to compute the basal coefficients are automatically computed in the course of the trace. However, this scheme is inapplicable to rays of half-field greater than 90° [1] and for decentred surfaces. If the surfaces are aspheric (see also Appendix D) and the system is rotationally symmetric, Ford's modified scheme[2] could be used with the same advantages.

A new scheme is presented below for use with decentred surfaces and aspheric surfaces which do not possess rotational symmetry. The equations of §9 are the basis of the scheme. The scheme differs from those of Ford in that no use can be made of the existence of quasi-invariants for the system (these presuppose a base-ray). If the new scheme is used for tracing the base-ray through an arbitrary system, many of the pseudo-parameters are again computed automatically in the course of the trace. Algebraic schemes do exist for tracing rays through arbitrary systems of rotationally symmetric aspheric surfaces[3], but these

---

[1] Ford (1960). Ford's basic scheme can be modified for rays of half-field greater than 90°. The modifications stem from the fact that at the first surface $R_A = -1$, $X_A = 2r$ in the correct paraxial limit for such rays, not 1 and 0 as is the usual case.

[2] Ford (1966).

[3] Allen and Snyder (1952).

are inconvenient for use with the basal theory and the following scheme is more general.

(b) The system Ⓚ may be specified as follows. First, each surface Ⓕ will be given by an equation of the form

$$x = \chi(y,z) = \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \chi_{\mu}^{(n)} y^{n-\mu} z^{\mu} \quad , \qquad (87.1)$$

where x, y, z are the coordinates of points on Ⓕ referred to a local coordinate system $Ⓒ_j$ whose origin lies on Ⓕ. Naturally the series (87.1) must be terminated at some stage. Second, let $Ⓒ_0$ be some suitably chosen coordinate system and let the origins of the coordinate systems $Ⓒ_j$ be the points x_{0j}, y_{0j}, z_{0j} in $Ⓒ_0$. (The x-axis of $Ⓒ_0$ would correspond to the pseudo-axis of §1.) If the $Ⓒ_j$ are not related to $Ⓒ_0$ by a translation, then the direction cosines of their coordinate axes must also be specified. It is anticipated that the $Ⓒ_j$ will be chosen so that they are in fact obtained from $Ⓒ_0$ by a translation. If some surfaces are spherical or have some special form (e.g., ellipsoids) so that their equation may be given in closed form not involving powers higher than the second in x, y or z, it is of advantage to specify them in the closed form. The abscisse of the point of intersection of a ray with the surface can then be determined exactly as the solution of a quadratic equation (cf. below). In that case the modifications to the scheme are merely a matter of detail. The scheme consists of three stages: determination of the point of incidence of a ray with the surface, the refraction of the ray and finally, transfer to the next surface.

88. Points of Intersection of Rays with Aspheric Surfaces

Any ray Ⓡ is specified at each surface by its canonical variables $\underset{\sim}{Y}$ and $\underset{\sim}{V}_A$ defined with respect to Ⓒ as in §5a. (Note: These are not the basal canonical variables. Throughout this appendix, subscripts "A" are omitted from many quantities, for example, from $\underset{\sim}{V}$, $\bar{\underset{\sim}{R}}$ and $\underset{\sim}{I}$. Since no basal variables are used at all in this appendix, no confusion need arise.) Before refraction at Ⓕ, the equation of Ⓡ is

$$\underset{\sim}{y} = \underset{\sim}{Y} + x\underset{\sim}{V} \quad . \tag{88.1}$$

The point of intersection (X,y,z) of Ⓡ with Ⓕ is obtained by eliminating $\underset{\sim}{y}$ from (87.1) by means of (88.1) and solving the resulting equation in x to give X. Suppose all the coefficients $\chi_\mu^{(n)}$ are zero for $n > 2$. Then X is the solution of the quadratic equation

$$\rho_2 X^2 + 2\rho_1 X + \rho_0 = 0 \quad , \tag{88.2}$$

where

$$\rho_2 = \chi_0^{(2)}V^2 + \chi_1^{(2)}VW + \chi_2^{(2)}W^2 \quad ,$$

$$\rho_1 = \chi_0^{(2)}YV + \tfrac{1}{2}\chi_1^{(2)}(YV+ZW) + \chi_2^{(2)}ZW + \tfrac{1}{2}(\chi_0^{(1)}V+\chi_1^{(1)}W-1) \quad ,$$

$$\rho_0 = \chi_0^{(2)}Y^2 + \chi_1^{(2)}YZ + \chi_2^{(2)}Z^2 + \chi_0^{(1)}Y + \chi_1^{(1)}Z \quad .$$

The relevant solution of (88.2) is that which reduces to zero when $\underset{\sim}{Y} \to 0$. Now suppose that all the $\chi_\mu^{(n)}$ are not zero for $n > 2$. Let X_n be an approximate solution for X. If $X_n + \delta X$ is an improved solution, then (87.1), (88.1) yield

$$X_n + \delta X = \chi_n + \delta X\, \underset{\sim}{V}\cdot\chi_{,\underset{\sim}{y}n}$$

where

$$\chi_n = \chi(Y+X_nV\ ,\ Z+X_nW) \quad , \quad \chi_{,\underset{\sim}{y}n} = \left(\frac{\partial\chi}{\partial y}\ ,\ \frac{\partial\chi}{\partial z}\right) \quad ,$$

with the derivatives calculated at $x = X_n$. Hence

$$X_{n+1} = (\chi_n - X_n\,\underset{\sim}{V}\cdot\chi_{,\underset{\sim}{y}n})/(1-\underset{\sim}{V}\cdot\chi_{,\underset{\sim}{y}n}) \quad . \tag{88.3}$$

From (87.1)

$$\frac{\partial\chi}{\partial y} = \sum_{n=1}^{\infty}\sum_{\mu=0}^{n} (n-\mu)\chi_{\mu}^{(n)} y^{n-\mu-1} z^{\mu} \quad , \quad \frac{\partial\chi}{\partial z} = \sum_{n=1}^{\infty}\sum_{\mu=0}^{n} \mu\chi_{\mu}^{(n)} y^{n-\mu} z^{\mu-1} . \quad (88.4)$$

For a general surface the point of intersection of the ray $\underset{\sim}{Y}$, $\underset{\sim}{V}$ with Ⓕ would be determined as follows: an approximate solution for X is obtained from (88.2) and used as the initial value of an iterative process whereby X is determined to the desired accuracy. Using the previous approximate value for X compute $\underset{\sim}{y}$ from (88.1) and then χ and $\chi_{,\underset{\sim}{y}}$ from (87.1), (88.4). Substitute these into (88.3) to determine the improved value for X. If the difference $X_{n+1} - X_n$ is greater than some prescribed value, repeat the iteration using the new value of X (i.e., X_{n+1}) as the initial value. When $X_{n+1} - X_n$ is less than the prescribed value the iteration is terminated.

89. Refraction

The normal to Ⓕ at the point of incidence (X,y,z) has the components

$$\vec{n} = (1, -\chi_{,y}, -\chi_{,z})/\sqrt{1+\chi_{,\underset{\sim}{y}}\cdot\chi_{,\underset{\sim}{y}}} \quad , \quad (89.1)$$

where the derivatives are evaluated for $\underset{\sim}{y} = \underset{\sim}{Y} + X\underset{\sim}{V}$. Using the previously determined value of X the components of $\vec{n}$ can be found from (89.1). Prior to calculating the refraction increments $\Delta\underset{\sim}{V}$ and $\Delta\underset{\sim}{Y}$, $\bar{R}$ must be found. $\underset{\sim}{I}$ is determined from (9.3) by using the known values of $\underset{\sim}{V}$ and the calculated values of $\underset{\sim}{n}$ and n_x. The coefficients μ, σ and τ can then be determined from (9.8), and (9.7) solved for $\bar{R}$. The relevant solutions of (9.7) are

$$\bar{R} = [(\sigma^2-\mu\tau)-\sigma]/\tau \quad , \quad \bar{R} = -2\sigma/\tau \quad , \quad (89.2)$$

for lenses and mirrors (k = -1) respectively. (9.2,4) now yield the canonical variables after refraction:

$$\underset{\sim}{V}' = \underset{\sim}{V} - \bar{R}\underset{\sim}{I} \quad , \quad \underset{\sim}{Y}' = \underset{\sim}{Y} + \bar{R}X\underset{\sim}{I} \quad . \tag{89.3}$$

90. Transfer

Finally, it is necessary to transfer these variables to the next surface. Suppose that the coordinate systems Ⓒ$_j$ are related to Ⓒ$_0$ by a translation without rotation. Then

$$\underset{\sim}{V}_{j+1} = \underset{\sim}{V}_j \quad . \tag{90.1}$$

After refraction at Ⓕ$_j$, Ⓡ intersects the plane $x = x_{0j}$ (coordinates in Ⓒ$_0$) in the point $\underset{\sim}{y} = \underset{\sim}{Y}_j' + \underset{\sim}{y}_{0j}$. Thus, in Ⓒ$_0$, the equation of Ⓡ is

$$\underset{\sim}{y} = \underset{\sim}{Y}_j' + \underset{\sim}{y}_{0j} + (x - x_{0j})\underset{\sim}{V}_j' \quad .$$

$\underset{\sim}{Y}$ was defined as the point of intersection of the ray with the x-plane of the coordinate system Ⓒ, referred to the origin of Ⓒ. It follows that Ⓡ will intersect the plane $x = x_{0j+1}$ in the point $\underset{\sim}{Y}_{j+1} + \underset{\sim}{y}_{0j+1}$. Thus, from the above equation,

$$\underset{\sim}{Y}_{j+1} = \underset{\sim}{Y}_j' + (x_{0j+1} - x_{0j})\underset{\sim}{V}_j' + \underset{\sim}{y}_{0j} - \underset{\sim}{y}_{0j+1} \quad . \tag{90.2}$$

(90.1,2) enable the canonical variables of Ⓡ to be determined before refraction at Ⓕ$_{j+1}$. Suppose now that the Ⓒ$_j$ are rotated with respect to Ⓒ$_0$. Then $\underset{\sim}{V}_{j+1}$ must be determined from $\underset{\sim}{V}_j'$ by a rotation of coordinates. The x-plane of Ⓒ is now not the plane $x = x_{0j}$ in Ⓒ$_0$ but is inclined to the x-axis. $\underset{\sim}{Y}_{j+1}$ will be determined by referring the point of intersection of Ⓡ with the x-plane of Ⓒ$_{j+1}$ to the point $(x_{0j+1}, y_{0j+1}, z_{0j+1})$ and then performing a rotation of coordinates (see §5b).

In this manner the ray is traced through the system. In the image space of the system as a whole $\underset{\sim}{Y}_k'$ and $\underset{\sim}{V}_k'$ will be available for computing the intersection point of the ray with the image plane. Using this scheme image heights are determined rather than aberrations and it is not possible to determine the contributions by the individual surfaces to the final aberration, however that may be defined. This was possible in the two schemes of Ford simply because the optical axis was the actual path of a ray and this could be taken as a base-ray, in which case the quasi-invariants were known and available for use.

APPENDIX D: THE PARABASAL COEFFICIENTS FOR ASPHERIC SURFACES

91. Aspheric Surfaces and the Basal Theory

(a) Throughout the derivation of the equations required to compute the basal coefficients it was assumed that the surfaces were spherical. It is highly desirable that these be generalised to surfaces which are aspheric and plane symmetric. It is known that aspherical surfaces can be used to great advantage in symmetrical systems. Undoubtedly they will eventually be incorporated into systems of large field angle in order to remove objectionable aberrations, thereby increasing the speed of the system, or to adjust distortion to a certain prescribed value (see §75c). The generalised basal theory would then be a valuable tool in the design and analysis of such systems. Secondly, if a plane symmetric system is required for some purpose it is almost certain that the required properties could be obtained more economically through the use of aspherical surfaces than from the use of a strongly decentred system of spherical surfaces since many more degrees of freedom are available with aspherical surfaces. It has already been seen (Appendix C) how rays may be traced through completely general surfaces and the present appendix will derive the equations for computing the parabasal coefficients of plane symmetric aspheric surfaces and outline the procedure for computing the higher order coefficients of these surfaces.

(b) It is possible for plane symmetric systems to be constructed from aspheric surfaces so that there exists a straight line normal to all the surfaces. This line is analogous to the optical axis of symmetric systems since a ray along it will be undeviated by the system. This ray would be a logical choice for a base-ray, in which case considerable simplifications would ensue since many of the pseudo-parameters would take on simple values. It would also be possible (in principle) to derive algebraic expressions for the derivatives of the coefficients with respect to the proper parameters of the system and for the chromatic aberration coefficients. However, such a system is probably academic.

Naturally it would be possible to compute the coefficients for anamorphotic systems using base-rays in either plane of symmetry or along the common section of the two planes of symmetry. In the latter case it would of course be preferable to derive anew the equations for the coefficients taking account of the existence of the two planes of symmetry and of the fact that the base-ray is undeviated by the system. (Such systems have been considered by various authors,[1] for example, Smith has considered the first order coefficients and Wynne considered the primary or third order coefficients for a system of cylindrical surfaces.) However, for the present purpose Ⓡ$_B$ will be an arbitrary meridional ray in an explicitly plane symmetrical system.

(c) When the base-ray is traced through a system of aspheric surfaces the surfaces are expressed in the form (87.1). For use in the basal theory the equation of the surface must be given in the coordinate system whose origin is the point of intersection of the base-ray with the surface. This equation may be obtained from (87.1) as follows. Let the point of intersection of Ⓡ$_B$ with Ⓕ have the coordinates $\vec{y}_B$ in the coordinate system Ⓒ of §87b. Then, provided the coordinate system used in the basal theory is related to Ⓒ by a translation, replace $\vec{y}$ in (87.1) by $\vec{y} + \vec{y}_B$ and expand in powers of $\vec{y}$ and $\vec{z}$. There results an equation of the form (87.1) for x (in the basal coordinate system) in terms of $\underset{\sim}{y}$.

---

[1] Smith (1928); Wynne (1954).

The coefficients in this expansion express themselves as series in $\underset{\sim}{y}_B$. Since (87.1) is itself an infinite series these will be infinite series. However, if $\underline{m}$ terms of (87.1) were sufficient for tracing the base-ray, it will be assumed that the coefficients in (87.1) for $\underline{n} > \underline{m}$ are all zero. The basal $\chi_\mu^{(n)}$ will then be expressed as finite series of $\underline{m}$ or less terms in $\underset{\sim}{y}_B$.

Translated coordinates and canonical variables are defined as in §5. The analysis of §8,9 was valid for any surface and $\underset{\sim}{I}$ and I_x are defined as in the equation following (10.3). Since (10.5,6,7) depended only on (9.2,4) and the definitions of $\underset{\sim}{V}$ and $\underset{\sim}{I}$, they hold for general surfaces and will form the basis of development of the basal theory for aspheric surfaces.

92. The Parabasal Coefficients

The abscisse X of the point of incidence of a ray with the surface will have the formal expansion (31.2). If this is substituted into the equation

$$X = \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \chi_\mu^{(n)} [Y+X(V+V_B)]^{n-\mu}(Z+XW)^\mu$$

(where μ is even since the surface is plane symmetric), and the right hand member expanded, it is found that

$$\hat{x}^{(1)} = \chi_0^{(1)}(Y+\hat{x}^{(1)}V_B) \quad . \tag{92.1}$$

Thus

$$\hat{P}_{x1} = \chi_0^{(1)}/(1-V_B\chi_0^{(1)}) \quad , \quad \hat{P}_{x2} = 0 \quad .$$

The normal $\vec{n}$ to the surface is given by (89.1). At the point of intersection with Ⓕ of the ray whose canonical variables are $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, $\vec{n}$ will have the formal expansions

$$n_x = n_{Bx} + \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \hat{n}_{x\mu\nu\tau}{}^{(n)} Y^{n-\mu} V^{\mu-\nu} Z^{\nu-\tau} W^{\tau} \quad ,$$

$$n_y = n_{By} + \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \hat{n}_{y\mu\nu\tau}{}^{(n)} Y^{n-\mu} V^{\mu-\nu} Z^{\nu-\tau} W^{\tau} \quad ,$$

$$n_z = \sum_{n=1}^{\infty} \sum_{\mu\nu\tau}^{n} \hat{n}_{z\mu\nu\tau}{}^{(n)} Y^{n-\mu} V^{\mu-\nu} Z^{\nu-\tau} W^{\tau} \quad , \qquad (92.2)$$

where ν is even in n_x and n_y and odd in n_z . Express $\vec{n}$ as a power series in y and z . From (88.4)

$$\chi_{,y} = \chi_0^{(1)} + 2\chi_0^{(2)} y + 0(2) \quad , \quad \chi_{,z} = 2\chi_2^{(2)} z + 0(2) \quad .$$

Hence,

$$n_x = (1+\chi_{,\underset{\sim}{y}} \cdot \chi_{,\underset{\sim}{y}})^{-\frac{1}{2}} = n_{Bx} + 2n_{By} n_{Bx}{}^2 \chi_0^{(2)} y + 0(2) \quad ,$$

$$n_y = n_{By} - 2n_{Bx}{}^3 \chi_0^{(2)} y + 0(2) \quad ,$$

$$n_z = -2n_{Bx}\chi_2^{(2)} z + 0(2) \quad , \qquad (92.3)$$

where

$$n_{Bx} = (1+\chi_0^{(1)2})^{-\frac{1}{2}} \quad , \quad n_{By} = -\chi_0^{(1)} n_{Bx} \quad , \quad n_{Bz} = 0 \quad , \qquad (92.4)$$

are the components of $\vec{n}$ for the base-ray. From the second of these it is evident that $\hat{p}_{x1}$ has precisely the value it does for spherical surfaces:

$$\hat{p}_{x1} = -\alpha_B n_{By} / \cos I_B = -\lambda_1 n_{By} \quad .$$

(This is due to the fact that to the first order, X is determined as if Ⓕ were replaced by its tangent plane at $Ⓟ_B$.) Thus, by substituting in (92.3) for $\underset{\sim}{y}$ according to

$$\underset{\sim}{y} = \underset{\sim}{Y} + X(\underset{\sim}{V} + \underset{\sim}{V}_B) \quad ,$$

the first order coefficients in (92.2) are found to be

$$\hat{p}_{nx1} = 2\chi_0^{(2)} n_{By} n_{Bx}{}^2 (1+\hat{p}_{x1} V_B) = 2\chi_0^{(2)} n_{By} n_{Bx}{}^3 \lambda_1 \quad ,$$

$$\hat{p}_{ny1} = -2\chi_0^{(2)} n_{Bx}{}^3 (1+\hat{p}_{x1} V_B) = -2\chi_0^{(2)} n_{Bx}{}^4 \lambda_1 \quad ,$$

$$\hat{p}_{nz3} = -2\chi_2^{(2)} n_{Bx} \quad ,$$

$$\hat{p}_{nx2} = \hat{p}_{ny2} = \hat{p}_{nz4} = 0 \quad . \qquad (92.5)$$

Now

$$\underset{\sim}{I} = \underset{\sim}{I}_A - \underset{\sim}{I}_B = n_x\underset{\sim}{V} + (n_x - n_{Bx})\underset{\sim}{V}_B - \underset{\sim}{n} + \underset{\sim}{n}_B \quad , \tag{92.6}$$

and the first order $\hat{\underset{\sim}{i}}$-coefficients are

$$\hat{p}_{iy1} = 2\chi_0^{(2)} n_{Bx}^3 \quad , \quad \hat{p}_{iy2} = n_{Bx} \quad ,$$
$$\hat{p}_{iz3} = 2\chi_2^{(2)} n_{Bx} \quad , \quad \hat{p}_{iz4} = n_{Bx} \quad . \tag{92.7}$$

The higher order $\hat{i}$-coefficients are also readily obtained from (92.6). (13.2) still holds since $\hat{p}_{x1}$ is unchanged and (13.4) are again obtained. Consider the parabasal terms of $\underset{\sim}{I}$ for aspheric surfaces and compare these with (10.4). The difference between these is the appearance in the aspheric $\underset{\sim}{i}$ of $\hat{p}_{iy1}$ in place of c in i_y and of $\hat{p}_{iz3}$ in place of c in i_z. Consequently, as can be checked by examining the analysis leading to (13.9), the same changes are required in (13.9) in order that this set of equations be valid for aspheric surfaces. Hence the canonical variables of a parabasal ray after refraction are given in terms of those before refraction by the equations

$$y' = (\lambda_1/\lambda_1')y \quad , \quad v' = k_y(\lambda_1'/\lambda_1)v - (\hat{p}_{iy1}\bar{R}_B\lambda_1/\alpha_B'\cos I_B')y \quad ,$$
$$z' = z \quad , \quad w' = k_z w - \hat{p}_{iz3}\bar{R}_B z \quad . \tag{92.8}$$

Since the transfer equations are independent of the nature of the surface, the parabasal coefficients can be computed from (92.8) in the usual manner.

93. The Higher Order Coefficients

The modifications required to compute the higher order coefficients for aspheric surfaces will be considered in outline only. They centre on the fact that for aspheric surfaces $\underset{\sim}{I}$ is no longer linear in $\underset{\sim}{Y}$ and $\underset{\sim}{V}$. First, the $\hat{x}$-coefficients must be determined. This is a simple matter:

the higher order analogues of (92.1) may be iteratively obtained from the equation preceding (92.1), and the $\hat{\ell}$-coefficients are read off from these. Next the second and higher order coefficients in (92.2) must be determined. The general method is along the lines of the derivation of (92.3,5) and is tedious. Once the coefficients in (92.2) have been determined the $\hat{\imath}$-coefficients may be read off (92.6). (27.3,4) depend on the general equations (9.7,8) and the definitions of $\underset{\sim}{I}$ and $\underset{\sim}{V}$, and consequently also applie to aspheric surfaces. The equations corresponding to those of Table 32/1 may be obtained by substituting the series for $\underset{\sim}{I}$ and n_x into (27.4) and carrying out the necessary expansions. The $\hat{r}$-coefficients follow as before.

Neither the principles nor the details of iteration and pseudo-expansions are altered by the generalisation to aspheric surfaces. Consequently the r*- and p*-coefficients and the r- and p-coefficients may be determined in the same manner as for spherical surfaces. The $\underset{\sim}{i}$-coefficients (as distinct from the $\hat{\underset{\sim}{i}}$-coefficients) must be determined by substituting the series (25.6) for $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ (which remain valid for aspheric surfaces) into the surface expansion of $\underset{\sim}{I}$. This essentially involves the use of iteration and pseudo-expansions. However, since only the second order $\underset{\sim}{i}$-coefficients are required in order to determine the third order $\underset{\sim}{g}$-coefficients, the $\underset{\sim}{i}$-coefficients are in fact best obtained from the second order equations corresponding to (39.2) for two-vectors. The $\underset{\sim}{r}$- and $\underset{\sim}{p}$-coefficients are again obtained from Table 40/1,2 and the $\underset{\sim}{g}$-coefficients follow as usual from (41.1). A considerable amount of work remains to be done with respect to aspheric surfaces, especially an investigation into the most suitable manner in which to specify the surface. For example, it may be of advantage in some cases to specify the surface by

$$c^2\vec{y}.\vec{y} - 2c\vec{y}.\vec{n}_B = \phi(y,z) \quad ,$$

where $\phi(y,z)$ is an odd function of z. In this manner $\phi(y,z)$ represents the deviation of the surface from a sphere of curvature c.

APPENDIX E: THE PRINCIPLE OF DUALITY

94. Dual Transformations and Transforms

The paracanonical coordinates $\underset{\sim}{S}$ and $\underset{\sim}{T}$ were introduced into the theory in a symmetrical fashion - there is nothing in (6.1) to distinguish between $\underset{\sim}{S}$ and $\underset{\sim}{T}$. Consequently, for any equation occurring in the theory either $\underset{\sim}{S}$ and $\underset{\sim}{T}$ must appear symmetrically in the equation, or there must exist a second equation corresponding to the original so that where $\underset{\sim}{S}$ occurs in one, $\underset{\sim}{T}$ occurs in the other, and vice versa. This may be expressed as follows:

A formal interchange of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ transforms any valid equation into another valid equation.

(Compare the case of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ which do not appear symmetrically in the theory - one represents a point and the other a direction. An interchange of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ does not necessarily transform a valid equation into another valid equation. For example, the pair of equations (9.2,4) should be transformed into each other whereas in fact they are not.) A formal interchange of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ will be called a dual transformation and under a dual transformation any quantity Q (this includes vector and scalar quantities and coefficients but not the parameters of the system) is transformed into its dual (transform) $Q^{\#}$. If two successive dual transformations are made there will in effect have been no change made at all. Thus the dual of the dual of a quantity is the original quantity. A quantity which satisfies

$$Q^{\#} = Q \quad , \tag{94.1}$$

is said to be self-dual. Note that

$$\underset{\sim}{S}^{\#} = \underset{\sim}{T} \quad , \quad \underset{\sim}{T}^{\#} = \underset{\sim}{S} \quad . \tag{94.2}$$

Let Q be a self-dual quantity which is linearly dependent on the paracanonical coordinates s_y and t_y (or s_z and t_z) of an arbitrary parabasal ray. Then

$$Q = Q_a s_y + Q_b t_y \quad .$$

Under a dual transformation this must be transformed into a valid equation. Thus, using (94.2),

$$Q^{\#} = Q_a{}^{\#} t_y + Q_b{}^{\#} s_y$$

for all s_y and t_y . Since Q is self-dual, it follows that

$$Q_a{}^{\#} = Q_b \quad , \quad Q_b{}^{\#} = Q_a \quad . \tag{94.3}$$

Next, suppose that $\underset{\sim}{Q}$ is a self-dual two-vector which may be expanded as in (24.1). Perform a dual transformation on (24.1) and use (94.2). (<u>Note</u>: An interchange of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ means an interchange of S_y and T_y, S_z and T_z.) Then, since $\underset{\sim}{Q}$ is a self-dual, the duals of the paracanonical coefficients are given by[1]

$$q_{\mu\nu\tau}{}^{(n)\#} = q_{(n-\mu+\nu)\nu\tau}{}^{(n)} \quad . \tag{94.4}$$

Note that the interchange of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ is equivalent to an interchange of $\underset{\sim}{\sigma}$ and $\underset{\sim}{\tau}$, $\bar{\underset{\sim}{\sigma}}$ and $\bar{\underset{\sim}{\tau}}$ in (6.1). Thus, from (6.2),

$$\underset{\sim}{g}^{\#} = -\underset{\sim}{g} \quad . \tag{94.5}$$

Obviously, any quantity which has physical meaning for the system (or for any ray traversing the system), irrespective of whether or not paracanonical coordinates have been defined, must be self-dual. Thus $\underset{\sim}{Y}$, $\underset{\sim}{V}$ and $\underset{\sim}{G}$ are particular examples of self-dual quantities. It is readily checked, using (94.2,3,5), that (25.6,7) are consistent with this.

---

1 Compare (6.3,5) of Buchdahl (1965). See also Table 29/3.

95. The Principle of Duality and its Applications

(a) A dual transformation will now be generalised to the following: a dual transformation is a formal interchange of all quantities and their duals. Consider for the moment the manner in which the expressions for the basal coefficients were derived. Throughout the derivation the basic operational procedure was as follows: some combination $\underline{\underline{C}}$ of functions of $\underset{\sim}{S}$ and $\underset{\sim}{T}$, or of $\underset{\sim}{Y}$ and $\underset{\sim}{V}$, was constructed and set identically equal to zero for all rays. $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ were expressed in terms of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ and the coefficients of the powers of $\underset{\sim}{S}$ and $\underset{\sim}{T}$ in $\underline{\underline{C}}$ were then zero. If a dual transformation is applied to $\underline{\underline{C}}$ it is readily seen that the theorem at the beginning of this appendix may be generalised to the Principle of Duality:

> Under a dual transformation valid equations are transformed into other valid equations.

That this is consistent with the principle of duality as expressed by Buchdahl[1] is obvious. As a simple example consider (14.3). Under a dual transformation $(y|v)$ becomes $y_b v_a - y_a v_b = -(y|v)$. N_y is unaffected since it is a function of the parameters of Ⓚ and g_y transforms according to (94.5). Thus the original (14.3) is recovered. Other examples may be found throughout this thesis, for instance, the iteration equations (Table 37/1), in particular the first and seventh members of these. (Note: (94.4) also applies to pseudo-coefficients.)

(b) There are two important applications of the principle of duality. The first is a check on the various equations appearing in the text of the

---

1 Buchdahl (1965) §6(ii).

thesis. From the principle of duality it follows that corresponding to any equation there exists a second equation formally similar to the original except that all quantities in the original are replaced by their duals. By comparing the relevant equations the existence of an error in one or the other can be ascertained. (e.g., the first and seventh members of Table 37/1 form such a pair.) The second application of the principle of duality is a computational aid.[1] The computer programme used for computing basal coefficients was constructed entirely from relations to which the principle of duality applies. It has already been observed that an interchange of $\underset{\sim}{Q}$ and $\underset{\sim}{T}$, $\bar{\underset{\sim}{Q}}$ and $\bar{\underset{\sim}{T}}$ in (6.1) is equivalent to a dual transformation. The numerical values of these quantities form the "initial data" of the a- and b-rays used to compute the coefficients. It follows that if the same programme is used with the numerical values of $\underset{\sim}{y}_{a1}$ and $\underset{\sim}{y}_{b1}$, $\underset{\sim}{v}_{a1}$ and $\underset{\sim}{v}_{b1}$ reversed, it will yield the correct values of the dual coefficients. When the programme was initially constructed this feature was a very valuable aid to "de-bugging".

The series for the surface and paracanonical expansions of any quantity are formally the same, except that where $\underset{\sim}{Y}$ and $\underset{\sim}{V}$ appear in the former, $\underset{\sim}{S}$ and $\underset{\sim}{T}$ appear in the latter. Thus, if (94.4) is applied to surface coefficients, equations such as those of Table 30/1 may be checked in the same manner as equations involving only paracanonical coefficients. Finally, the dual transformation can be built up from two independent transformations - for example, an interchange of S_y, T_y followed by an interchange of S_z, T_z. Each of these transformations forms a dual transformation in its own right and equations analogous to (94.1-5) may be obtained in each case. A principle of duality exists for each transformation. This is not the case in the axial theory since an interchange of S_{Ay}, T_{Ay} alone does not preserve the rotational invariance of the equations, in particular of ξ , η , ζ .

APPENDIX F: LISTING OF PROGRAMME TIRIKI

This Appendix consists of a listing of the computer programme used to compute the basal coefficients appearing in this thesis. The listing presented here was obtained from a successful production run.

```
C
C
C     PROGRAMME TIRIKI                                  P.J.SANDS, MARCH 1967.
C     ********* ******
C
C
C               THIS PROGRAMME COMPUTES THE BASAL COEFFICIENTS OF
C     THE FIRST THREE ORDERS FOR A SYMMETRIC OPTICAL SYSTEM OF LESS
C     THAN 16 SPHERICAL SURFACES.  THE BASE-RAY IS MERIDIONAL AND
C     MUST CORRESPOND TO A HALF-FIELD OF LESS THAN 90 DEGREES.  THE
C     RELEVANT THEORY CAN BE FOUND IN P. SANDS' THESIS (A.N.U. 1967)
C     AND THE NOTATION OF THIS PROGRAMME IS BASED ON THAT OF SANDS
C     WITH THE TWO MAJOR EXCEPTIONS (A)  S AND T REPLACE P AND R RESPEC-
C     TIVELY AND (B)  THE COEFFICIENTS ARE NUMBERED CONSECUTIVELY.
C     THE COORDINATES EMPLOYED ARE OT-BARRED AND THE BASE-RAY MUST
C     BE SPECIFIED IN AXIAL OT-BARRED COORDINATES.
C
C               THE LOGICAL UNITS ARE AS FOLLOWS.  1 IS A CARD
C     READER, 2 A CARD PUNCH AND 3 A LINE PRINTER.  THE INPUT MEDIUM
C     IS 80-COLUMN PUNCHED CARDS AND IS AS FOLLOWS.
C
C     CARD 1        AN IDENTIFYING STRING OF 36 CHARACTERS, THE NUMBER
C                   OF SURFACES, THE FOCAL LENGTH AND THE APERTURE.
C
C     CARD 2, ETC   (WITH ONE CARD FOR EACH SURFACE)    THE CURVATURE,
C                   DISTANCE FROM THE PREVIOUS SURFACE, REFRACTIVE INDEX
C                   IN THE OBJECT SPACE OF THE SURFACE AND THE CLEAR
C                   APERTURE OF THE SURFACE.
C
C     CARD KK+2     THE POSITION OF THE ENTRANCE PUPIL, RECIPROCAL OF
C                   THE OBJECT DISTANCE AND THE REFRACTIVE INDEX
C                   IN THE IMAGE SPACE OF THE SYSTEM.
C
C     CARD KK+3     THE SURFACE IN WHOSE OBJECT SPACE THE DIAPHRAGM IS
C                   SITUATED AND THE DISTANCE OF THE DIAPHRAGM FROM THE
C                   POLAR TANGENT PLANE OF THE SURFACE.
C
C     CARD KK+4     THE NUMBER OF BASE-RAYS TO BE COMPUTED (LESS THAN
C                   50).
C
C     REMAINING CARDS    THE AXIAL OT-BARRED COORDINATES OF MERIDI-
C                   ONAL BASE-RAYS (Y-COMPONENTS ONLY).
C
C               THE BASIC PRINTOUT PRODUCED BY THE PROGRAMME CONSISTS
C     OF (A)  THE PARAMETERS OF THE SYSTEM AND PARAXIAL COEFFICIENTS
C     AND DATA,  (B)  DATA RELATING TO THE TRACE OF THE BASE-RAY AND
C     SOME PARABASAL DETAILS OF THE BASE-RAY,  (C)  THE PARABASAL
C     COEFFICIENTS AND THE SURFACE CONTRIBUTIONS TO THE A-COMPONENTS
C     OF THE G-COEFFICIENTS (WITH THE FINAL COEFFICIENTS IN THE LAST
C     COLUMN),  (D)  THE SURFACE CONTRIBUTIONS TO THE B-COMPONENTS
C     OF THE G-COEFFICIENTS AND  (E)  THE IMAGE HEIGHT, FINAL V-
C     AND DIAPHRAGM COEFFICIENTS.  IN ADDITION, IF IJK=2 THE SURFACE
```

286

```
      C        CONTRIBUTIONS TO THE IMAGE HEIGHT COEFFICIENTS ARE PRINTED.  MORE-
      C        OVER, IF THE BASE-RAY IS A PRINCIPAL RAY (RHODPG LESS THAN
      C        0.0001) THE ENTRANCE PUPIL COEFFICIENTS AND PRINCIPAL RAY
      C        IMAGE HEIGHT COEFFICIENTS ARE COMPUTED AND PRINTED.  THE PUNCH-
      C        OUT IS (A) THE TITLE OF THE SYSTEM, ITS FOCAL LENGTH AND STOP
      C        NUMBER AND THE SPC OF THE BASE-RAY,  (B)  SOME DATA RELATING
      C        TO THE BASE-RAY IN THE OBJECT AND IMAGE SPACES,  (C)  THE
      C        PARAXIAL COEFFICIENTS AT THE FIRST SURFACE AND  (D)  THE IMAGE
      C        HEIGHT AND FINAL V-COEFFICIENTS AND THE DIAPHRAGM COEFFICIENTS.
      C
      C                  IN THE COMMON BLOCKS DIMENSIONED VARIABLES AA, AB
      C        ETC. ARE DUMMY VARIABLES FOR WHICH THE SECOND INDEX IS KK OR
      C        KK+1 AS THE CASE MAY BE.  A1, A2 ETC. ARE DUMMY VARIABLES
      C        INDEPENDENT OF THE NUMBER OF SURFACES.  THE SUBROUTINES REQUIRED
      C        BY THE PROGRAMME ARE PARAXL, BASRAY, XST123, PSEUDO, GCOEFS,
      C        ITRATE, ST2VEC, HDCOEF AND PRBRAY WITH ITRATE AND ST2VEC CALLED
      C        ONLY BY GCOEFS AND COEFOP CALLED BY BOTH GCOEFS AND HDCOEF.
      C
      C
0001           COMMON SYSTEM(9),KQ,FOCAL,STPNUM
0002           DIMENSION SYB(50),TYB(50)
      C
0003           CALL PARAXL(1,SYB,TYB,N,KK)
0004           IJK=2
      C
0005           DO 1 J=1,N
0006           WRITE(3,400)
0007           WRITE(3,401) SYSTEM,FOCAL,STPNUM
0008           WRITE(3,402) SYB(J),TYB(J)
0009           CALL BASRAY(SYB(J),TYB(J),KK)
0010           CALL PARBAS(AMUY,AMUZ)
0011           CALL XST123
0012           CALL PSEUDO
0013           CALL GCOEFS
0014           CALL HDCOEF(AMUY,AMUZ,IJK)
0015           CALL PRBRAY
0016         1 CONTINUE
      C
0017           STOP
      C
0018       400 FORMAT('1',15(/),16X,'PARABASAL, SECOND AND THIRD ORDER BASAL G-CO
              1EFFICIENTS AND BASAL IMAGE HEIGHT COEFFICIENTS'/16X,90('=')//34X,'
              2FOR A SYMMETRIC OPTICAL SYSTEM OF SPHERICAL SURFACES,'/34X,52('=')
              3//46X,'THE BASE RAY BEING MERIDIONAL.'/46X,29('='))
0019       401 FORMAT(10(/),35X,'THE SYSTEM IS  -   ',9A4/35X,13('=')//45X,'FOCAL
              1LENGTH IS  -   F ='F7.4//45X,'STOP NUMBER IS  -   NU ='F6.2)
0020       402 FORMAT(5(/),35X,'THE BASE RAY IS MERIDIONAL  -'/35X,26('=')//42X,
              1'SBY = ',E12.5,'   TBY = ',E12.5)
0021           END
```

```
0001          SUBROUTINE PARAXL(LPARAX,SYB,TYB,NRAY,KKK)                      TK1020
        C                                                                     TK1030
        C           SUBROUTINE PARAXL READS IN THE DATA OF A SYMMETRIC        TK1040
        C     SYSTEM OF SPHERICAL SURFACES, COMPUTES THE PARAXIAL COEFFICIENTS TK1050
        C     OF THE SYSTEM AND PRINTS OUT THE SYSTEM AND PARAXIAL DATA COR-  TK1060
        C     RECTLY NORMALISED TO A PRESCRIBED FOCAL LENGTH.  THE CONTROL    TK1070
        C     CHARACTER LPARAX GOVERNS THE FOLLOWING OPTIONS -                TK1080
        C                                                                     TK1090
        C     LPARAX = 1  READS THE DATA, COMPUTES THE PARAXIAL DETAILS AND   TK1100
        C                 PRINTS THESE,                                       TK1110
        C                                                                     TK1120
        C     LPARAX = 2  CAUSES THE READING OF THE DATA TO BE DELETED,       TK1130
        C                                                                     TK1140
        C     LPARAX = 3  MERELY PRINTS OUT THE SYSTEM AND PARAXIAL DATA ON   TK1150
        C                 THE ASSUMPTION THAT THE LATTER ARE KNOWN.           TK1160
        C                                                                     TK1170
        C     IF MORE THAN 15 SURFACES ARE SPECIFIED, EXECUTION IS TERMINATED. TK1180
        C     IF ENPUP = 0, IT IS RE-EVALUATED SO AS TO CORRESPOND TO DSTP,   TK1190
        C     OTHERWISE DSTP IS DETERMINED SO AS TO CORRESPOND TO ENPUP.      TK1200
        C     IN ORDER TO USE CANONICAL COORDINATES, SPECIFY JD = 1 AND       TK1210
        C     DSTP = 0.                                                       TK1220
        C                                                                     TK1230
0002          COMMON SYSTEM(9),KK,FOCAL,STPNUM,C(16),D(16),N(16),RHO(16),ENPUP TK1240
0003          COMMON OBJECT,JD,DSTP,K(15)                                     TK1250
0004          COMMON IA(16),IB(16),VA(16),VB(16),YA(16),YB(16),IMAGE(16)      TK1260
0005          COMMON RHOSTP,EXPUP                                             TK1270
0006          REAL K,N,IA,IB,IMAGE                                            TK1280
0007          DIMENSION SYB(50),TYB(50)                                       TK1290
        C                                                                     TK1300
0008          GO TO (100,200,300),LPARAX                                      TK1310
        C                                                                     TK1320
        C     PROCEDURE FOR READING SYSTEM PARAMETERS.                        TK1330
        C                                                                     TK1340
0009      100 READ (1,400) SYSTEM,KK,FOCAL,STPNUM                             TK1350
0010          WRITE(3,400) SYSTEM,KK,FOCAL,STPNUM                             TK1360
0011          IF(KK-15)2,2,1                                                  TK1370
0012        1   WRITE(3,401)KK                                                TK1380
0013            STOP                                                          TK1390
0014        2   KP=KK+1                                                       TK1400
0015          KKK=KK                                                          TK1410
0016          DO 3  J=1,KK                                                    TK1420
0017            READ (1,402) C(J),D(J),N(J),RHO(J)                            TK1430
0018        3   WRITE(3,402) C(J),D(J),N(J),RHO(J)                            TK1440
0019          READ(1,402) ENPUP,OBJECT,N(KP)                                  TK1450
0020          WRITE(3,402)ENPUP,OBJECT,N(KP)                                  TK1460
0021          READ(1,403) JD,DSTP                                             TK1470
0022          WRITE(3,403)JD,DSTP                                             TK1480
0023          DO 4  J=1,KK                                                    TK1490
0024        4   K(J)=N(J)/N(J+1)                                              TK1500
        C                                                                     TK1510
0025          READ(1,404) NRAY                                                TK1520
0026          WRITE(3,404) NRAY                                               TK1530
```

```
              DO 5 J=1,NRAY
                READ(1,402) SYB(J),TYB(J)
      5         WRITE(3,402) SYB(J),TYB(J)
C
C             PROCEDURE FOR CALCULATING PARAXIAL COEFFICIENTS.
C
    200       M=1
      6       D(KP)=0.0
              VA(1)=-OBJECT
              YA(1)=1.
      7       VB(1)=1./(1.-ENPUP*OBJECT)
              YB(1)=-ENPUP*VB(1)
C
              DO 8 J=2,KP
                JJ=J-1
                IA(JJ)=C(JJ)*YA(JJ)+VA(JJ)
                IB(JJ)=C(JJ)*YB(JJ)+VB(JJ)
                X=K(JJ)-1.
                VA(J)=X*IA(JJ)+VA(JJ)
                VB(J)=X*IB(JJ)+VB(JJ)
                YA(J)=D(J)*VA(J)+YA(JJ)
      8         YB(J)=D(J)*VB(J)+YB(JJ)
C
              IF(ENPUP)10,9,10
      9         ENPUP=(YB(JD)+DSTP*VB(JD))/(YA(JD)+DSTP*VA(JD))
                IF(ENPUP)7,10,7
     10         DSTP=-YB(JD)/VB(JD)
              GO TO (11,13),M
C
     11       FOCALL=-1./FOCAL/VA(KP)
              DO 12 J=1,KK
                C(J)=C(J)*FOCALL
                D(J)=D(J)/FOCALL
     12         RHO(J)=RHO(J)/FOCALL
              ENPUP=ENPUP/FOCALL
              OBJECT=OBJECT*FOCALL
              DSTP=DSTP/FOCALL
                M=2
                GO TO 6
C
     13       EXPUP=-YB(KP)/VB(KP)
              DO 14 J=1,KK
     14         IMAGE(J)=-YA(J)/VA(J+1)
              FOCAL=-1./VA(KP)
              RHOSTP=N(1)*VB(1)*FOCAL/2/N(JD)/VB(JD)/STPNUM
C
C             PROCEDURE FOR PRINTING THE SYSTEM AND PARAXIAL DATA.
C
    300       WRITE(3,500) SYSTEM,FOCAL,STPNUM
              WRITE(3,501)
              DO 15 J=1,KK
     15         WRITE(3,502) J,C(J),N(J),K(J),D(J),RHO(J)
```

```
0069              WRITE(3,503) RHOSTP,DSTP,JD                                  TK2070
0070              IF(OBJECT)16,17,16                                           TK2080
0071        16      OB=1./OBJECT                                               TK2090
0072                WRITE(3,504) ENPUP,EXPUP,OB,IMAGE(KK)                      TK2100
0073                GO TO 18                                                   TK2110
0074        17      WRITE(3,508) ENPUP,EXPUP,IMAGE(KK)                         TK2120
0075        18    WRITE(3,506)                                                 TK2130
0076              DO 19 J=1,KK                                                 TK2140
0077        19      WRITE(3,502) J,YA(J),VA(J),YB(J),VB(J),IMAGE(J)            TK2150
0078              WRITE(3,509)                                                 TK2160
0079              DO 20 J=1,KK                                                 TK2170
0080        20      WRITE(3,502) J,IA(J),IB(J)                                 TK2180
0081              WRITE(3,507) YA(KP),VA(KP),YB(KP),VB(KP)                     TK2190
          C                                                                    TK2200
0082              RETURN                                                       TK2210
          C                                                                    TK2220
0083       400    FORMAT(1X,9A4,4X,I2,1X,E10.3,1X,F4.1)                        TK2230
0084       401    FORMAT(' SYSTEM IS TOO LARGE, K = ',I2)                      TK2240
0085       402    FORMAT(4(1X,E12.5))                                          TK2250
0086       403    FORMAT(1X,I2,1X,E12.5)                                       TK2260
0087       404    FORMAT(1X,I3)                                                TK2270
0088       500    FORMAT('1'///1X,9A4,4X,'FOCAL LENGTH = ',F8.5,',  STOP = F/',F5.2/ TK2280
                 1/)                                                           TK2290
0089       501    FORMAT(/35X,'CONSTRUCTION OF SYSTEM'/35X,22('=')//15X,'C',15X,'N', TK2300
                 115X,'K',15X,'D',14X,'RHO'/)                                  TK2310
0090       502    FORMAT(3X,I2,7(4X,E12.5))                                    TK2320
0091       503    FORMAT('0','DIAPHRAGM IS ',E12.5,' IN DIAMETER, AND SITUATED ', TK2330
                 1E12.5,' BEFORE THE ',I2,'TH SURFACE.')                       TK2340
0092       504    FORMAT('0ENTRANCE PUPIL AT P = ',E12.5,', EXIT PUPIL AT P'' = ', TK2350
                 1E12.5/'0OBJECT AT L = ',E12.5,', IDEAL IMAGE AT L'' = ',E12.5) TK2360
0093       506    FORMAT(////35X,'PARAXIAL COEFFICIENTS'/35X,21('=')///14X,'YA',14 TK2370
                 1X,'VA',14X,'YB',14X,'VB',14X,'L'''/)                         TK2380
0094       507    FORMAT(//' FINAL COEFFICIENTS -'/11X,'YA = ',E12.5,', VA = ',E12.5 TK2390
                 1,', YB = ',E12.5,', VB = ',E12.5)                            TK2400
0095       508    FORMAT('0ENTRANCE PUPIL AT P = ',E12.5,', EXIT PUPIL AT P'' = ', TK2410
                 1E12.5/'0OBJECT AT INFINITY, IDEAL IMAGE AT L'' = ',E12.5)    TK2420
0096       509    FORMAT(//14X,'IA',14X,'IB'/)                                 TK2430
0097              END                                                          TK2440
```

```
0001          SUBROUTINE BASRAY(SYY,TYY,KKK)
        C
        C                THE TRACE OF THE BASE-RAY IS BASED OF FORD'S RAY
        C     TRACE SCHEME (P. FORD, JOSA., 50, 528-533, 1960).  THE BASIC
        C     SCHEME HAS BEEN MODIFIED SO AS TO TRACE MERIDIONAL RAYS ONLY
        C     AND COMPUTE THE PSEUDO-PARAMETERS AND OTHER QUANTITIES OF
        C     INTEREST ASSOCIATED WITH THE BASE-RAY.  VIGNETTING IS DETER-
        C     MINED AND ANNOUNCED BUT OTHERWISE IGNORED.  IF THE BASE-RAY
        C     IS TOTALLY INTERNALLY REFLECTED OR MISSES A SURFACE, THE SYSTEM
        C     IS TRUNCATED BY DELETING THE OFFENDING AND SUBSEQUENT SURFACES.
        C     PROCESSING THEN CONTINUES BUT BEFORE CONSIDERING THE NEXT BASE-
        C     RAY, THE SYSTEM IS CONVERTED TO ITS ORIGINAL FORM.  IF THE
        C     BASE-RAY MISSES THE FIRST SURFACE, EXECUTION IS TERMINATED.
        C
        C                NOTE - AS IT STANDS, THIS PROGRAMME CANNOT ACCEPT
        C     SYSTEMS INCORPORATING MIRRORS.  THE REQUIRED MODIFICATIONS ARE
        C     SIMPLE.
        C
0002          COMMON SYSTEM(9),KK,FOCAL,STPNUM,C(16),D(16),N(16),RHO(16),ENPUP
0003          COMMON OBJECT,JD,DSTP,K(15)
0004          COMMON IA(16),IB(16),VA(16),VB(16),YA(16),YB(16),IMAGE(16)
0005          COMMON RHOSTP,EXPUP
0006          COMMON SYB,TYB,VBB(16),IBY(15),NBX(15),NBY(15),SB(15),XB(15)
0007          COMMON YBB(15),ALPHA(16),BETA(16),HBY,RHODPG,COSINC(15),COSRFC(15)
0008          REAL K,N,IA,IB,IMAGE,IBY,NBY,NBX
0009          DIMENSION THETA(15),ANGINC(15),ANGRFC(15),PSI(15)
0010          ASIN(X)=ATAN(X/SQRT(1-X*X))
        C
0011          TYB=TYY
0012          SYB=SYY
0013          KK=KKK
        C     (NOTE - THIS RESTORES THE SYSTEM TO ITS ORIGINAL FORM IF
        C     TRUNCATION OCCURED FOR THE PREVIOUS BASE RAY BY EITHER
        C     S.0031 OR S.0043.)
0014          Y=YA(1)*SYB+YB(1)*TYB
0015          V=VA(1)*SYB+VB(1)*TYB
0016          VBB(1)=V
0017          SY=SYY
0018          TY=TYY
        C
0019          DO 13 J=1,KK
0020            IBY(J)=C(J)*Y+VBB(J)
0021            IF(J-JD)3,1,3
0022      1       RHODPG=Y+V*DSTP
0023              IF(RHOSTP-ABS(RHODPG))2,3,3
0024      2       WRITE(3,403)RHODPG
0025      3     P2=1.+V*V
0026            P1=1.-P2-C(J)*Y*V
0027            P0=C(J)*C(J)*Y*Y-P2-2.*P1
0028            P=P1*P1-P0*P2
0029            IF(P)4,5,5
0030      4       WRITE(3,404)J
```

```
                 KK=J-1
                 GO TO 14
       5      P=(SQRT(P)-P1)/P2
              NBX(J)=P
              NBY(J)=-IBY(J)+NBX(J)*VBB(J)
              U=K(J)*K(J)
              U0=(U-1.)*P2
              U1=U*P1+P2*P
              U2=U*(P0+1.)-P2*P*P
              S=U1*U1-U0*U2
              IF(S)6,7,7
       6         WRITE(3,405)J
                 KK=J-1
                 GO TO 14
       7      S=(SQRT(S)-U1)/U2
              SB(J)=S
              IF(C(J))9,8,9
       8         XB(J)=0.
                 YBB(J)=Y
                 GO TO 10
       9         XB(J)=(1.-P)/C(J)
                 YBB(J)=Y+XB(J)*V
      10         T=S*XB(J)
                 IF(RHO(J)-ABS(YBB(J)))11,12,12
      11            WRITE(3,406)J,YBB(J)
      12      Q=K(J)+S-1.
              JJ=J+1
              DA=N(JJ)*(YA(J)*Q+VA(JJ)*T)/N(1)
              DB=N(JJ)*(YB(J)*Q+VB(JJ)*T)/N(1)
              SY=SY+DB*IBY(J)
              TY=TY-DA*IBY(J)
              Y=YA(JJ)*SY+YB(JJ)*TY
              V=VA(JJ)*SY+VB(JJ)*TY
              VBB(JJ)=V
              ALPHA(J)=SQRT(1./P2)
              BETA(J)=ALPHA(J)*VBB(J)
              THETA(J)=ASIN(NBY(J))/1.745329E-02
              PSI(J)=ATAN(VBB(J))/1.745329E-02
              ANGRFC(J)=ASIN(K(J)*ALPHA(J)*IBY(J))/1.745329E-02
              ANGINC(J)=ASIN(ALPHA(J)*IBY(J))/1.745329E-02
              COSINC(J)=COS(ANGINC(J)*1.745329E-02)
      13      COSRFC(J)=COS(ANGRFC(J)*1.745329E-02)
C
      14      KP=KK+1
              IF(KK)15,15,16
      15        WRITE(3,407)
                STOP
      16      HBY=Y+IMAGE(KK)*V
              AMAG=-N(1)/N(KP)/VA(KP)
              EPSBY=HBY-AMAG*TYB
              ALPHA(KP)=1./SQRT(1.+VBB(KP)*VBB(KP))
              BETA(KP)=ALPHA(KP)*VBB(KP)
```

292

```
            C
0082              WRITE(3,400)  SYSTEM,FOCAL,STPNUM
0083              WRITE(3,401)
0084              WRITE(3,402)  SYB,TYB
0085              WRITE(3,408)
0086              DO 17  J=1,KK
0087        17      WRITE(3,409)J,PSI(J),THETA(J),ANGINC(J),ANGRFC(J),XB(J),YBB(J)
0088              PSI(KP)=ATAN(VBB(KP))/1.745329E-02
0089              WRITE(3,410)  PSI(KP),HBY,EPSBY,RHODPG
            C
0090              RETURN
            C
0091        400   FORMAT('1'1X,9A4,4X,'FOCAL LENGTH =',F6.4,',  STOP = ',F4.1)
0092        401   FORMAT(///26X,'DATA RELATING TO BASE RAY'/26X,25('='))
0093        402   FORMAT(//' BASE RAY IS MERIDIONAL - SY = ',E12.5,',  TY = ',E12.5)
0094        403   FORMAT(/' BASE RAY VIGNETTED BY DIAPHRAGM, RHODPG = ',E12.5)
0095        404   FORMAT(/' BASE RAY MISSES ',I2,'TH. SURFACE')
0096        405   FORMAT(/' BASE RAY IS TOTALLY INTERNALLY REFLECTED AT ',I2,'TH.
                 1SURFACE')
0097        406   FORMAT(/' BASE RAY VIGNETTED BY ',I2,'TH. SURFACE, RHO = ',E12.5)
0098        407   FORMAT(/' TERMINATED')
0099        408   FORMAT(///12X,'PSI',5X,'THETA',6X,'I',7X,'I''',9X,'XB',13X,'YB'/)
0100        409   FORMAT(3X,I2,1X,4(3X,F6.1),2(3X,E12.5))
0101        410   FORMAT(3X,'K''',4X,F6.1//' INTERSECTION HEIGHT OF BASE RAY',8X,'=
                 1',E12.5/' ABERRATION OF BASE RAY',17X,'= ',E12.5/' BASE RAY INTERS
                 2ECTS DIAPHRAGM WITH RHO = ',E12.5)
0102              END
```

```
0001            SUBROUTINE PARBAS(MUY,MUZ)
          C
          C                THE PARABASAL COEFFICIENTS ARE DETERMINED BY TRACING
          C     THE MERIDIONAL AND SAGITTAL A- AND B-RAYS USING THE PARABASAL
          C     RAY TRACE EQUATIONS.  THE PARABASAL I-COEFFICIENTS, THE POSI-
          C     TIONS OF THE MERIDIONAL AND SAGITTAL FOCI AND THE AUGMENTING
          C     FACTORS ARE ALSO DETERMINED.  THE PARABASAL COEFFICIENTS ARE
          C     PRINTED BY SUBROUTINE COEFOP, SOME PARABASAL DETAILS ARE HOW-
          C     EVER PRINTED BY SUBROUTINE PARBAS.
          C
0002            COMMON SYSTEM(9),KK,FOCAL,STPNUM,C(16),D(16),N(16),RHO(16),ENPUP
0003            COMMON OBJECT,JD,DSTP,K(15)
0004            COMMON IAA(16),IAB(16),VAA(16),VAB(16),YAA(16),YAB(16),IMAGE(16)
0005            COMMON RHOSTP,EXPUP
0006            COMMON SYB,TYB,VBB(16),IBY(15),NBX(15),NBY(15),SB(15),XB(15)
0007            COMMON YBB(15),ALPHA(16),BETA(16),HBY,EPSBY,COSINC(15),COSRFC(15)
0008            COMMON NY(16),NZ(16),YA(16),VA(16),IYA(16),YAP(16),YB(16),VB(16)
0009            COMMON IYB(16),YBP(16),ZA(16),WA(16),ZB(16),WB(16),IZA(16),IZB(16)
0010            COMMON DELTAM(15),DELTAS(15),GY,GZ
0011            DIMENSION DD(16),X(16),KY(16),KZ(16),LAMBDA(16)
0012            REAL K,N,IAA,IAB,IMAGE,IBY,NBY,NBX,NY,NZ,IYA,IYB,IZA,IZB,KY,KZ
0013            REAL LAMBDA,MUY,MUZ
          C
0014            KP=KK+1
0015            DO 1 J=1,KK
0016              NZ(J)=N(J)*ALPHA(J)
0017              NY(J)=NZ(J)*ALPHA(J)*ALPHA(J)
0018              LAMBDA(J)=ALPHA(J)/COSINC(J)
0019        1     X(J)=LAMBDA(J)*COSRFC(J)/ALPHA(J+1)
0020            NZ(KP)=N(KP)*ALPHA(KP)
0021            NY(KP)=NZ(KP)*ALPHA(KP)*ALPHA(KP)
0022            DO 2 J=2,KK
0023        2     DD(J)=D(J)+XB(J)-XB(J-1)
0024            DD(KP)=0.
0025            P=ENPUP-XB(1)
          C
0026            YA(1)=1.-OBJECT*XB(1)
0027            VA(1)=VAA(1)
0028            VB(1)=VAB(1)
0029            YB(1)=-P*VB(1)
0030            ZA(1)=YA(1)
0031            ZB(1)=YB(1)
0032            WA(1)=VA(1)
0033            WB(1)=VB(1)
0034            GY=NY(1)
0035            GZ=NZ(1)
          C     (NOTE - FOR OT-BARRED COORDINATES G=1. THUS GY AND GZ ARE THE
          C     QUANTITIES ASSOCIATED WITH THE ANTE-PRIME.)
          C
0036            DO 3 J=2,KP
0037              JJ=J-1
0038              YAP(JJ)=X(JJ)*YA(JJ)
```

294

```
0039           YBP(JJ)=X(JJ)*YB(JJ)
0040           KY(JJ)=NY(JJ)/NY(J)/X(JJ)
0041           KZ(JJ)=NZ(JJ)/NZ(J)
0042           Z=-C(JJ)*SB(JJ)
0043           WA(J)=KZ(JJ)*WA(JJ)+Z*ZA(JJ)
0044           WB(J)=KZ(JJ)*WB(JJ)+Z*ZB(JJ)
0045           Z=Z*LAMBDA(JJ)/ALPHA(J)/COSRFC(JJ)
0046           VA(J)=VA(JJ)*KY(JJ)+Z*YA(JJ)
0047           VB(J)=VB(JJ)*KY(JJ)+Z*YB(JJ)
0048           YA(J)=YAP(JJ)+DD(J)*VA(J)
0049           YB(J)=YBP(JJ)+DD(J)*VB(J)
0050           ZA(J)=ZA(JJ)+DD(J)*WA(J)
0051           ZB(J)=ZB(JJ)+DD(J)*WB(J)
0052           IYA(JJ)=C(JJ)*YA(JJ)+NBX(JJ)*VA(JJ)
0053           IYB(JJ)=C(JJ)*YB(JJ)+NBX(JJ)*VB(JJ)
0054           IZA(JJ)=C(JJ)*ZA(JJ)+NBX(JJ)*WA(JJ)
0055           IZB(JJ)=C(JJ)*ZB(JJ)+NBX(JJ)*WB(JJ)
0056           OB=IMAGE(JJ)-XB(JJ)
0057           DELTAM(JJ)=-OB-YAP(JJ)/VA(J)
0058     3     DELTAS(JJ)=-OB-ZA(JJ)/WA(J)
       C
0059         MUZ=-NZ(KP)*WA(KP)
0060         MUY=-NY(KP)*VA(KP)
0061         WRITE(3,401)
0062         WRITE(3,405)
0063         DO 6 J=1,KK
0064     6   WRITE(3,403)J,NY(J),NZ(J),DELTAM(J),DELTAS(J)
0065         WRITE(3,403)KP,NY(KP),NZ(KP)
0066         WRITE(3,406)
0067         WRITE(3,407)YA(KP),VA(KP),YB(KP),VB(KP),ZA(KP),WA(KP),ZB(KP),WB(KP
            1)
0068         WRITE(3,402) MUY,MUZ
0069         WRITE(3,408)P,OB
       C
0070         RETURN
       C
0071   401   FORMAT(////26X,'SOME PARABASAL DETAILS'/26X,22('=')//)
0072   402   FORMAT(//11X,'AUGMENTING FACTORS ARE -', 1X,'MUY = ',F8.5,',   MUZ
            1= ',F8.5)
0073   403   FORMAT(1X,I2,2X,8(2X,E12.5))
0074   405   FORMAT(//11X,'NY',12X,'NZ',10X,'DELTA M',8X,'DELTA S'/)
0075   406   FORMAT(//11X,'FINAL COEFFICIENTS ARE -'/)
0076   407   FORMAT( 4X,'YA = ',E12.5,', VA = ',E12.5,', YB = ',E12.5,', VB = ',
            1E12.5/ 4X,'ZA = ',E12.5,', WA = ',E12.5,', ZB = ',E12.5,', WB = ',
            2E12.5)
0077   408   FORMAT(//11X,'P (BASAL) = ',E12.5,',   L'' (BASAL) = ',E12.5)
0078         END
```

```
0001        SUBROUTINE XST123                                                  TK4750
      C                                                                        TK4760
      C               SUBROUTINE XST123 COMPUTES IN FOUR STAGES THE FIRST,    TK4770
      C     SECOND AND THIRD ORDER X-, S- AND T- COEFFICIENTS.  THE FOUR      TK4780
      C     STAGES ARE   (A)  THE COMPUTATION OF THE X-COEFFICIENTS,   (B)    TK4790
      C     THE COMPUTATION OF THE MU-, SIGMA- AND TAU-COEFFICIENTS,  (C)     TK4800
      C     THE COMPUTATION OF THE S-COEFFICIENTS AND FINALLY  (D)  THE       TK4810
      C     COMPUTATION OF THE T-COEFFICIENTS.  THE ENTIRE EXECUTABLE PORTION TK4820
      C     OF THE SUBROUTINE IS EXECUTED ONCE FOR EACH SURFACE.  THE S-      TK4830
      C     COEFFICIENTS ARE STORED IN MEMORY LOCATIONS LABELLED AS FOR Q-    TK4840
      C     COEFFICIENTS AND, AFTER THEY HAVE BEEN COMPUTED, THE T-COEFFIC-   TK4850
      C     IENTS ARE STORED IN THE LOCATIONS OF THE ORIGINAL X-COEFFICIENTS  TK4860
      C     FOR THE SURFACE UNDER COSIDERATION.                               TK4870
      C                                                                        TK4880
0002        COMMON SYSTEM(9),KK,FOCAL,STPNUM,C(16),D(16),N(16),RHO(16),ENPUP  TK4890
0003        COMMON OBJECT,JD,DSTP,K(15),AA(7,16),A1(4)                        TK4900
0004        COMMON VBB(16),IBY(15),NBX(15),NBY(15),SB(15),XB(15),YBB(15)      TK4910
0005        COMMON ALPHA(16),BETA(16),HBY,EPBY,COSINC(15),COSRFC(15),AC(18,16) TK4920
0006        COMMON PX(2,15),SX(6,15),TX(10,15),PQ(2,15),SQ(6,15),TQ(10,15)    TK4930
0007        DIMENSION LAMBDA(4),S1(2),T1(2),S2(6),T2(6),S3(10),T3(10),        TK4940
           1PS(2),SS(6),TS(6),PPSI(2),SPSI(6),SPHI(6),TPHI(10)                TK4950
0008        REAL K,N,IBY,NBY,NBX,LAMBDA,M1,M2                                 TK4960
      C                                                                        TK4970
0009        DO 3  J=1,KK                                                      TK4980
0010        LAMBDA(1)=ALPHA(J)/COSINC(J)                                      TK4990
0011        LAMBDA(2)=1./2./K(J)/ALPHA(J)/COSRFC(J)                           TK5000
0012        LAMBDA(3)=K(J)*K(J)*ALPHA(J)*ALPHA(J)*LAMBDA(2)                   TK5010
0013        LAMBDA(4)=-2.*SB(J)*NBX(J)                                        TK5020
      C                                                                        TK5030
      C     PROCEDURE FOR COMPUTATION OF X-COEFFICIENTS.                      TK5040
      C                                                                        TK5050
0014          PX(1,J)=-NBY(J)*LAMBDA(1)                                       TK5060
0015          PX(2,J)=0.                                                      TK5070
0016          SX(4,J)=C(J)*LAMBDA(1)/2.                                       TK5080
0017          Z=LAMBDA(1)*LAMBDA(1)                                           TK5090
0018          SX(1,J)=SX(4,J)*Z                                               TK5100
0019          SX(2,J)=PX(1,J)*PX(1,J)                                         TK5110
0020          SX(3,J)=0.                                                      TK5120
0021          SX(5,J)=0.                                                      TK5130
0022          SX(6,J)=0.                                                      TK5140
0023          Z=Z*C(J)*IBY(J)                                                 TK5150
0024          TX(1,J)=Z*SX(1,J)                                               TK5160
0025          TX(2,J)=3.*PX(1,J)*SX(1,J)                                      TK5170
0026          TX(3,J)=PX(1,J)*SX(2,J)                                         TK5180
0027          TX(4,J)=Z*SX(4,J)                                               TK5190
0028          TX(6,J)=0.                                                      TK5200
0029          TX(7,J)=0.                                                      TK5210
0030          TX(8,J)=SX(4,J)*PX(1,J)                                         TK5220
0031          TX(5,J)=2.*TX(8,J)                                              TK5230
0032          TX(9,J)=0.                                                      TK5240
0033          TX(10,J)=0.                                                     TK5250
      C                                                                        TK5260
```

296

```
C     PROCEDURE FOR COMPUTATION OF MU-, SIGMA- AND TAU-COEFFICIENTS.
C
      CLAM2=C(J)*LAMBDA(2)
      CLAM3=C(J)*LAMBDA(3)
      A2=ALPHA(J)*ALPHA(J)
      AB=2.*ALPHA(J)*BETA(J)
      T0=(1-K(J)*K(J))*LAMBDA(2)/SB(J)-1.
      M1=(T0+1.)*AB
      M2=(T0+1.)*A2
      S1(1)=C(J)*(LAMBDA(3)*VBB(J)+LAMBDA(2)*PX(1,J))
      S1(2)=LAMBDA(3)*(IBY(J)+NBX(J)*VBB(J))-AB*LAMBDA(2)*NBX(J)
      XX=NBY(J)*LAMBDA(3)
      T1(1)=LAMBDA(4)*S1(1)+2.*C(J)*SB(J)*XX
      T1(2)=LAMBDA(4)/2.*(S1(2)-XX)
      S2(1)=CLAM2*SX(1,J)
      S2(2)=CLAM3+CLAM2*(AB*PX(1,J)+SX(2,J))
      S2(3)=NBX(J)*(LAMBDA(3)-A2*LAMBDA(2))
      S2(6)=S2(3)
      S2(4)=CLAM2*SX(4,J)
      S2(5)=CLAM3
      T2(1)=LAMBDA(4)*S2(1)+C(J)*SB(J)*(CLAM2*SX(2,J)-CLAM3)
      T2(2)=LAMBDA(4)*S2(2)
      T2(3)=LAMBDA(4)*S2(3)/2.
      T2(6)=T2(3)
      T2(4)=LAMBDA(4)*S2(4)-SB(J)*C(J)*CLAM3
      T2(5)=LAMBDA(4)*S2(5)
      S3(1)=CLAM2*TX(1,J)
      S3(2)=CLAM2*(TX(2,J)+AB*SX(1,J))
      S3(3)=CLAM2*(TX(3,J)+AB*SX(2,J)+A2*PX(1,J))
      S3(4)=CLAM2*TX(4,J)
      S3(5)=CLAM2*TX(5,J)
      S3(6)=CLAM2*A2*PX(1,J)
      S3(8)=CLAM2*(TX(8,J)+AB*SX(4,J))
      T3(1)=LAMBDA(4)*S3(1)+2.*C(J)*CLAM2*SB(J)*PX(1,J)*SX(1,J)
      T3(2)=LAMBDA(4)*S3(2)+C(J)*CLAM2*SB(J)*SX(2,J)*(AB+2.*PX(1,J))
      T3(3)=LAMBDA(4)*S3(3)
      T3(4)=LAMBDA(4)*S3(4)+C(J)*SB(J)*S3(5)
      T3(5)=LAMBDA(4)*S3(5)
      T3(6)=LAMBDA(4)*S3(6)
      T3(8)=LAMBDA(4)*S3(8)
C
C     PROCEDURE FOR COMPUTATION OF S-COEFFICIENTS.
C
      DO 21 I=1,6
   21   SPHI(I)=T2(I)+2.*S2(I)
      DO 22 I=1,10
   22   TPHI(I)=T3(I)+2.*S3(I)
      PS(1)=T1(1)+2.*S1(1)
      PS(2)=T1(2)+2.*S1(2)+M1
      PPSI(1)=T0*PS(1)+2.*S1(1)+2.*T1(1)
      PPSI(2)=T0*PS(2)+2.*S1(2)+2.*T1(2)
      X=T2(3)+2.*S2(3)+M2
```

```
0081            SS(1)=SPHI(1)+PS(1)*PPSI(1)                                    TK5790
0082            SS(2)=SPHI(2)+PS(1)*PPSI(2)+PS(2)*PPSI(1)                      TK5800
0083            SS(3)=X+PS(2)*PPSI(2)                                          TK5810
0084            SS(4)=SPHI(4)                                                  TK5820
0085            SS(5)=SPHI(5)                                                  TK5830
0086            SS(6)=X                                                        TK5840
0087            DO 23 I=1,6                                                    TK5850
0088      23      SPSI(I)=T0*SS(I)+2.*S2(I)+2.*T2(I)                           TK5860
0089            TS(1)=TPHI(1)+PS(1)*(SPSI(1)+T1(1)*PS(1))+SS(1)*PPSI(1)        TK5870
0090            TS(2)=TPHI(2)+PS(1)*(SPSI(2)+T1(1)*PS(2)+T1(2)*PS(1))+PS(2)*(SPSI( TK5880
               11)+T1(1)*PS(1))+PPSI(1)*SS(2)+PPSI(2)*SS(1)                   TK5890
0091            TS(3)=TPHI(3)+PS(1)*(SPSI(3)+T1(2)*PS(2))+PS(2)*(SPSI(2)+PS(2)*T1( TK5900
               11)+PS(1)*T1(2))+PPSI(1)*SS(3)+PPSI(2)*SS(2)                   TK5910
0092            TS(4)=TPHI(4)+PS(1)*SPSI(4)+SS(4)*PPSI(1)                      TK5920
0093            TS(5)=TPHI(5)+PS(1)*SPSI(5)+SS(5)*PPSI(1)                      TK5930
0094            TS(6)=TPHI(6)+PS(1)*SPSI(6)+SS(6)*PPSI(1)                      TK5940
0095            TS(7)=PS(2)*(SPSI(3)+T1(2)*PS(2))+SS(3)*PPSI(2)                TK5950
0096            TS(8)=TPHI(8)+PS(2)*SPSI(4)+SS(4)*PPSI(2)                      TK5960
0097            TS(9)=PS(2)*SPSI(5)+SS(5)*PPSI(2)                              TK5970
0098            TS(10)=PS(2)*SPSI(6)+SS(6)*PPSI(2)                             TK5980
0099            PQ(1,J)=PS(1)                                                  TK5990
0100            PQ(2,J)=PS(2)                                                  TK6000
0101            DO 1 I=1,6                                                     TK6010
0102       1    SQ(I,J)=SS(I)                                                  TK6020
0103            DO 2 I=1,10                                                    TK6030
0104       2    TQ(I,J)=TS(I)                                                  TK6040
          C                                                                    TK6050
          C     PROCEDURE FOR COMPUTATION OF T-COEFFICIENTS.                   TK6060
          C                                                                    TK6070
0105              ST1=SX(1,J)+PX(1,J)*PQ(1,J)                                  TK6080
0106              ST2=SX(2,J)+PX(1,J)*PQ(2,J)                                  TK6090
0107              TX(1,J)=TX(1,J)+SX(1,J)*PQ(1,J)+PX(1,J)*SQ(1,J)              TK6100
0108              TX(2,J)=TX(2,J)+SX(2,J)*PQ(1,J)+SX(1,J)*PQ(2,J)+PX(1,J)*SQ(2,J) TK6110
0109              TX(3,J)=TX(3,J)+SX(2,J)*PQ(2,J)+PX(1,J)*SQ(3,J)              TK6120
0110              TX(4,J)=TX(4,J)+SX(4,J)*PQ(1,J)+PX(1,J)*SQ(4,J)              TK6130
0111              TX(5,J)=TX(5,J)+PX(1,J)*SQ(5,J)                              TK6140
0112              TX(6,J)=PX(1,J)*SQ(6,J)                                      TK6150
0113              TX(8,J)=TX(8,J)+SX(4,J)*PQ(2,J)                              TK6160
0114              SX(1,J)=ST1                                                  TK6170
0115       3      SX(2,J)=ST2                                                  TK6180
          C                                                                    TK6190
0116            RETURN                                                         TK6200
0117            END                                                            TK6210
```

298

```
0001          SUBROUTINE PSEUDO                                                 TK6220
        C                                                                       TK6230
        C                 SUBROUTINE PSEUDO PRODUCES THE PSEUDO-COEFFICIENTS    TK6240
        C     OF S AND T.   THE FOUR GROUPS OF AUXILIARY QUANTITIES ARE COMPUTED TK6250
        C     ONCE AND USED TO DETERMINE BOTH SETS OF PSEUDO-COEFFICIENTS.      TK6260
        C     AGAIN THE ENTIRE SUBROUTINE IS EXECUTED ONCE FOR EACH SURFACE.    TK6270
        C     THE PSEUDO-COEFFICIENTS ARE STORED IN THE LOCATIONS OF THE        TK6280
        C     CORRESPONDING SURFACE COEFFICIENTS.                               TK6290
        C                                                                       TK6300
0002          COMMON SYSTEM(9),KK                                               TK6310
0003          COMMON AA(25,16),A(3),YA(16),VA(16),IYA(16),YAP(16),YB(16),VB(16) TK6320
0004          COMMON IYB(16),YBP(16),ZA(16),WA(16),ZB(16),WB(16),AB(4,16)       TK6330
0005          COMMON PT(2,15),ST(6,15),TT(10,15),PS(2,15),SS(6,15),TS(10,15)    TK6340
0006          REAL IYA,IYB                                                      TK6350
0007          DIMENSION A1(3,3,15),A2(3,3,15),B1(4,4,15),B2(6,6,15),XX(6),XY(6) TK6360
0008          DO 7 J=1,KK                                                       TK6370
        C                                                                       TK6380
        C     PROCEDURE FOR COMPUTATION OF AUXILIARY QUANTITIES.                TK6390
        C                                                                       TK6400
0009             A1(1,1,J)=YA(J)*YA(J)                                          TK6410
0010             A1(1,2,J)=YA(J)*VA(J)                                          TK6420
0011             A1(1,3,J)=VA(J)*VA(J)                                          TK6430
0012             A1(2,1,J)=2.*YA(J)*YB(J)                                       TK6440
0013             A1(2,2,J)=YA(J)*VB(J)+YB(J)*VA(J)                              TK6450
0014             A1(2,3,J)=2.*VA(J)*VB(J)                                       TK6460
0015             A1(3,1,J)=YB(J)*YB(J)                                          TK6470
0016             A1(3,2,J)=YB(J)*VB(J)                                          TK6480
0017             A1(3,3,J)=VB(J)*VB(J)                                          TK6490
0018             A2(1,1,J)=ZA(J)*ZA(J)                                          TK6500
0019             A2(1,2,J)=ZA(J)*WA(J)                                          TK6510
0020             A2(1,3,J)=WA(J)*WA(J)                                          TK6520
0021             A2(2,1,J)=2.*ZA(J)*ZB(J)                                       TK6530
0022             A2(2,2,J)=ZA(J)*WB(J)+ZB(J)*WA(J)                              TK6540
0023             A2(2,3,J)=2.*WA(J)*WB(J)                                       TK6550
0024             A2(3,1,J)=ZB(J)*ZB(J)                                          TK6560
0025             A2(3,2,J)=ZB(J)*WB(J)                                          TK6570
0026             A2(3,3,J)=WB(J)*WB(J)                                          TK6580
0027             X=A1(2,2,J)+ YA(J)*VB(J)                                       TK6590
0028             Y=A1(2,2,J)+ YB(J)*VA(J)                                       TK6600
0029             B1(1,1,J)=A1(1,1,J)*YA(J)                                      TK6610
0030             B1(1,2,J)=A1(1,1,J)*VA(J)                                      TK6620
0031             B1(1,3,J)=A1(1,2,J)*VA(J)                                      TK6630
0032             B1(1,4,J)=A1(1,3,J)*VA(J)                                      TK6640
0033             B1(2,1,J)=3.*A1(1,1,J)*YB(J)                                   TK6650
0034             B1(2,2,J)=YA(J)*Y                                              TK6660
0035             B1(2,3,J)=VA(J)*X                                              TK6670
0036             B1(2,4,J)=3.*A1(1,3,J)*VB(J)                                   TK6680
0037             B1(3,1,J)=3.*A1(3,1,J)*YA(J)                                   TK6690
0038             B1(3,2,J)=YB(J)*X                                              TK6700
0039             B1(3,3,J)=VB(J)*Y                                              TK6710
0040             B1(3,4,J)=3.*VA(J)*A1(3,3,J)                                   TK6720
0041             B1(4,1,J)=YB(J)*A1(3,1,J)                                      TK6730
```

```
0042             B1(4,2,J)=VB(J)*A1(3,1,J)                                      TK6740
0043             B1(4,3,J)=YB(J)*A1(3,3,J)                                      TK6750
0044             B1(4,4,J)=VB(J)*A1(3,3,J)                                      TK6760
0045             DO 1 I=1,3                                                     TK6770
0046               II=I+3                                                       TK6780
0047               DO 1 L=1,3                                                   TK6790
0048                 LL=L+3                                                     TK6800
0049                 B2(L,I,J)=YA(J)*A2(L,I,J)                                  TK6810
0050                 B2(L,II,J)=VA(J)*A2(L,I,J)                                 TK6820
0051                 B2(LL,I,J)=YB(J)*A2(L,I,J)                                 TK6830
0052        1        B2(LL,II,J)=VB(J)*A2(L,I,J)                                TK6840
          C                                                                     TK6850
          C     PROCEDURE FOR COMPUTATION OF FIRST ORDER PSEUDO-COEFFICIENTS.   TK6860
          C                                                                     TK6870
0053             X=YA(J)*PS(1,J)+VA(J)*PS(2,J)                                  TK6880
0054             Y=YB(J)*PS(1,J)+VB(J)*PS(2,J)                                  TK6890
0055             PS(1,J)=X                                                      TK6900
0056             PS(2,J)=Y                                                      TK6910
0057             X=YA(J)*PT(1,J)+VA(J)*PT(2,J)                                  TK6920
0058             Y=YB(J)*PT(1,J)+VB(J)*PT(2,J)                                  TK6930
0059             PT(1,J)=X                                                      TK6940
0060             PT(2,J)=Y                                                      TK6950
          C                                                                     TK6960
          C     PROCEDURE FOR COMPUTATION OF SECOND ORDER PSEUDO-COEFFICIENTS.  TK6970
          C                                                                     TK6980
0061             DO 2 I=1,3                                                     TK6990
0062               II=I+3                                                       TK7000
0063               XY(I)=ST(1,J)*A1(I,1,J)+ST(2,J)*A1(I,2,J)+ST(3,J)*A1(I,3,J)  TK7010
0064               XY(II)=ST(4,J)*A2(I,1,J)+ST(5,J)*A2(I,2,J)+ST(6,J)*A2(I,3,J) TK7020
0065               XX(I)=SS(1,J)*A1(I,1,J)+SS(2,J)*A1(I,2,J)+SS(3,J)*A1(I,3,J)  TK7030
0066        2      XX(II)=SS(4,J)*A2(I,1,J)+SS(5,J)*A2(I,2,J)+SS(6,J)*A2(I,3,J) TK7040
0067             DO 3 I=1,6                                                     TK7050
0068               ST(I,J)=XY(I)                                                TK7060
0069        3      SS(I,J)=XX(I)                                                TK7070
          C                                                                     TK7080
          C     PROCEDURE FOR COMPUTATION OF THIRD ORDER PSEUDO-COEFFICIENTS.   TK7090
          C                                                                     TK7100
0070             DO 4 I=1,4                                                     TK7110
0071               XY(I)=TT(1,J)*B1(I,1,J)+TT(2,J)*B1(I,2,J)+TT(3,J)*B1(I,3,J)+ TK7120
               1   TT(7,J)*B1(I,4,J)                                            TK7130
0072        4      XX(I)=TS(1,J)*B1(I,1,J)+TS(2,J)*B1(I,2,J)+TS(3,J)*B1(I,3,J)+ TK7140
               1   TS(7,J)*B1(I,4,J)                                            TK7150
0073             DO 5 I=1,3                                                     TK7160
0074               TT(I,J)=XY(I)                                                TK7170
0075        5      TS(I,J)=XX(I)                                                TK7180
0076             TT(7,J)=XY(4)                                                  TK7190
0077             TS(7,J)=XX(4)                                                  TK7200
0078             DO 6 I=1,6                                                     TK7210
0079               XY(I)=TT(4,J)*B2(I,1,J)+TT(5,J)*B2(I,2,J)+TT(6,J)*B2(I,3,J)+ TK7220
               1   TT(8,J)*B2(I,4,J)+TT(9,J)*B2(I,5,J)+TT(10,J)*B2(I,6,J)       TK7230
0080        6      XX(I)=TS(4,J)*B2(I,1,J)+TS(5,J)*B2(I,2,J)+TS(6,J)*B2(I,3,J)+ TK7240
               1   TS(8,J)*B2(I,4,J)+TS(9,J)*B2(I,5,J)+TS(10,J)*B2(I,6,J)       TK7250
```

```
0081                DO 7 I=1,3                                                   TK7260
0082                   TT(I+3,J)=XY(I)                                           TK7270
0083                   TT(I+7,J)=XY(I+3)                                         TK7280
0084                   TS(I+3,J)=XX(I)                                           TK7290
0085          7        TS(I+7,J)=XX(I+3)                                         TK7300
             C                                                                   TK7310
0086                RETURN                                                       TK7320
0087                END                                                          TK7330
```

301

```
0001        SUBROUTINE GCOEFS
      C
      C               THE COMPUTATION OF THE SECOND AND THIRD ORDER G-
      C     COEFFICIENTS ACCOMPLISHED IN SUBROUTINE GCOEFS IS BASED ON THE
      C     FACT THAT THE SURFACE CONTRIBUTIONS TO THE QUASI-INVARIANTS ARE
      C     LINEAR COMBINATIONS TO THE TWO-VECTORS S AND T.  THE SUBROUTINE
      C     MAY BE DIVIDED INTO THE FOLLOWING FOUR BASIC STEPS - (1)  THE
      C     COMPUTATION OF THE COEFFICIENTS REQUIRED WHEN FORMING THE ABOVE
      C     LINEAR COMBINATIONS,  (2)  THE COMPUTATION OF THE SECOND ORDER
      C     G-COEFFICIENTS,  (3)  THE COMPUTATION OF THE THIRD ORDER G-
      C     COEFFICIENTS AND (4)  THE OUTPUT OF THE RESULTS OF STEPS (2)
      C     AND (3).
      C
      C               STEPS  (2) AND (3) CONSIST BASICALLY OF  (A)
      C     ITERATION,  (B)  THE COMPUTATION OF THE COEFFICIENTS FOR THE
      C     TWO-VECTORS S AND T AND FINALLY  (C)  THE COMPUTATION OF THE
      C     SURFACE CONTRIBUTIONS AND THE INTERMEDIATE COEFFICIENTS.  TWO
      C     SPECIAL SUBROUTINES (ITRATE AND ST2VEC) ARE PROVIDED FOR THIRD
      C     ORDER ITERATION AND THE DETERMINATION OF THE THIRD ORDER COEFFIC-
      C     IENTS OF THE TWO-VECTORS S AND T.  THESE TWO SUBROUTINES ARE
      C     RESPECTIVELY DIRECT TRANSLATIONS OF TABLES 37/1 AND 40/2 OF
      C     SANDS' THESIS.  TO FACILITATE THE IDENTIFICATION OF COEFFICIENTS
      C     IN THE PRINTOUT, THE LAST SECTION OF THIS PROGRAMME DETERMINES
      C     THE ORIGINAL NUMBERING OF THE COEFFICIENTS AS GIVEN IN SANDS'
      C     THESIS.
      C
0002        COMMON SYSTEM(9),KK,A1(2),C(16),AA(12,16),A2(7),IBY(15),NBX(15)
0003        COMMON NBY(15),SB(15),AC(6,15),A3(4)
0004        COMMON NY(16),NZ(16),YA(16),VA(16),IYA(16),YAP(16),YB(16),VB(16)
0005        COMMON IYB(16),YBP(16),ZA(16),WA(16),ZB(16),WB(16),IZA(16),IZB(16)
0006        COMMON DELTAM(15),DELTAS(15),GY,GZ
0007        COMMON PT(2,15),ST(6,15),TT(10,15),PS(2,15),SS(6,15),TS(10,15)
0008        COMMON SGYA(6,16),SGYB(6,16),SGZA(4,16),SGZB(4,16)
0009        COMMON TGYA(10,16),TGYB(10,16),TGZA(10,16),TGZB(10,16)
0010        COMMON SY(6,16),SV(6,16),SZ(4,16),SW(4,16)
0011        COMMON TY(10,16),TV(10,16),TZ(10,16),TW(10,16)
0012        COMMON LY(10),LZ(10)
0013        DIMENSION BAR(8,15),SGGYA(6,15),SGGYB(6,15),SGGZA(4,15),SGGZB(4,15
           1),TGGYA(10,15),TGGYB(10,15),TGGZA(10,15),TGGZB(10,15)
0014        DIMENSION SAY(6),TAY(6),SAZ(6),TAZ(6),SIZ(4,15)
0015        DIMENSION DY(10,15),DV(10,15),DZ(10,15),DW(10,15),SIY(6,15)
0016        REAL IBY,NY,NZ,IYA,IYB,IZA,IZB,NBX,NBY
      C
      C     PROCEDURE FOR COMPUTATION OF Y- AND V-BARRED COEFFICIENTS.
      C
0017        KP=KK+1
0018        DO 1 J=1,KK
0019          JJ=J+1
0020          X=-NY(JJ)*SB(J)
0021          BAR(1,J)=X*VA(JJ)
0022          BAR(2,J)=X*YAP(J)
0023          BAR(3,J)=X*VB(JJ)
```

```
                BAR(4,J)=X*YBP(J)
                X=-NZ(JJ)*SB(J)
                BAR(5,J)=X*WA(JJ)
                BAR(6,J)=X*ZA(J)
                BAR(7,J)=X*WB(JJ)
        1       BAR(8,J)=X*ZB(J)
      C
      C     PROCEDURE FOR COMPUTATION OF SECOND ORDER COEFFICIENTS.
      C
            DO 4 J=1,KK
              DO 2 I=1,6
                SAY(I)=SS(I,J)*IBY(J)
        2       TAY(I)=ST(I,J)*IBY(J)
              SAY(1)=SAY(1)+PS(1,J)*IYA(J)
              SAY(2)=SAY(2)+PS(1,J)*IYB(J)+PS(2,J)*IYA(J)
              SAY(3)=SAY(3)+PS(2,J)*IYB(J)
              TAY(1)=TAY(1)+PT(1,J)*IYA(J)
              TAY(2)=TAY(2)+PT(1,J)*IYB(J)+PT(2,J)*IYA(J)
              TAY(3)=TAY(3)+PT(2,J)*IYB(J)
              SAZ(1)=PS(1,J)*IZA(J)
              SAZ(2)=PS(1,J)*IZB(J)
              SAZ(3)=PS(2,J)*IZA(J)
              SAZ(4)=PS(2,J)*IZB(J)
              TAZ(1)=PT(1,J)*IZA(J)
              TAZ(2)=PT(1,J)*IZB(J)
              TAZ(3)=PT(2,J)*IZA(J)
              TAZ(4)=PT(2,J)*IZB(J)
      C
              DO 3 I=1,6
                SGGYA(I,J)=SAY(I)*BAR(2,J)+TAY(I)*BAR(1,J)
        3       SGGYB(I,J)=SAY(I)*BAR(4,J)+TAY(I)*BAR(3,J)
              DO 4 I=1,4
                SGGZA(I,J)=SAZ(I)*BAR(6,J)+TAZ(I)*BAR(5,J)
        4       SGGZB(I,J)=SAZ(I)*BAR(8,J)+TAZ(I)*BAR(7,J)
            DO 5 I=1,6
              SGYA(I,1)=0.
              SGYB(I,1)=0.
              DO 5 J=1,KK
                JJ=J+1
                SGYA(I,JJ)=SGGYA(I,J)+SGYA(I,J)
        5       SGYB(I,JJ)=SGGYB(I,J)+SGYB(I,J)
            DO 6 I=1,4
              SGZA(I,1)=0.
              SGZB(I,1)=0.
              DO 6 J=1,KK
                JJ=J+1
                SGZA(I,JJ)=SGGZA(I,J)+SGZA(I,J)
        6       SGZB(I,JJ)=SGGZB(I,J)+SGZB(I,J)
            DO 8 J=1,KP
              DO 7 I=1,6
                SV(I,J)=-(VA(J)*SGYB(I,J)-VB(J)*SGYA(I,J))/GY
        7       SY(I,J)=-(YA(J)*SGYB(I,J)-YB(J)*SGYA(I,J))/GY
```

```
0072              DO 8 I=1,4                                                 TK8380
0073                SW(I,J)=-(WA(J)*SGZB(I,J)-WB(J)*SGZA(I,J))/GZ             TK8390
0074       8        SZ(I,J)=-(ZA(J)*SGZB(I,J)-ZB(J)*SGZA(I,J))/GZ             TK8400
          C                                                                  TK8410
          C     PROCEDURE FOR COMPUTATION OF THIRD ORDER COEFFICIENTS.       TK8420
          C                                                                  TK8430
0075            CALL ITRATE(TS,SS,SGYA,SGZA,SGYB,SGZB,GY,GZ,KK)              TK8440
0076            CALL ITRATE(TT,ST,SGYA,SGZA,SGYB,SGZB,GY,GZ,KK)              TK8450
          C                                                                  TK8460
0077            DO 10 J=1,KK                                                 TK8470
0078              DO 9 I=1,6                                                 TK8480
0079       9        SIY(I,J)=C(J)*SY(I,J)+NBX(J)*SV(I,J)                     TK8490
0080              DO 10 I=1,4                                                TK8500
0081      10        SIZ(I,J)=C(J)*SZ(I,J)+NBX(J)*SW(I,J)                     TK8510
0082            DO 11 J=1,KK                                                 TK8520
0083              DO 11 I=1,6                                                TK8530
0084                SS(I,J)=SS(I,J)-(SGYB(I,J)*PS(1,J)-SGYA(I,J)*PS(2,J))/GY TK8540
0085      11        ST(I,J)=ST(I,J)-(SGYB(I,J)*PT(1,J)-SGYA(I,J)*PT(2,J))/GY TK8550
0086            CALL ST2VEC(DY,DZ,TS,SS,PS,SIY,SIZ,IYA,IYB,IZA,IZB,IBY,KK)   TK8560
0087            CALL ST2VEC(DV,DW,TT,ST,PT,SIY,SIZ,IYA,IYB,IZA,IZB,IBY,KK)   TK8570
          C                                                                  TK8580
0088            DO 12 J=1,KK                                                 TK8590
0089              DO 12 I=1,10                                               TK8600
0090                TGGYA(I,J)=BAR(2,J)*DY(I,J)+BAR(1,J)*DV(I,J)             TK8610
0091                TGGYB(I,J)=BAR(4,J)*DY(I,J)+BAR(3,J)*DV(I,J)             TK8620
0092                TGGZA(I,J)=BAR(6,J)*DZ(I,J)+BAR(5,J)*DW(I,J)             TK8630
0093      12        TGGZB(I,J)=BAR(8,J)*DZ(I,J)+BAR(7,J)*DW(I,J)             TK8640
0094            DO 13 I=1,10                                                 TK8650
0095              TGYA(I,1)=0.                                               TK8660
0096              TGYB(I,1)=0.                                               TK8670
0097              TGZA(I,1)=0.                                               TK8680
0098              TGZB(I,1)=0.                                               TK8690
0099              DO 13 J=1,KK                                               TK8700
0100                JJ=J+1                                                   TK8710
0101                TGYA(I,JJ)=TGYA(I,J)+TGGYA(I,J)                          TK8720
0102                TGYB(I,JJ)=TGYB(I,J)+TGGYB(I,J)                          TK8730
0103                TGZA(I,JJ)=TGZA(I,J)+TGGZA(I,J)                          TK8740
0104      13        TGZB(I,JJ)=TGZB(I,J)+TGGZB(I,J)                          TK8750
0105            DO 14 J=1,KP                                                 TK8760
0106              DO 14 I=1,10                                               TK8770
0107              TY(I,J)=-(YA(J)*TGYB(I,J)-YB(J)*TGYA(I,J))/GY              TK8780
0108              TV(I,J)=-(VA(J)*TGYB(I,J)-VB(J)*TGYA(I,J))/GY              TK8790
0109              TZ(I,J)=-(ZA(J)*TGZB(I,J)-ZB(J)*TGZA(I,J))/GZ              TK8800
0110      14      TW(I,J)=-(WA(J)*TGZB(I,J)-WB(J)*TGZA(I,J))/GZ              TK8810
          C                                                                  TK8820
          C     PROCEDURE FOR DETERMINING THE NUMBERING OF THE COEFFICIENTS. TK8830
          C                                                                  TK8840
0111            LY(1)=1                                                      TK8850
0112            LY(2)=2                                                      TK8860
0113            LY(3)=5                                                      TK8870
0114            LY(4)=8                                                      TK8880
0115            LY(5)=9                                                      TK8890
```

304

```
0001            SUBROUTINE ITRATE(Q,SQ,SGYA,SGZA,SGYB,SGZB,Y,Z,KK)                    TK9080
          C                                                                           TK9090
0002            DIMENSION Q(10,15),SGYA(6,15),SGYB(6,15),SGZA(4,15),SGZB(4,15),      TK9100
               1QQ(10),SQ(6,15)                                                       TK9110
          C                                                                           TK9120
0003            GY=-Y                                                                 TK9130
0004            GZ=-Z                                                                 TK9140
0005            DO 1 J=1,KK                                                           TK9150
0006               QQ(1)=(2.*SGYB(1,J)*SQ(1,J)-SGYA(1,J)*SQ(2,J))/GY                  TK9160
0007               QQ(2)=(2.*SGYB(2,J)*SQ(1,J)+(SGYB(1,J)-SGYA(2,J))*SQ(2,J)-2.*     TK9170
               1          SGYA(1,J)*SQ(3,J))/GY                                       TK9180
0008               QQ(3)=(2.*SGYB(3,J)*SQ(1,J)+(SGYB(2,J)-SGYA(3,J))*SQ(2,J)-2.*     TK9190
               1          SGYA(2,J)*SQ(3,J))/GY                                       TK9200
0009               QQ(4)=(2.*SGYB(4,J)*SQ(1,J)-SGYA(4,J)*SQ(2,J))/GY+(2.*SGZB(1,J)*  TK9210
               1          SQ(4,J)-SGZA(1,J)*SQ(5,J))/GZ                               TK9220
0010               QQ(5)=(2.*SGYB(5,J)*SQ(1,J)-SGYA(5,J)*SQ(2,J))/GY+(2.*SGZB(2,J)*  TK9230
               1          SQ(4,J)+(SGZB(1,J)-SGZA(2,J))*SQ(5,J)-2.*SGZA(1,J)*SQ(6,J)  TK9240
               2          )/GZ                                                        TK9250
0011               QQ(6)=(2.*SGYB(6,J)*SQ(1,J)-SGYA(6,J)*SQ(2,J))/GY+(SGZB(2,J)*     TK9260
               1          SQ(5,J)-2.*SGZA(2,J)*SQ(6,J))/GZ                            TK9270
0012               QQ(7)=(SGYB(3,J)*SQ(2,J)-2.*SGYA(3,J)*SQ(3,J))/GY                  TK9280
0013               QQ(8)=(SGYB(4,J)*SQ(2,J)-2.*SGYA(4,J)*SQ(3,J))/GY+(2.*SGZB(3,J)*  TK9290
               1          SQ(4,J)-SGZA(3,J)*SQ(5,J))/GZ                               TK9300
0014               QQ(9)=(SGYB(5,J)*SQ(2,J)-2.*SGYA(5,J)*SQ(3,J))/GY+(2.*SGZB(4,J)*  TK9310
               1          SQ(4,J)+(SGZB(3,J)-SGZA(4,J))*SQ(5,J)-2.*SGZA(3,J)*SQ(6,J)  TK9320
               2          )/GZ                                                        TK9330
0015               QQ(10)=(SGYB(6,J)*SQ(2,J)-2.*SGYA(6,J)*SQ(3,J))/GY+(SGZB(4,J)*    TK9340
               1           SQ(5,J)-2.*SGZA(4,J)*SQ(6,J))/GZ                           TK9350
0016               DO 1 I=1,10                                                        TK9360
0017         1       Q(I,J)=Q(I,J)+QQ(I)                                              TK9370
          C                                                                           TK9380
0018            RETURN                                                                TK9390
0019            END                                                                   TK9400
```

```
            LY(6)=10
            LY(7)=11
            DO 15 I=8,10
   15         LY(I)=I+6
            LZ(1)=3
            LZ(2)=4
            LZ(3)=6
            LZ(4)=7
            LZ(5)=12
            LZ(6)=13
            DO 16 I=7,10
   16         LZ(I)=10+I
C
            CALL COEFOP(1,SGYA,SGZA,TGYA,TGZA,LY,LZ)
            CALL COEFOP(2,SGYB,SGZB,TGYB,TGZB,LY,LZ)
C
            RETURN
            END
```

```
0001                SUBROUTINE ST2VEC(QY,QZ,TQ,SQ,PQ,SIY,SIZ,IYA,IYB,IZA,IZB,IBY,KK)
              C
0002                DIMENSION QY(10,15),QZ(10,15),TQ(10,15),SQ(6,15),PQ(2,15),SIY(6,15
                   1),SIZ(4,15),IYA(16),IYB(16),IZA(16),IZB(16),IBY(15)
0003                REAL IYA,IZA,IYB,IZB,IBY
              C
0004                DO 4 J=1,KK
0005                  QY(1,J)=SQ(1,J)*IYA(J)+PQ(1,J)*SIY(1,J)
0006                  QY(2,J)=SQ(1,J)*IYB(J)+SQ(2,J)*IYA(J)+PQ(1,J)*SIY(2,J)+PQ(2,J)*
                   1              SIY(1,J)
0007                  QY(3,J)=SQ(2,J)*IYB(J)+SQ(3,J)*IYA(J)+PQ(1,J)*SIY(3,J)+PQ(2,J)*
                   1              SIY(2,J)
0008                  DO 1 I=4,6
0009          1         QY(I,J)=SQ(I,J)*IYA(J)+PQ(1,J)*SIY(I,J)
0010                  DO 2 I=7,10
0011                      II=I-4
0012          2         QY(I,J)=SQ(II,J)*IYB(J)+PQ(2,J)*SIY(II,J)
0013                  DO 3 I=1,10
0014          3       QY(I,J)=QY(I,J)+IBY(J)*TQ(I,J)
              C
0015                  QZ(1,J)=SQ(1,J)*IZA(J)+PQ(1,J)*SIZ(1,J)
0016                  QZ(2,J)=SQ(1,J)*IZB(J)+PQ(1,J)*SIZ(2,J)
0017                  QZ(3,J)=SQ(2,J)*IZA(J)+PQ(1,J)*SIZ(3,J)+PQ(2,J)*SIZ(1,J)
0018                  QZ(4,J)=SQ(2,J)*IZB(J)+PQ(1,J)*SIZ(4,J)+PQ(2,J)*SIZ(2,J)
0019                  QZ(5,J)=SQ(3,J)*IZA(J)+PQ(2,J)*SIZ(3,J)
0020                  QZ(6,J)=SQ(3,J)*IZB(J)+PQ(2,J)*SIZ(4,J)
0021                  QZ(7,J)=SQ(4,J)*IZA(J)
0022                  QZ(8,J)=SQ(4,J)*IZB(J)+SQ(5,J)*IZA(J)
0023                  QZ(9,J)=SQ(5,J)*IZB(J)+SQ(6,J)*IZA(J)
0024          4       QZ(10,J)=SQ(6,J)*IZB(J)
0025                RETURN
              C
0026                END
```

```
0001          SUBROUTINE COEFOP(LJK,SGYA,SGZA,TGYA,TGZA,LY,LZ)
        C
        C                SUBROUTINE COEFOP IS USED TO PRINT THE PARABASAL,
        C     G- OR H-COEFFICIENTS.  THE OPTIONS GOVERNED BY LJK ARE -
        C
        C     LJK = 1    THE PARABASAL COEFFICIENTS AND THE SURFACE CONTRI-
        C                BUTIONS TO THE A-COMPONENTS OF THE G-COEFFICIENTS
        C                AND THE CORRESPONDING FINAL COEFFICIENTS ARE PRINTED,
        C
        C     LJK = 2    THE B-COMPONENTS OF THE SURFACE CONTRIBUTIONS TO
        C                THE G-COEFFICIENTS AND THE CORRESPONDING FINAL
        C                COEFFICIENTS ARE PRINTED,
        C
        C     LJK = 3    THE SURFACE CONTRIBUTIONS TO THE FINAL IMAGE HEIGHT
        C                COEFFICIENTS AND THE CORRESPONDING FINAL COEFFIC-
        C                IENTS ARE PRINTED.
        C
        C     THE SURFACE CONTRIBUTIONS ARE COMPUTED BY TAKING THE DIFFERENCES
        C     BETWEEN THE INTERMEDIATE COEFFICIENTS.
        C
        C     NOTE - IF THE BASE-RAY IS ON THE VERGE OF MISSING A SURFACE OR OF
        C     SUFFERING TOTAL INTERNAL REFLECTION, THE F-FORMAT IN THE OUT-PUT
        C     SHOULD BE CHANGED TO E-FORMAT OTHERWISE THE FORMAT SPECIFICATION
        C     WILL BE EXCEEDED.
        C
0002          COMMON SYSTEM(9),KK,FOCAL,STPNUM,AA(12,15),A1(17),SYB,TYB
0003          COMMON AB(13,15),A2(7),YA(16),VA(16),IYA(16),YAP(16),YB(16),VB(16)
0004          COMMON IYB(16),YBP(16),ZA(16),WA(16),ZB(16),WB(16)
0005          DIMENSION  SGGYA(6,15),SGGZA(4,15),TGGYA(10,15),TGGZA(10,15)
0006          DIMENSION SGYA(6,16),SGZA(4,16),TGYA(10,16),TGZA(10,16)
0007          DIMENSION LY(10),LZ(10)
0008          REAL IYA,IYB
        C
        C     PROCEDURE FOR DETERMINING SURFACE CONTRIBUTIONS.
        C
0009          KP=KK+1
0010          DO 22 J=1,KK
0011            JJ=J+1
0012            DO 20 I=1,6
0013     20       SGGYA(I,J)=SGYA(I,JJ)-SGYA(I,J)
0014            DO 21 I=1,4
0015     21       SGGZA(I,J)=SGZA(I,JJ)-SGZA(I,J)
0016            DO 22 I=1,10
0017              TGGYA(I,J)=TGYA(I,JJ)-TGYA(I,J)
0018     22       TGGZA(I,J)=TGZA(I,JJ)-TGZA(I,J)
        C
0019          WRITE(3,420) SYSTEM,FOCAL,STPNUM,SYB,TYB
0020          GO TO (10,40,50),LJK
        C
        C     PROCEDURE FOR PRINTING PARABASAL COEFFICIENTS.
        C
0021     10   WRITE(3,400)
```

```
0022              WRITE(3,401) (J,J=1,KP)                                        TK0260
0023              WRITE(3,410)                                                   TK0270
0024              WRITE(3,402)  (YA(J),J=1,KP)                                   TK0280
0025              WRITE(3,403)  (VA(J),J=1,KP)                                   TK0290
0026              WRITE(3,404)  (YB(J),J=1,KP)                                   TK0300
0027              WRITE(3,405)  (VB(J),J=1,KP)                                   TK0310
0028              WRITE(3,410)                                                   TK0320
0029              WRITE(3,406)  (ZA(J),J=1,KP)                                   TK0330
0030              WRITE(3,407)  (WA(J),J=1,KP)                                   TK0340
0031              WRITE(3,408)  (ZB(J),J=1,KP)                                   TK0350
0032              WRITE(3,409)  (WB(J),J=1,KP)                                   TK0360
          C                                                                      TK0370
          C       PROCEDURE FOR PRINTING 2ND AND 3RD ORDER G-COEFFS., A-COMPONENTS.  TK0380
          C                                                                      TK0390
0033              WRITE(3,411)                                                   TK0400
0034       30     WRITE(3,401)(J,J=1,KK)                                         TK0410
0035              WRITE(3,410)                                                   TK0420
0036              DO 1 I=1,6                                                     TK0430
0037        1       WRITE(3,412) LY(I),(SGGYA(I,J),J=1,KK),SGYA(I,KP)            TK0440
0038              WRITE(3,410)                                                   TK0450
0039              DO 2 I=1,4                                                     TK0460
0040        2       WRITE(3,413) LZ(I),(SGGZA(I,J),J=1,KK),SGZA(I,KP)            TK0470
0041              WRITE(3,410)                                                   TK0480
0042              DO 3 I=1,10                                                    TK0490
0043        3       WRITE(3,414) LY(I),(TGGYA(I,J),J=1,KK),TGYA(I,KP)            TK0500
0044              WRITE(3,410)                                                   TK0510
0045              DO 4 I=1,10                                                    TK0520
0046        4       WRITE(3,415) LZ(I),(TGGZA(I,J),J=1,KK),TGZA(I,KP)            TK0530
          C                                                                      TK0540
0047              RETURN                                                         TK0550
          C                                                                      TK0560
          C       PROCEDURE FOR PRINTING 2ND AND 3RD ORDER G-COEFFS., B-COMPONENTS.  TK0570
          C                                                                      TK0580
0048       40     WRITE(3,416)                                                   TK0590
0049              GO TO 30                                                       TK0600
          C                                                                      TK0610
          C       PROCEDURE FOR PRINTING SURFACE CONTRIBUTIONS TO H-COEFFICIENTS.   TK0620
          C                                                                      TK0630
0050       50     WRITE(3,417)                                                   TK0640
0051              GO TO 30                                                       TK0650
          C                                                                      TK0660
0052      400     FORMAT(//' PARABASAL COEFFICIENTS'/1X,22('=')/)                TK0670
0053      401     FORMAT(14X,I2,10(8X,I2)/)                                      TK0680
0054      402     FORMAT(3X,'YA',4X,11F10.5)                                     TK0690
0055      403     FORMAT(3X,'VA',4X,11F10.5)                                     TK0700
0056      404     FORMAT(3X,'YB',4X,11F10.5)                                     TK0710
0057      405     FORMAT(3X,'VB',4X,11F10.5)                                     TK0720
0058      406     FORMAT(3X,'ZA',4X,11F10.5)                                     TK0730
0059      407     FORMAT(3X,'WA',4X,11F10.5)                                     TK0740
0060      408     FORMAT(3X,'ZB',4X,11F10.5)                                     TK0750
0061      409     FORMAT(3X,'WB',4X,11F10.5)                                     TK0760
0062      410     FORMAT(1X)                                                     TK0770
```

```
0063        411  FORMAT(/ ' SECOND, THIRD ORDER G-COEFFICIENTS, A-COMPONENTS -'/
                149('=')/)
0064        412  FORMAT(3X,'SGY',I2,1X,11F10.5)
0065        413  FORMAT(3X,'SGZ',I2,1X,11F10.5)
0066        414  FORMAT(3X,'TGY',I2,1X,11F10.5)
0067        415  FORMAT(3X,'TGZ',I2,1X,11F10.5)
0068        416  FORMAT(//' SECOND, THIRD ORDER G-COEFFICIENTS, B-COMPONENTS -'/
                149('=')/)
0069        417  FORMAT(//' SECOND AND THIRD ORDER IMAGE HEIGHT COEFFICIENTS -'/
                149('=')/)
0070        420  FORMAT('1',9A4,4X,'FOCAL LENGTH =',F6.4,' ,STOP =',F6.2/1X,
                1'BASE RAY IS MERIDIONAL - SBY =',E12.5,', TBY = ',E12.5)
0071             END
```

```
0001          SUBROUTINE HDCOEF(MUY,MUZ,IJK)
        C
        C               IF IJK = 1, SUBROUTINE HDCOEF COMPUTES THE IMAGE
        C     HEIGHT COEFFICIENTS IN THE IDEAL IMAGE PLANE OF THE SYSTEM AND
        C     THE DIAPHRAGM COEFFICIENTS.  THESE, AS WELL AS THE V-COEFFIC-
        C     IENTS IN THE IMAGE SPACE, ARE PRINTED AND PUNCHED ONTO CARDS.
        C     (THE DETAILS OF THE PUNCHING HAVE BEEN GIVEN IN THE MAIN PRO-
        C     GRAMME.)  IF IJK = 2 THE SURFACE CONTRIBUTIONS TO THE IMAGE
        C     HEIGHT COEFFICIENTS ARE ALSO COMPUTED AND PRINTED.
        C
0002          COMMON SYSTEM(9),KK,FOCAL,STPNUM,AA(4,16),ENPUP,OBJECT,JD,DSTPP
0003          COMMON AB(3,15),AI(2),VAA(16),VAB(16),YAA(16),YAB(16),IMAGE(16)
0004          COMMON RHOSTP,EXP,SYB,TYB,VBB(16),AC(4,15),XB(15),ACC(3,15),A2(2)
0005          COMMON HBY,EPSBY,AD(4,15),A3(2)
0006          COMMON YA(16),VA(16),AE(2,16),YB(16),VB(16),AF(2,16),ZA(16),WA(16)
0007          COMMON ZB(16),WB(16),AG(2,16),DELTAM(15),DELTAS(15),GY,GZ
0008          COMMON AH(36,15)
0009          COMMON SGYA(6,16),SGYB(6,16),SGZA(4,16),SGZB(4,16)
0010          COMMON TGYA(10,16),TGYB(10,16),TGZA(10,16),TGZB(10,16)
0011          COMMON SY(6,16),SV(6,16),SZ(4,16),SW(4,16)
0012          COMMON TY(10,16),TV(10,16),TZ(10,16),TW(10,16)
0013          COMMON LY(10),LZ(10)
0014          COMMON PEY(2),PEZ(2),SEY(6),SEZ(4),TEY(10),TEZ(10)
0015          COMMON PDY(2),PDZ(2),SDY(6),SDZ(4),TDY(10),TDZ(10)
0016          DIMENSION SHY(6,16),SHZ(4,16),THY(10,16),THZ(10,16)
0017          REAL IMAGE,MUY,MUZ,IMAGEB
        C
        C     PROCEDURE FOR COMPUTING FINAL IMAGE HEIGHT COEFFICIENTS.
        C
0018          KP=KK+1
0019          IMAGEB=IMAGE(KK)-XB(KK)
0020          PEY(1)=YA(KP)+IMAGEB*VA(KP)
0021          PEY(2)=YB(KP)+IMAGEB*VB(KP)
0022          PEZ(1)=ZA(KP)+IMAGEB*WA(KP)
0023          PEZ(2)=ZB(KP)+IMAGEB*WB(KP)
0024          DO 1 I=1,6
0025        1   SEY(I)=SGYA(I,KP)/MUY-DELTAM(KK)*SV(I,KP)
0026          DO 2 I=1,4
0027        2   SEZ(I)=SGZA(I,KP)/MUZ-DELTAS(KK)*SW(I,KP)
0028          DO 3 I=1,10
0029            TEY(I)=TGYA(I,KP)/MUY-DELTAM(KK)*TV(I,KP)
0030        3   TEZ(I)=TGZA(I,KP)/MUZ-DELTAS(KK)*TW(I,KP)
        C
        C     PROCEDURE FOR COMPUTING SURFACE CONTRIBUTIONS TO H-COEFFICIENTS.
        C
0031          GO TO (31,30),IJK
0032       30 DO 20 J=1,KP
0033            DO 21 I=1,6
0034       21     SHY(I,J)=SGYA(I,J)/MUY-DELTAM(KK)*SV(I,J)
0035            DO 22 I=1,4
0036       22     SHZ(I,J)=SGZA(I,J)/MUZ-DELTAS(KK)*SW(I,J)
0037            DO 20 I=1,10
```

```
0038               THY(I,J)=TGYA(I,J)/MUY-DELTAM(KK)*TV(I,J)
0039       20      THZ(I,J)=TGZA(I,J)/MUZ-DELTAS(KK)*TW(I,J)
0040            CALL COEFOP(3,SHY,SHZ,THY,THZ,LY,LZ)
          C
          C     PROCEDURE FOR COMPUTING DIAPHRAGM COEFFICIENTS.
          C
0041            DSTP = DSTPP-XB(JD)
0042       31   PDY(1)=YA(JD)+DSTP*VA(JD)
0043            PDY(2)=YB(JD)+DSTP*VB(JD)
0044            PDZ(1)=ZA(JD)+DSTP*WA(JD)
0045            PDZ(2)=ZB(JD)+DSTP*WB(JD)
0046            DO 4 I=1,6
0047        4     SDY(I)=SY(I,JD)+DSTP*SV(I,JD)
0048            DO 5 I=1,4
0049        5     SDZ(I)=SZ(I,JD)+DSTP*SW(I,JD)
0050            DO 6 I=1,10
0051              TDY(I)=TY(I,JD)+DSTP*TV(I,JD)
0052        6     TDZ(I)=TZ(I,JD)+DSTP*TW(I,JD)
          C
          C     PROCEDURE FOR PRINTING H- V- AND DIAPHRAGM COEFFICIENTS.
          C
0053            WRITE(3,400) SYSTEM,FOCAL,STPNUM,SYB,TYB
0054            WRITE(3,401)
0055            WRITE(3,402) PEY(1),PEZ(1),VA(KP),WA(KP),PDY(1),PDZ(1)
0056            WRITE(3,4021)PEY(2),PEZ(2),VB(KP),WB(KP),PDY(2),PDZ(2)
0057            WRITE(3,405)
0058            DO 7 J=1,4
0059        7     WRITE(3,403) LY(J),SEY(J),LZ(J),SEZ(J),LY(J),SV(J,KP),LZ(J),
               1  SW(J,KP),LY(J),SDY(J),LZ(J),SDZ(J)
0060            DO 8 J=5,6
0061        8     WRITE(3,4031)LY(J),SEY(J),LY(J),SV(J,KP),LY(J),SDY(J)
0062            WRITE(3,405)
0063            DO 9 J=1,10
0064        9     WRITE(3,404) LY(J),TEY(J),LZ(J),TEZ(J),LY(J),TV(J,KP),LZ(J),
               1  TW(J,KP),LY(J),TDY(J),LZ(J),TDZ(J)
          C
          C     PROCEDURE FOR PUNCHING DATA REQUIRED FOR PREDICTING ABERRATIONS.
          C
0065            WRITE(2,500) SYSTEM,FOCAL,STPNUM,SYB,TYB
0066            WRITE(2,501) HBY,EPSBY,RHOSTP,VBB(KP),VBB(1)
0067            WRITE(2,501) VAA(1),YAA(1),VAB(1),YAB(1)
0068            WRITE(2,501) PEY,PEZ,SEY,SEZ,TEY,TEZ,VA(KP),VB(KP),WA(KP),WB(KP),
               1(SV(I,KP),I=1,6),(SW(I,KP),I=1,4),(TV(I,KP),I=1,10),(TW(I,KP),
               2I=1,10),PDY,PDZ,SDY,SDZ,TDY,TDZ
          C
0069            RETURN
          C
0070      400   FORMAT('1',9A4,4X,'FOCAL LENGTH =',F7.4,', STOP =',F4.1/' BASE RAY
               1 IS MERIDIONAL - SBY = ',E12.5,', TBY = ',E12.5///)
0071      401   FORMAT(4X,'FINAL IMAGE-HEIGHT COEFFICIENTS',11X,'FINAL V-, W-COFFI
               1CIENTS',13X,'DIAPHRAGM COEFFICIENTS'/4X,31('='),11X,23('='),13X,22
               2('=')/)
```

312

```
0072      402  FORMAT(' HYA    =',F10.5,', HZA    =',F10.5,', VA    =',F10.5,', WA     TK3070
              1 =',F10.5,', YAD    =',F10.5,', ZAD    =',F10.5)                    TK3080
0073     4021  FORMAT(' HYB    =',F10.5,', HZB    =',F10.5,', VB    =',F10.5,', WA     TK3090
              1 =',F10.5,', YBD    =',F10.5,', ZBD    =',F10.5)                    TK3100
0074      403  FORMAT(' SHY',I2,' =',F10.5,', SHZ',I2,' =',F10.5,', SV',I2,' =',   TK3110
              1F10.5,', SW',I2,' =',F10.5,', SY',I2,'D =',F10.5,', SZ',I2,'D =',   TK3120
              2F10.5)                                                              TK3130
0075      404  FORMAT(' THY',I2,' =',F10.5,', THZ',I2,' =',F10.5,', TV',I2,' =',F  TK3140
              110.5,', TW',I2,' =',F10.5,', TY',I2,'D =',F10.5,', TZ',I2,'D =',    TK3150
              2F10.5)                                                              TK3160
0076     4031  FORMAT(' SHY',I2,' =',F10.5,',',20X,'SV',I2,' =',F10.5,',',19X,     TK3170
              1'SY',I2,'D =',F10.5,',')                                            TK3180
0077      405  FORMAT(2X)                                                          TK3190
0078      500  FORMAT('V',1X,9A4,3X,F8.5,1X,F4.1,2(1X,E12.5))                      TK3200
0079      501  FORMAT('V',6(1X,E12.5))                                             TK3210
0080           END                                                                 TK3220
```

313

```
0001          SUBROUTINE PRBRAY
        C
        C              FOR THE PURPOSES OF SUBROUTINE PRBRAY, THE BASE-
        C     RAY IS REGARDED AS BEING PRINCIPAL IF RHODPG IS LESS THAN 0.0001.
        C     IF THIS IS THE CASE, THE SERIES FOR THE DIAPHRAGM COORDINATES
        C     IS INVERTED TO GIVE THE FIRST AND SECOND ORDER ENTRANCE PUPIL
        C     COEFFICIENTS AND SELECTED COEFFICIENTS OF THE THIRD ORDER.  THE
        C     COEFFICIENTS FOR THE IMAGE HEIGHT OF THE PRINCIPAL RAY ARE DE-
        C     TERMINED AND, TOGETHER WITH THE ENTRANCE PUPIL COEFFICIENTS,
        C     PRINTED.  IF THE BASE-RAY IS NOT PRINCIPAL CONTROL IS RETURNED
        C     TO THE MAIN PROGRAMME AND A NEW BASE-RAY IS CONSIDERED.
        C
0002          COMMON AA(21,16),A1(14),RHODPG,AB(138,16),AC(38,15),LY(10),LZ(10)
0003          COMMON PEY(2),PEZ(2),SEY(6),SEZ(4),TEY(10),TEZ(10)
0004          COMMON PDY(2),PDZ(2),SDY(6),SDZ(4),TDY(10),TDZ(10)
0005          DIMENSION PSY(2),PSZ(2),SSY(6),SSZ(4),TSY(10),TSZ(10)
        C
0006          IF(ABS(RHODPG)-0.0001) 2,2,1
0007        1 RETURN
0008        2 KP=KK+1
        C
        C     PROCEDURE FOR COMPUTATION OF ENTRANCE PUPIL COEFFICIENTS.
        C
0009          PSY(1)=1./PDY(1)
0010          PSZ(1)=1./PDZ(1)
0011          PSY(2)=-PDY(2)/PDY(1)
0012          PSZ(2)=-PDZ(2)/PDZ(1)
        C
0013          SSY(1)=-SDY(1)*PSY(1)*PSY(1)
0014          SSY(2)=-PSY(1)*(2.*SDY(1)*PSY(2)+SDY(2))
0015          SSY(3)=-PSY(2)*PSY(2)*SDY(1)-PSY(2)*SDY(2)-SDY(3)
0016          SSY(4)=-SDY(4)*PSZ(1)*PSZ(1)
0017          SSY(5)=-PSZ(1)*(2.*SDY(4)*PSZ(2)+SDY(5))
0018          SSY(6)=-PSZ(2)*PSZ(2)*SDY(4)-PSZ(2)*SDY(5)-SDY(6)
0019          DO 4 I=1,6
0020        4   SSY(I)=SSY(I)/PDY(1)
0021          SSZ(1)=-SDZ(1)*PSY(1)*PSZ(1)
0022          SSZ(2)=-PSY(1)*(SDZ(1)*PSZ(2)+SDZ(2))
0023          SSZ(3)=-PSZ(1)*(PSY(2)*SDZ(1)+SDZ(3))
0024          SSZ(4)=-PSY(2)*(SDZ(1)*PSZ(2)+SDZ(2))-SDZ(3)*PSZ(2)-SDZ(4)
0025          DO 5 I=1,4
0026        5 SSZ(I)=SSZ(I)/PDZ(1)
        C
0027          X=2.*SDY(1)*PSY(2)+SDY(2)
0028          TSY(7)=-(X*SSY(3)+TDY(1)*PSY(2)**3+TDY(2)*PSY(2)**2+TDY(3)*PSY(2)
             1+TDY(7))/PDY(1)
0029          TSY(10)=-(X*SSY(6)+(2.*SDY(4)*PSZ(2)+SDY(5))*SSZ(4)+(TDY(4)*
             1PSZ(2)**2+TDY(5)*PSZ(2)+TDY(6))*PSY(2)+TDY(8)*PSZ(2)**2+TDY(9)
             2*PSZ(2)+TDY(10))/PDY(1)
0030          X=SDZ(1)*PSZ(2)+SDZ(2)
0031          TSZ(6)=-(X*SSY(3)+(SDZ(1)*PSY(2)+SDZ(3))*SSZ(4)+(TDZ(1)*PSZ(2)+
             1TDZ(2))*PSY(2)**2+(TDZ(3)*PSZ(2)+TDZ(4))*PSY(2)+TDZ(5)*PSZ(2)+
```

```
            2TDZ(6))/PDZ(1)
0032         TSZ(10)=-(X*SSY(6)+TDZ(7)*PSZ(2)**3+TDZ(8)*PSZ(2)**2+TDZ(9)*
            1PSZ(2)+TDZ(10))/PDZ(1)
      C
      C      PROCEDURE FOR COMPUTATION OF H-COEFFICIENTS (PRINCIPAL RAYS).
      C
0033         PHY2=PEY(1)*PSY(2)+PEY(2)
0034         SHY3=PEY(1)*SSY(3)+SEY(1)*PSY(2)**2+SEY(2)*PSY(2)+SEY(3)
0035         SHY6=PEY(1)*SSY(6)+SEY(4)*PSY(2)**2+SEY(5)*PSY(2)+SEY(6)
0036         X=2.*SEY(1)*PSY(2)+SEY(2)
0037         THY7 =PEY(1)*TSY(7)+X*SSY(3)+TEY(1)*PSY(2)**3+TEY(2)*PSY(2)**2
            1+TEY(3)*PSY(2)+TEY(7)
0038         THY10=PEY(1)*TSY(10)+X*SSY(6)+(2.*SEY(4)*PSZ(2)+SEY(5))*SSZ(4)+
            1(TEY(4)*PSZ(2)**2+TEY(5)*PSZ(2)+TEY(6))*PSY(2)+TEY(8)*PSZ(2)**2
            2+TEY(9)*PSZ(2)+TEY(10)
      C
0039         PHZ2=PEZ(1)*PSZ(2)+PEZ(2)
0040         X=SEZ(1)*PSZ(2)+SEZ(2)
0041         SHZ4=PEZ(1)*SSZ(4)+X*PSY(2)+SEZ(3)*PSZ(2)+SEZ(4)
0042         THZ6 =PEZ(1)*TSZ(6)+X*SSY(3)+(SEZ(1)*PSY(2)+SEZ(3))*SSZ(4)
            1+(TEZ(1)*PSZ(2)+TEZ(2))*PSY(2)**2+(TEZ(3)*PSZ(2)+TEZ(4))*PSY(2)
            2+TEZ(5)*PSZ(2)+TEZ(6)
0043         THZ10=PEZ(1)*TSZ(10)+X*SSY(6)+TEZ(7)*PSZ(2)**3+TEZ(8)*PSZ(2)**2
            1+TEZ(9)*PSZ(2)+TEZ(10)
      C
      C      PROCEDURE FOR PRINTING ENTRANCE PUPIL AND PRINCIPAL RAY H-COEFFS.
      C
0044         WRITE(3,400)
0045         WRITE(3,401) PSY(1),PSZ(1)
0046         WRITE(3,4011) PSY(2),PSZ(2),PHY2,PHZ2
0047         WRITE(3,405)
0048         WRITE(3,402) SSY(1),SSZ(1)
0049         WRITE(3,4021)SSY(2),SSZ(2)
0050         WRITE(3,4022)SSY(3),SSZ(3),SHY3
0051         WRITE(3,4023)SSY(4),SSZ(4),SHZ4
0052         WRITE(3,4024)SSY(5)
0053         WRITE(3,4025)SSY(6),SHY6
0054         WRITE(3,405)
0055         WRITE(3,403) TSY(7),TSZ(6),THY7,THZ6
0056         WRITE(3,4031)TSY(10),TSZ(10),THY10,THZ10
      C
0057         RETURN
      C
0058    400  FORMAT(///6X,'ENTRANCE PUPIL COEFFICIENTS',13X,'PRINCIPAL RAY H-CO
            1EFFICIENTS'/6X,27('='),13X,28('=')/)
0059    401  FORMAT(' PSY 1 =',F10.5,', PSZ 3 =',F10.5)
0060    4011 FORMAT(' PSY 2 =',F10.5,', PSZ 4 =',F10.5,',   HYB   =',F10.5,', H
            1ZB    =',F10.5)
0061    402  FORMAT(' SSY 1 =',F10.5,', SSZ 3 =',F10.5)
0062    4021 FORMAT(' SSY 2 =',F10.5,', SSZ 4 =',F10.5)
0063    4022 FORMAT(' SSY 5 =',F10.5,', SSZ 6 =',F10.5,',    SHY 5 =',F10.5,',')
0064    4023 FORMAT(' SSY 8 =',F10.5,', SSZ 7 =',F10.5,',',22X,'SHZ 7 =',F10.5)
```

```
0065      4024 FORMAT(' SSY 9 =',F10.5,',')                                               TK1950
0066      4025 FORMAT(' SSY10 =',F10.5,',',22X,'SHY10 =',F10.5,',')                        TK1960
0067      403  FORMAT(' TSY11 =',F10.5,', TSZ13 =',F10.5,',    THY11 =',F10.5,', T      TK1970
              1HZ13 =',F10.5)                                                             TK1980
0068      4031 FORMAT(' TSY16 =',F10.5,', TSZ20 =',F10.5,',    THY16 =',F10.5,', T      TK1990
              1HZ20 =',F10.5)                                                             TK2000
0069      405  FORMAT(2X)                                                                 TK2010
0070           END                                                                        TK2020
```

316

APPENDIX G: TABLES OF SURFACE CONTRIBUTIONS TO G- AND h-COEFFICIENTS

The following three tables are the surface contributions to, respectively, the basal G_a-, G_b and h-coefficients. The final coefficient is given in the last column and the contribution by the *i*th surface is given in the *i*th column. (See §72,73.)

F.HAVLICEK EXTRA WIDE-ANGLE JOSA 41   FOCAL LENGTH =1.0000 ,STOP = 10.00
BASE RAY IS MERIDIONAL - SBY = 0.71902E-01, TBY = 0.10989E 01

PARABASAL COEFFICIENTS

	1	2	3	4	5
YA	1.00000	0.87042	1.57947	1.67042	1.68140
VA	0.0	-0.21354	1.08817	0.99345	-0.81070
YB	-0.51284	-0.33919	-0.20131	-0.19750	-0.19880
VB	1.00000	0.49253	0.13107	0.02197	0.31184
ZA	1.00000	0.94533	1.82616	1.94589	1.94589
WA	0.0	-0.19535	1.18052	1.15552	-0.93320
ZB	-0.51284	-0.33492	-0.05306	-0.03017	-0.03017
WB	1.00000	0.63573	0.37776	0.22094	0.38111

SECOND, THIRD ORDER G-COEFFICIENTS, A-COMPONENTS -

	1	2	3	4	
SGY 1	-0.02106	-1.10166	0.02145	0.99014	-0.11112
SGY 2	-0.07808	0.69655	-0.08519	-0.48288	0.05039
SGY 5	-0.08531	-0.11134	-0.01802	0.05247	-0.16221
SGY 8	-0.01125	-0.58181	0.01760	0.49429	-0.08118
SGY 9	-0.02331	0.18073	-0.07861	-0.07465	0.00416
SGY10	-0.04812	-0.03136	-0.02205	0.00468	-0.09685
SGZ 3	-0.01886	-1.11548	-0.01513	0.98946	-0.16002
SGZ 4	-0.03517	0.20915	-0.09490	-0.07587	0.00321
SGZ 6	-0.08070	0.56080	-0.00492	-0.40881	0.06637
SGZ 7	-0.15049	-0.10515	-0.03087	0.03135	-0.25516
TGY 1	-0.00688	3.15885	0.31802	-23.32140	-19.85141
TGY 2	-0.00833	-1.59583	0.04725	10.51085	8.95394
TGY 5	0.01779	0.29584	-0.03222	-1.61805	-1.33665
TGY 8	-0.00980	4.81096	0.31817	-35.08791	-29.96858
TGY 9	-0.00488	-0.85513	0.04556	4.79536	3.98091
TGY10	-0.01386	0.19151	0.02761	-0.31524	-0.10998
TGY11	0.02192	-0.02798	0.03216	0.09655	0.12265
TGY14	-0.01769	-1.24198	0.07784	5.61098	4.42915
TGY15	-0.00143	0.27664	-0.04282	-0.81624	-0.58384
TGY16	0.00919	-0.04705	0.02834	0.05868	0.04915
TGZ 3	-0.00705	4.22834	0.18890	-34.39285	-29.98268
TGZ 4	-0.01314	-0.69659	0.17576	2.51295	1.97897
TGZ 6	-0.00135	-1.54094	-0.03082	10.49355	8.92045
TGZ 7	-0.00252	0.29121	-0.06585	-0.80907	-0.58623
TGZ12	0.00642	0.14003	-0.01724	-0.78245	-0.65324
TGZ13	0.01197	-0.05172	0.04009	0.08305	0.08339
TGZ17	-0.00981	6.40526	0.01968	-51.72598	-45.31087
TGZ18	-0.03418	-3.05428	0.10707	11.63596	8.65456
TGZ19	-0.06581	0.66329	-0.10422	-1.02504	-0.53177
TGZ20	-0.06750	-0.05921	-0.01069	0.03350	-0.10390

E.HAVLICEK EXTRA WIDE-ANGLE JOSA 41 FOCAL LENGTH =1.0000 ,STOP = 10.00
BASE RAY IS MERIDIONAL - SBY = 0.71902E-01, TBY = 0.10989E 01

SECOND, THIRD ORDER G-COEFFICIENTS, B-COMPONENTS -

	1	2	3	4	
SGY 1	0.00648	0.10370	0.02262	0.00844	0.14123
SGY 2	0.07469	-0.06066	0.00599	0.02158	0.04161
SGY 5	0.02712	0.01070	0.00251	-0.00376	0.03656
SGY 8	-0.01750	0.15682	0.03149	-0.07958	0.09122
SGY 9	0.03583	-0.02090	0.00806	0.00948	0.03247
SGY10	0.01856	0.00345	0.00284	-0.00056	0.02429
SGZ 3	0.03183	-0.15079	0.00107	0.29188	0.17399
SGZ 4	0.05935	0.02827	0.00674	-0.02238	0.07198
SGZ 6	0.03003	0.01408	0.00006	-0.02999	0.01418
SGZ 7	0.05599	-0.00264	0.00039	0.00230	0.05604
TGY 1	-0.00716	-0.77748	-0.04047	3.66012	2.83501
TGY 2	0.00863	0.36380	-0.00170	-1.66604	-1.29530
TGY 5	-0.00117	-0.06055	-0.00198	0.23763	0.17392
TGY 8	-0.01109	-1.23482	-0.03810	5.54038	4.25638
TGY 9	0.02929	0.05980	-0.02113	-0.62920	-0.56124
TGY10	0.00198	-0.01971	-0.00547	0.02862	0.00543
TGY11	-0.01503	0.00346	-0.00352	-0.01199	-0.02707
TGY14	-0.01340	0.30739	0.00863	-0.88569	-0.58306
TGY15	0.01852	-0.04607	0.00021	0.11017	0.08283
TGY16	-0.01016	0.00460	-0.00324	-0.00574	-0.01455
TGZ 3	-0.01101	-0.83485	-0.00529	2.58186	1.73071
TGZ 4	-0.02053	0.16954	0.03845	-0.23468	-0.04722
TGZ 6	0.02012	0.21087	0.00268	-0.74466	-0.51099
TGZ 7	0.03751	-0.03390	-0.01137	0.04442	0.03665
TGZ12	-0.00941	-0.00030	0.00022	0.02732	0.01783
TGZ13	-0.01754	-0.00058	-0.00049	-0.00040	-0.01902
TGZ17	-0.01750	-1.37446	0.00847	3.95972	2.57624
TGZ18	-0.00137	0.30035	0.05301	-0.55622	-0.20422
TGZ19	0.07091	-0.00298	-0.00079	0.01399	0.08113
TGZ20	0.02357	-0.00094	0.00020	0.00041	0.02325

E.HAVLICEK EXTRA WIDE-ANGLE JOSA 41 FOCAL LENGTH =1.0000 ,STOP = 10.00
BASE RAY IS MERIDIONAL - SBY = 0.71902E-01, TBY = 0.10989E 01

SECOND AND THIRD ORDER IMAGE HEIGHT COEFFICIENTS -
==

	1	2	3	4	
SGY 1	-0.03024	-1.59876	0.02296	1.43661	-0.16944
SGY 2	-0.11299	1.01663	-0.11947	-0.71395	0.07023
SGY 5	-0.12255	-0.16123	-0.02819	0.07525	-0.23671
SGY 8	-0.01581	-0.83787	0.02223	0.70844	-0.12302
SGY 9	-0.03396	0.26492	-0.11354	-0.11344	0.00398
SGY10	-0.06918	-0.04495	-0.03297	0.00561	-0.14149
SGZ 3	-0.02112	-1.25255	-0.02597	1.11265	-0.18699
SGZ 4	-0.03939	0.23897	-0.10503	-0.09443	0.00012
SGZ 6	-0.08933	0.62802	-0.00171	-0.46358	0.07340
SGZ 7	-0.16657	-0.11559	-0.03682	0.03146	-0.28752
TGY 1	-0.00972	4.54318	0.49589	-33.89677	-28.86743
TGY 2	-0.01206	-2.29744	0.05298	15.27816	13.02164
TGY 5	0.02548	0.42698	-0.04435	-2.35076	-1.94266
TGY 8	-0.01384	6.91504	0.51355	-50.99287	-43.57813
TGY 9	-0.00746	-1.23997	0.05735	6.97857	5.78848
TGY10	-0.01988	0.27760	0.04173	-0.45875	-0.15929
TGY11	0.03162	-0.04144	0.04696	0.14181	0.17895
TGY14	-0.02510	-1.78832	0.10214	8.14894	6.43766
TGY15	-0.00235	0.40163	-0.06031	-1.18794	-0.84897
TGY16	0.01331	-0.06904	0.04106	0.08660	0.07194
TGZ 3	-0.00767	4.66904	0.24732	-38.46085	-33.55215
TGZ 4	-0.01429	-0.76902	0.19401	2.80065	2.21135
TGZ 6	-0.00169	-1.70614	-0.04726	11.73734	9.98225
TGZ 7	-0.00315	0.32568	-0.07240	-0.90628	-0.65616
TGZ12	0.00718	0.15573	-0.01817	-0.87479	-0.73005
TGZ13	0.01338	-0.05887	0.04476	0.09473	0.09400
TGZ17	-0.01065	7.06613	0.07612	-57.83490	-50.70331
TGZ18	-0.03769	-3.39222	0.09821	13.00240	9.67070
TGZ19	-0.07331	0.74450	-0.11250	-1.15620	-0.59750
TGZ20	-0.07470	-0.06539	-0.01306	0.03606	-0.11710

APPENDIX H: TABLES OF SYMBOLS AND AFFIXES

This Appendix is divided into three parts: (a) a table of special symbols and conventions, (b) a table of kernel symbols, and (c) a table of indices and other affixes. These tables are by no means exhaustive and contain only repeatedly used symbols etc. If a symbol has some particular meaning in only one section, that meaning is not given below. As in the text, Q and q are reserved for general kernel symbols and are only used in examples. Only the generic form of coefficients is given and §29 should be referred to for an interpretation of coefficients. Section references refer to the first use of the symbol in question.

(a) Special Symbols and Conventions

○	around a letter indicates script type
=	below a letter indicates German type
~	below a letter indicates block type, used only for two-vectors (§4)
→	above a symbol denotes a vector (§4)
*	as a superscript } are often used to indicate a variation of the usual meaning of the kernel symbol
-	above the kernel } are often used to indicate a variation of the usual meaning of the kernel symbol
(f|g)	$f_a g_b - f_b g_a$ (§14)
$\binom{n}{r}$	$n!/(n-r)!r!$ (§36)
O(m)	terms of the <u>m</u>th and higher orders in given variables

(b) Kernel Symbols

The symbols representing parabasal quantities are not shown separately but are indicated in brackets after the corresponding general symbol. Symbols whose kernel is a "w" or "z", are the second component of the corresponding two-vector whose kernel is "v" or "y".

Ⓐ	pseudo-axis (§1)
Ⓒ, Ⓒ$_d$, Ⓒ$_0$	coordinate systems associated with: Ⓕ (§5), Ⓓ (§63), Ⓐ (§1)
Ⓒ*	coordinate system associated with Ⓗ (§7)
Ⓒ*	coordinate system obtained by transformation of Ⓒ (§5)
Ⓓ	diaphragm (§1,61)
Ⓔ	entrance pupil (§6,61)
Ⓕ	boundary between two optical media (§1)
Ⓗ	object (or image) surface, usually plane (§6,7)
Ⓚ, Ⓚ$_S$	optical system (§1), symmetric system (§54)
Ⓛ	line in Ⓜ and, in object space of Ⓚ, normal to Ⓡ$_B$ (§18)
Ⓜ	plane of symmetry of Ⓚ, meridional plane (§1,3)
Ⓟ	point of incidence of Ⓡ with Ⓕ (§8)
Ⓡ, Ⓡ$_B$	arbitrary ray (§5), base-ray (§2)
Ⓢ	sagittal plane (§3)

a	matrix element in rotation of coordinates (§5)
a	α/α_B (§59)
a_y	coefficient of Y in ΔG_y (§26)
A	cross sectional area of pencil (§64)
A_E	area of Ⓔ (§64)

b	matrix element in rotation of coordinates (§5)
b_y	coefficient of V in ΔG_y (§26)

c	curvature of spherical surface (§1)
C	centre of spherical surface (§1)
d_d	location of plane of Ⓓ (§63)
d'	separation of x- and x_+-planes (§10)
$\underset{\approx}{E}$	optical invariant of $Ⓚ_S$ (§57) (denoted by $E_x{}^*$ in M)
E_B	basal point of Ⓔ (§6)
$\underset{\sim}{F}(\underset{\sim}{f})$	out-of-focus term (§22)
F_M, F_S	meridional and sagittal foci (§17)
$\underset{\sim}{F}_\nu{}^{(n)}$	coefficient of $\rho^{n-\nu}H^\nu$ in $\underset{\sim}{\epsilon}$ (§49)
$\underset{\sim}{g}$, g	$\underset{\sim}{\sigma}\bar{\tau} - \underset{\sim}{\bar{\sigma}}\tau$, usually denoted by g (§6)
$\underset{\sim}{G}(\underset{\sim}{g})$	quasi-invariant (§22) (denoted by $\underset{\sim}{\Lambda}(\underset{\sim}{\lambda})$ in M)
$\underset{\sim}{H}(\underset{\sim}{h})$	object height (§6,7)
h_1	object height of a parabasal, meridional object (§18)
H	distance of O from basal point of Ⓗ, i.e., length of $\underset{\sim}{H}$ (§48,49)
I_{Ax} , $\underset{\sim}{I}_A$	$n_z V_A - n_y W_A$, $n_x \underset{\sim}{V}_A - \underset{\sim}{n}$ (§9)
I_x , $\underset{\sim}{I}(i_x, \underset{\sim}{i})$	$I_{Ax} - I_{Bx}$, $\underset{\sim}{I}_A - \underset{\sim}{I}_B$ (§10)
I	angle of incidence (§2)
$\underline{k}$	number of surfaces of Ⓚ (§1)
k	refractance, $k = N/N'$ (§8)
k_y , k_z	$k_y = N_y/N_y{}'$, $k_z = N_z/N_z{}'$ (§13)
K_M , K_S	meridional, sagittal comatic asymmetries (§53)

$\underline{\underline{\ell}}$; $\underline{\underline{\ell}}_M$, $\underline{\underline{\ell}}_S$	distance of O_B from Ⓟ (§17) ; value of $\underline{\underline{\ell}}$ for F_M , F_S (§20)
ℓ	position of object plane or distance of O from x-plane (§6,7)
ℓ_M , ℓ_S	distances of F_M , F_S from x-plane (§20)
$\underline{m}$	magnification (§7)
$\underline{\underline{N}}_{gn}$	number of $\underline{n}$th order g-coefficients (§25)
$\underline{\underline{N}}_{In}$	number of $\underline{n}$th order extremal identities (§43)
$\underline{\underline{N}}_{yn}$, $\underline{\underline{N}}_{zn}$	number of $\underline{n}$th order terms in expansions of Q_y , Q_z (§24)
$\vec{n}$	normal to Ⓕ (§8)
N	refractive index (§1)
N_y, N_z	$N_y = N\alpha_B^3$, $N_z = N\alpha_B$ (§13)
N , N_{tr}	number of rays forming spot diagram, N_{tr} from ray traces (§77)
$\underline{n}$	order of term of coefficient (§24)
O, O_B	point object (§6) , basal object or basal point of Ⓗ (§6)
p	location of plane of Ⓔ (§6)
P(p)	RX (§10) (denoted by $\underline{\underline{X}}$ in M)
$P_{q\alpha}$, $P_{q\alpha}$	generic form of first order coefficients of Q (§29)
$q_{\mu\nu\tau}^{(n)}$, $\underline{\underline{q}}_{\mu\nu\tau}^{(n)}$ $\hat{q}_{\mu\nu\tau}^{(n)}$	generic form of $\underline{n}$th order coefficients: paracanonical (§24) , pseudo- (§35) , surface (§26)
$\hat{q}^{(n)}$, $q^{(n)}$	generic form of $\underline{n}$th order terms of expansion of Q in terms of: $\underline{Y}$ and $\underline{V}$ (§26) , $\underline{S}$ and $\underline{T}$ (§24)

$\bar{R}_A$	$(\Delta N \cos I)/N'\alpha'$ (§8) (denoted by $\bar{\underline{\underline{S}}}$ in M)
$R(r)$	defined by $\bar{R}_A = R\bar{R}_B$ (§8) (denoted by $\underline{\underline{S}}$ in M)
$\underset{\sim}{S}(\underset{\sim}{s})$	paracanonical coordinates (§6)
$S_{q\alpha}$, $s_{q\alpha}$	generic form of second order coefficients of Q (§29)
S_0	S_y for principal ray of pencil containing R_B (§65)
$\underset{\sim}{T}(\underset{\sim}{t})$	paracanonical coordinates (§6)
$T_{q\alpha}$, $t_{q\alpha}$	generic form of third order coefficients in Q (§29)
U	$N_z(ZV-YW)$ (§59)
$\underline{\underline{V}}$	point characteristic (§44)
$\underset{\sim}{V}(\underset{\sim}{v})$	$\underset{\sim}{V}_A - \underset{\sim}{V}_B$ (§5)
$\underset{\sim}{V}_A$	$\underset{\sim}{\beta}/\alpha$ (§5)
$\underset{\sim}{v}_p$, $\underset{\sim}{v}_q$, $\underset{\sim}{v}_a$, $\underset{\sim}{v}_b$	paracanonical coefficients (§14)
$\bar{\underset{\sim}{v}}$	$-\bar{R}_B(N_y'v_M' , N_z'w_S')$ (§26)
$\underset{\sim}{x}_M$	(x_M , x_S), displacement of F_M , F_S from Ⓗ (§22)
x	displacement of image plane from initial position (§51)
x , $\bar{x}$	abscisse of point on Ⓡ (§5,17)
X(x)	abscisse of Ⓟ (§9)
$\underset{\sim}{Y}$, $\underset{\sim}{Y}_d(\underset{\sim}{y},\underset{\sim}{y}_d)$	point of intersection of R with: x-plane (§5), Ⓓ (§63)
$\underset{\sim}{y}_p$, $\underset{\sim}{y}_q$, $\underset{\sim}{y}_a$, $\underset{\sim}{y}_b$	parabasal coefficients (§14)
$\underset{\sim}{y}$, $\bar{\underset{\sim}{y}}$	ordinates of point on R (§5,17)
$\bar{\underset{\sim}{y}}$	$-\bar{R}_B(N_y'y_M' , N_z'z_S')$ (§26)

$\vec{\beta}$	(α,β,γ) , direction of ray (§5)
$\underset{\sim}{\delta}_y$, $\underset{\sim}{\delta}_v$	$-\sum_i \Delta`\underset{\sim}{G}_{bi}$ , $\sum_i \Delta`\underset{\sim}{G}_{ai}$ (§25)
δ	grid mesh used in construction of spot diagrams (§74)
Δ, Δ_j, ∇	defined by: $\Delta Q = Q' - Q$ (§8) , $\Delta_j Q = Q_j - Q_1$ (§44) , $\nabla Q = Q_+ - Q'$ (§10)
$\underset{\sim}{\epsilon}$	aberration (§7)
ζ	$\underset{\sim}{T}_A \cdot \underset{\sim}{T}_A$ (§60)
η	$\underset{\sim}{S}_A \cdot \underset{\sim}{T}_A$ (§60)
θ	angle variable for polar coordinates in Ⓔ (§21) or Ⓓ (§64)
$\lambda_1,\lambda_2,\lambda_3,\lambda_4$	combinations of pseudo-parameters (§13,32)
$\underset{\sim}{\mu}$	augmenting factor (§22)
μ	constant term in equation for $\bar{R}_A$ or R (§9,27)
$\bar{\mu}$	$\alpha_B{}^2 \underset{\sim}{V}_A \cdot \underset{\sim}{V}_A$ (§32)
ξ	$\underset{\sim}{S}_A \cdot \underset{\sim}{S}_A$ (§60)
Π	product sign
ρ_0, ρ_1, ρ_2	coefficients in quadratic equation for X (§27)
ρ	radial coordinates in Ⓔ (§21) or Ⓓ (§64)
ρ_{tr}	radial coordinate in Ⓓ determined by tracing (§75)
ρ_d	radius of Ⓓ (§63)
σ_α	second order aberration coefficients (§52)
σ	coefficient of linear term in equation for $\bar{R}_A$ or R (§9,27)
$\underset{\sim}{\sigma}$, $\bar{\underset{\sim}{\sigma}}$	coefficients of $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$ in $\underset{\sim}{S}$, usually denoted by σ, $\bar{\sigma}$. (§6)
Σ	summation sign
τ_α	third order aberration coefficients (§53)
τ	coefficient of quadratic term in equation of $\bar{R}_A$ or R (§9,27)
$\underset{\sim}{\tau}$, $\bar{\underset{\sim}{\tau}}$	coefficients of $\underset{\sim}{Y}_1$ and $\underset{\sim}{V}_1$ in $\underset{\sim}{T}$, usually denoted by τ, $\bar{\tau}$ (§6)
ϕ	angle coordinate for polar coordinates in Ⓗ (§49)
$\hat{\phi}^{(n)}$	$Q_0{}^2\hat{\tau}^{(n)} + 2Q_0\hat{\sigma}^{(n)} + \hat{\mu}^{(n)}$ (§30)
χ	$[1+(\alpha_B b/\alpha_B{}^*)V]^{-1}$ (§5)

$\chi(y,z)$	equation of aspheric surface (§87)
$\chi_\mu(n)$	coefficients in expansion of $\chi(y,z)$ in terms of y,z (§87)
$\hat{\psi}^{(n)}$	$\tau_0\hat{q}^{(n)} + 2\hat{\sigma}^{(n)} + 2Q_0\hat{\tau}^{(n)}$ (§30)
ψ	angle of inclination of ray with axis (§69)

(c) Indices

The various indices characterising coefficients were described in §29, in particular in Tables 29/1,2. The only entries from these tables which are repeated below are those which are also relevant to quantities other than coefficients. In the table below Sb is an abbreviation for "subscript" and Sp an abbreviation for "superscript". The number in parentheses after these abbreviations gives the rank (§29) of the index.

`	Sp(-1)	ante-prime: $\grave{Q} = (g/N_{y1})Q$ (§25)
^	Sp(0)	indicates that a quantity is either expressed in terms of $\underset{\sim}{Y}$, $\underset{\sim}{V}$ or associated with such an expression (§26)
~	Sp(0)	implies multiplication by $\underset{\sim}{\mu}$ (§22)
*	Sp(1)	indicates that a quantity is either of the second order in $\underset{\sim}{Y}$, $\underset{\sim}{V}$ or associated with such a quantity (§26)
†	Sp(1)	same as * but with $\underset{\sim}{Y}$, $\underset{\sim}{V}$ replaced by $\underset{\sim}{S}$, $\underset{\sim}{T}$ (§24)
+	Sp(5)	as a surface indicator, an abbreviation for j+1 (§10)
'	Sp(7)	prime: implies association with the image space (§1)
"	Sp(7)	attached to a quantity referring to the object space of Ⓕ, implies definition in $Ⓒ_H'$ (§84)
#	Sp [1]	denotes dual transformation (§94)
,	Sp [1]	implies differentiation by variables following "," (§44)

---

[1] No rank has been assigned. The affix goes immediately to the right of the last affix characterising the quantity being transformed. Further affixes refer to the transformed quantity.

A	Sb(1)	implies reference to the axial theory (§5,54)
a , b	Sb(4)	denotes a-, b-components (§22)
B	Sb(1)	implies association with $Ⓡ_B$, Q_B is value of Q_A for $Ⓡ_B$ (§2,54)
d	Sb(5)	as a surface indicator refers to Ⓓ (§63)
H	Sb(1)	implies association with hamiltonian coordinates (§18)
i, j, k	Sb(5)	surface indicators: denote particular surface of Ⓚ (§1)
M	Sb(1)	implies association with meridional rays (§17)
N	Sb(1)	implies association with normal coordinates (§14)
(n)	Sp(6)	denotes order of term or coefficient (§24)
0	Sb(1)	denotes a special value, usually zeroth order term (§24)
P	Sb(1)	implies association with principal ray (§7)
S	Sb(1)	implies association with sagittal rays (§17)
y , z	Sb(2)	indicates y- , z-components of two-vectors (§4)

BIBLIOGRAPHY

ALLEN, W.A. and SNYDER, J.R., J.O.S.A., 42, 243-8 (1952).

BARAKAT, R. and HOUSTON, A., Optica Acta, 13, 1-30 (1966).

BORN, M. and WOLF, E., "Principles of Optics" (Pergamon Press, 1964).

BUCHDAHL, H.A., "Optical Aberration Coefficients" (O.U.P., 1954).

J.O.S.A., 48, 747-756 (1958).

J.O.S.A., 49, 1113-1121 (1959).

J.O.S.A., 50, 678-683 (1960).

J.O.S.A., 55, 641-649 (1965).

CONRADY, A.E., Roy. Astron. Soc. MN., 79, 384-390 (1919).

"Applied Optics and Optical Design, I" (Dover, 1957).

COURANT, R., "Differential and Integral Calculus, Vol.1" (Blackie and Son, 1959).

EPSTEIN, L.I., J.O.S.A., 39, 847-53 (1949).

FLÜGGE, J., "Das Photographische Objektiv" (Springer-Verlag, Berlin, Vienna, 1955).

FORD, P.W., J.O.S.A., 50, 528-533 (1960).

"The Use of Optical Aberration Coefficients" (Thesis, Tasmania, 1962).

J.O.S.A., 56, 209-212 (1966).

GARDNER, I.C., J. Res. Nat. Bur. Stand., Wash., 39, 213-19 (1947).

GARDNER, I.C. and WASHER, F.E., J. Res. Nat. Bur. Stand., Wash., 40, 93-103 (1948).

HAVLICÉK, F.J., J.O.S.A., 41, 1058-1059 (1951).

HERZBERGER, M., J.O.S.A., 47, 583-594 (1957).

"Modern Geometrical Optics" (Int. Sci. Publishers, 1958).

HILL, R., Roy. Met. Soc., Quart. Jour., 50, 227-235 (1924).

KIUTI, M., Sci. Pap. Coll. Gen. Educ., Univ. Tokyo., 1, 19-36 (1951).

KUBOTA, H. and MIYAMOTO, K., Rep. of the Inst. Indust. Sci., Univ. Tokyo, 13, 38-53 (1963).

LUNEBERG, R.K., "Mathematical Theory of Optics" (Univ. Calf. Press, 1964).

MARÉCHAL, A., Comptes Rendus, 228, 668-70 (1949).

MAXWELL, J.C., "The Scientific Papers of James Clerk Maxwell, Vol.2", p.144 (C.U.P., 1890).

MONTEL, M., Revue d'Optique, 32, 585-600 (1953).

NIJBOER, B.R.A., Physica, 10, 679-692 (1943).

REISS, M., J.O.S.A., 38, 980-986 (1948).

SLUSSAREFF, G., Jour. of Physics, U.S.S.R., 4, 537-545 (1941).

SMITH, T., Trans. Opt. Soc., 29, 71-87 (1928a).

Trans. Opt. Soc., 29, 167-178 (1928b).

Trans. Opt. Soc., 31, 131-156 (1930).

STAVROUDIS, O.N. and SUTTON, L.E., "Spot Diagrams for the Prediction of Lens Performance from Design Data", N.B.S. Monograph 93 (U.S. Govt. Printing Office, 1965).

STEPHAN, W.G., App. Sci. Res. B., 1, 273-283 (1949).

STEWARD, C.G., "The Symmetrical Optical System" (Camb. Tracts in Math. and Math. Physics, No.25, C.U.P., 1958).

SYNGE, J.D., "Geometrical Optics and an Introduction to Hamilton's Method" (Camb. Tracts in Math. and Math. Physics, No.37, C.U.P., 1937).

WEINSTEIN, W., Proc. Phys. Soc., 62B, 726-40 (1949).

Proc. Phys. Soc., 63B, 709-23 (1950).

WYNNE, G.G., Proc. Phys. Soc., 67B, 529-37 (1954).

INDEX