UNIVERSITY OF LONDON

Imperial College of Science and Technology

Department of Physics

Optics Section

FOCUSSING OPTICS FOR LASER COMPRESSION EXPERIMENTS

Nicholas Edwin Ellis

Thesis submitted for the degree of Doctor of Philosophy

of the University of London (1982)

ABSTRACT

The purpose of the focussing optics for laser compression and fusion experiments is to uniformly illuminate the spherical target at normal incidence. Consequently, such optics are classified as energy delivery systems.

After a brief description of the compression process, the typical laser used and some initial considerations for the optics, Chapter 2 shows how the target illumination for any system can be calculated. It is seen that the energy transformation depends on the aberrations and the principal surface shape. The inverse problem of finding the latter from the desired energy transformation is of importance and particular conditions for uniformity are deduced. Other examples are also considered; the aplanat being an important case.

Knowing the principal surface required for a particular energy transformation, a method for designing a lens system with this surface is given. Two aspheric surfaces are required to do this and satisfy the constant optical path condition.

Chapters 4 and 5 are concerned with designs that use a single aspheric surface. It may be used to satisfy one of two conditions; either the constant optical path condition or to achieve the desired target illumination. A method for generating the aspheric for either case is developed in Chapter 4 and examples given.

Wavelength tunable lenses are discussed in Chapter 5. Here, the aspheric null corrects the lens at a visible wavelength, thus aiding the testing procedure during manufacture. It can

then be recorrected for the laser wavelength (usually infrared). They can actually be tuned to any wavelength over a wide range, so may be used at the laser harmonics.

This multiple wavelength use is then taken further by using two aspherics to satisfy the constant optical path condition or produce the correct illumination at two wavelengths (fundamental and second harmonic).

Finally, nonlinear effects and other problems arising from the high powers are discussed.

CONTENTS

		Page
Notation		8
Chapter 1.	INTRODUCTION	11
1.1)	Laser compression and fusion	11
1.2)	The laser	13
1.3)	The focussing optics : some initial considerations	18
1.4)	Mirrors versus lenses	24
Chapter 2.	TARGET ILLUMINATION	30
2.1)	Introduction	30
2.2)	Calculating the target intensity	31
2.3)	Conditions for uniform target illumination	35
2.4)	Uniform illumination from a Gaussian beam	39
2.5)	The aplanat	40
2.6)	Uniform and orthogonal illumination	45
2.7)	Lens-mirror combinations	50
2.8)	Spherical aberration required to uniformly illuminate the target	51
Chapter 3.	DESIGN OF LENSES WITH A SPECIFIED PRINCIPAL SURFACE	54
3.1)	Introduction	54
3.2)	The singlet	55
3.3)	Design examples	63
3.4)	The doublet	69
3.5)	Design examples	76
3.6)	Conclusions	84

Page

Chapter 4. DESIGNS CONTAINING A SINGLE ASPHERIC (I) 86

4.1) Introduction 86

4.2) Correcting spherical aberration by the use of an aspheric surface 88

4.3) Calculation of the spherical aberration coefficients 91

4.4) Initial conditions and procedure 104

4.5) Calculating the shape of the aspheric surface 107

4.6) Target intensity from the polynomial expansions 110

4.7) Examples corrected for constant optical path 114

4.8) Uniform illumination using one aspheric surface or corrector plate 120

Chapter 5. DESIGNS CONTAINING A SINGLE ASPHERIC (II): WAVELENGTH TUNABLE LENSES 127

5.1) Introduction 127

5.2) Recorrecting spherical aberration 129

5.3) Examples and discussion 131

5.4) Conclusions 147

Chapter 6. DUAL-WAVELENGTH LENSES 149

6.1) Introduction 149

6.2) Correcting spherical aberration at two wavelengths by the use of two aspheric surfaces 151

6.3) Design examples 152

Chapter 7. NONLINEAR EFFECTS AND DAMAGE 165

7.1) Nonlinear effects 165

7.2) Ghost images 168

7.3) Nonlinear phase: the B-integral 183

		Page
Chapter 8.	SUMMARY AND CONCLUSIONS	188
Appendix A.	THE ALDIS THEOREM	190
Appendix B.	RUNGE-KUTTA SOLUTION OF FIRST-ORDER DIFFERENTIAL EQUATIONS	196
Appendix C.	FORTRAN LISTING OF ROUTINES THAT CALCULATE THE ABERRATION COEFFICIENTS, $T^{(i)}$	199
Appendix D.	COMA, THE SINE RELATION, THE SINE CONDITION AND OFFENSE AGAINST THE SINE CONDITION	206
References		215
Acknowledgments		217

Notation

Roman (Upper case)

$A_2, A_4, \ldots$	Polynomial coefficients of $Y^2, Y^4, \ldots$ for a surface expansion in the form, $Z = A_2Y^2 + A_4Y^4 + \ldots$
B	Refraction invariant
$B_\int$	B-integral
C	Linear coma coefficients
$D_0, D_1, \ldots, D_5$	Dispersion formula coefficients
E	Electric field
F	Focal length
H	Lagrange invariant
I	Intensity
$K_{15}, K_{17}, \ldots$ $K_{35}, K_{37}, \ldots$	Constants relating the aberration coefficients for optimum correction
LA	Longitudinal spherical aberration
L	X direction cosine of a ray
M	Y direction cosine of a ray
N	Z direction cosine of a ray
OSC	Offense against the sine condition
P	Polarization
R_T	Target radius
R'	Radius of reference sphere
S_0	Target centre-paraxial focus distance
S	$=S_0 + LA$
T	Transverse ray aberration in the paraxial image plane
T_ζ	Change of transverse ray aberration due to a defocus

U	Finite ray angle
W	Wavefront aberration
X	X surface coordinate
X_p	X pupil coordinate
Y	Y surface coordinate
Y_p	Y pupil coordinate
Z	Z surface coordinate

Roman (Lower case)

$a_4, a_6, \ldots$	Polynomial coefficients of $Y^4, Y^6, \ldots$ for a surface expansion in the form, $$Z = \frac{cY^2}{1+\sqrt{1-c^2Y^2}} + a_4Y^4 + a_6Y^6 + \ldots$$
b	Back focal length
c	Paraxial curvature
d	Element spacing and thickness
f	Function of Y expressing the aspheric nature of a surface
g	Gaussian beam radius at e^{-1} central intensity
n	Refractive index
n_I	Nonlinear index
n_{NL}	Nonlinear refractive index ($=n+n_II$)
t	Time
u	Paraxial ray angle
y	Paraxial ray height

Greek (Lower case)

α	X direction cosine of surface normal
β	Y direction cosine of surface normal

γ	Z direction cosine of surface normal
χ_1 χ_2 χ_3	Susceptibility terms
ε_0	Permittivity of free space
η	Field coordinate
λ	Wavelength
ϕ	Target angle
ρ	Normalized ray height ($=Y/Y_m$)
θ	Azimuth angle
ζ	Defocus
σ	Binomial coefficients
ω	Angular frequency

N.B.

1.) $'$ refers to image space. When also referring to an intermediate image space a double prime $''$ will be used.

2.) $\bar{}$ refers to the principal ray.

3.) Subscripts number a set of variables and can denote an optical surface or medium. The following subscripts are also used.

B - beam; T - target; m - marginal ray; p - pupil; ζ - defocus

4.) A superscript in brackets, (i), denotes the order of a particular term.

5.) Distances measured from left to right and above the optical axis are defined to be positive while those from right to left and below the axis are negative.

6.) Ray angles are defined to be positive or negative when the Y-direction cosine is also positive or negative respectively.

CHAPTER 1

INTRODUCTION

1.1. Laser compression and fusion

The process of laser compression involves the illumination of a small spherical pellet by a very high powered and very short pulse of laser radiation. If the pellet is illuminated from all directions, uniformly and simultaneously this high energy flux will compress the target to a fraction of its original size and heat it to very high temperatures. The eventual purpose behind such experiments is the goal of achieving laser fusion. Targets made from the hydrogen isotopes deuterium and tritium, would be compressed and heated to the conditions under which the fusion of hydrogen nuclei to helium can occur. The energy released by the reaction being significantly greater than the initial input for a worthwhile enterprise.

Lasers powerful enough to do this do not exist at present. Today, the typical powers achievable are in the range of hundreds of GW to tens of TW with pulse durations of 1ns to 100ps. So in this short period the target can receive a dose of energy amounting to a few KJ. Energy fluxes over the target surface may be in the range of 10^{15} to 10^{16} W/cm^2.

A typical target size is 0·1mm diameter. They are usually illuminated by many beams, each focussed onto the target from a different direction in three-dimensional space, usually in a symmetrical fashion.

During the initial stages of the laser pulse, a plasma atmosphere develops around the target, which is usually

referred to as the "critical surface." It is here that maximum absorption occurs for the part of the pulse still to arrive at the target. This critical surface also has the effect of smoothing out any non-uniformities in the target illumination. Consequently, the uniformity tolerance provided by the optics is not nearly so stringent as it would otherwise have to be. Estimates for the uniformity range from 10-20%.(1)

The energy delivered to the target flows inwards by an electron transport mechanism to heat the pellet's surface. Material effectively "boils" away from the surface, causing the remaining mass to implode by reaction.

The very high density (about 10^{26} particles per cm^3) and temperature (10^8K) that could be achieved by this process mean that the plasma would only need to be held together for a very short time to allow the nuclear reaction to take place (10^{-12}sec). The inertia of the plasma itself is capable of confining it for this length of time. Consequently, laser fusion has been referred to as "inertial-confinement" fusion. This is contrasted by the other approach to fusion power. Here a low density plasma is confined in a magnetic "bottle" for periods of the order of 1sec. The inertia is not capable of holding the plasma together for this length of time, hence the magnetic confinement.

It is found both theoretically and in practice that the distance travelled by the radiation in the plasma before absorption, is proportional to the square of the wavelength. Thus shorter wavelengths are absorbed more efficiently. The threshold energy densities for various plasma instabilities

are also found to increase for shorter wavelengths, so these advantages have led to an investigation of the compression process at the laser harmonic frequencies. Nowadays, facilities for generating outputs at the fundamental, second and even the third harmonics are being constructed.

1.2. The Laser

In virtually all laboratories the Neodymium-glass laser is the type used for compression experiments although research is constantly exploring alternatives (e.g. the CO_2 laser at 10·6μm; Krypton Fluoride (KrF) at 0·249μm and Xenon Chloride (XeCl) at 0·308μm).

A glass rod doped with Neodymium ions (Nd^{3+}) pumped by Xenon flash lamps will lase in the infrared with a wavelength of 1·064μm. The output pulse is selected from a mode-locked oscillator and then shaped, spatially and temporally, before being split into a number of separate beams. Each of the beams is then amplified. This amplification does not occur in a single stage; the maximum gain is fixed by the damage threshold to the amplifying medium and other optical components.

At the very high intensities dealt with, the response of a medium to the electric field is no longer linear. At low and moderate energy fluxes, the polarization, P, induced in a medium by an electric field, E, is

$$P = \varepsilon_0 \chi_1 E \qquad (1.1)$$

where χ_1 is the susceptibility and ε_0 is the permittivity of free space. With a high flux, however, this expression must

be replaced by,

$$P = \varepsilon_0\chi_1 E + \varepsilon_0\chi_2 E^2 + \varepsilon_0\chi_3 E^3 + \ldots\ldots \qquad (1.2)$$

The terms in E^2, E^3,... lead to a host of new phenomena known as "nonlinear" effects which can damage the optical elements and break-up the beam. One important result of these higher order terms is the variation of refractive indices with intensity,

$$n_{NL} = n + n_I I \qquad (1.3)$$

where n is the normal refractive index, I the beam intensity and n_I the nonlinear index. The fact that the beam has to propagate through large amounts of glass increases the chances of damage from nonlinear effects.

The maximum gain is therefore set by the desire to keep the intensity below the damage threshold. This is typically between 5 and 20GW/cm^2. The lower limit is for surface coatings (reflecting or antireflecting) and the upper limit for bulk materials. If the beam is to stay below this threshold it must be expanded before the next amplification. The first power gains are achieved by glass rod amplifiers while the larger apertures required (10 to 20cm diameter) for the final stages are usually a set of Neodymium-glass discs. The latter are tilted at the Brewster angle so that a suitably polarized beam does not lose energy by reflection at the surfaces.

Once each beam has been amplified to the required power level they are then focussed onto the target so as to achieve an acceptable illumination. This thesis concerns itself with the design of such focussing optics. The illumination should

be uniform and reasonably orthogonal over the target surface. Although to meet the orthogonality condition completely it is necessary to illuminate the target from all the 4π solid angle, this cannot be done in practice; space around the target has to be left for target insertion into the chamber and the various detectors for observation of the compression process. Over the years as the laser power has increased from megawatts to terawatts the beam size and number of beams used has had to increase to cope with the greater energy fluxes. Target chambers have also had to be scaled up so that the optics and detectors can fit into the 4π solid angle.

Fig. (1.1) shows a schematic diagram of the type of laser system described. Not shown, are all the other optical and electro-optical devices for handling the beam, e.g. spatial filters and Faraday rotators. The latter can rotate the plane of polarization of any reflected pulse so that it can then be rejected by a polarizer. This prevents the pulse from being amplified and condensed into smaller areas as it propagates back down the system, inevitably leading to damage.

Although the laser in widespread use for fusion is the Neodymium-glass laser, it is not believed to be suitable for a practical fusion reactor. The reasons for this are its low efficiency (0·2%) and low cooling rate of the glass elements. The latter implies that to prevent over heating, the pulse rate is low. Properties required would be the opposite to this; high efficiency and pulse rate. Until research produces an alternative, however, the Nd:glass laser will remain the favoured type.

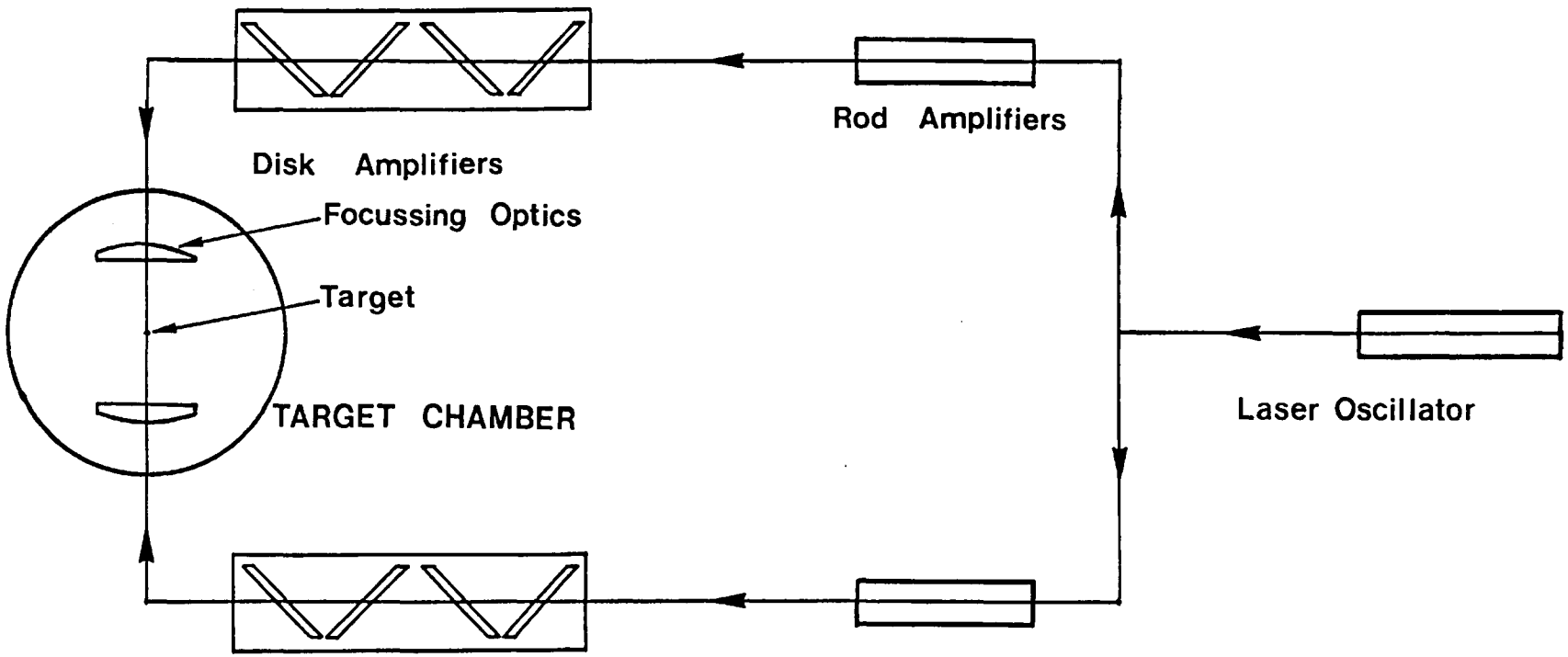


Fig. (1.1) Schematic view of a Nd:glass laser compression system with 2 beams.

By the deliberate introduction of a nonlinear effect known as "harmonic generation" the same laser system can be used to generate outputs at 0·532 and 0·355μm. for second and third harmonics respectively. Thus the advantages of using shorter wavelengths can be studied without the need for a different laser.

From equation (1.2) we can see how the harmonics originate. If the incident electric field is of the form,

$$E = E_0 e^{i\omega t} \tag{1.4}$$

then (1.2) gives,

$$P = \varepsilon_0 \chi_1 E_0 e^{i\omega t} + \varepsilon_0 \chi_2 E_0^2 e^{i(2\omega)t} + \varepsilon_0 \chi_3 E_0^3 e^{i(3\omega)t} + \ldots \tag{1.5}$$

The polarization of the medium is therefore forced to have harmonic variations of frequency 2ω, 3ω,.. etc. So re-radiated fields will contain a proportion of these frequencies. Conversion efficiencies greater than 50% have actually been obtained for second harmonic generation.

To generate these from the fundamental a Potassium Dihydrogen Phosphate (KDP) crystal is employed. It can be seen that for an isotropic medium, P must change sign if E changes sign. Hence terms in (1.2) that contain even powers of E (E^2, E^4,...) must be absent.

Thus second harmonic generation is not possible from an isotropic material. Anisotropic media are required. In such a case (1.2) has to be rewritten as,

$$P_i = \varepsilon_0 \chi_{iJ} E_J + \varepsilon_0 \chi_{iJk} E_J E_k + \varepsilon_0 \chi_{iJkl} E_J E_k E_l + \ldots \tag{1.6}$$

where the susceptibility terms are now tensors rather

than scalars. The summation convention over repeated indices applies here. Second harmonic generation is now possible, provided $\underline{P}$ does not change sign when $\underline{E}$ changes sign i.e. provided the crystal does not have inversion symmetry.

1.3. The focussing optics: some initial considerations

The purpose of the focussing optics is to illuminate the target uniformly and as close to normal incidence as possible. They have to transform a certain beam intensity to a uniform profile over the target. By doing this, it is believed that the chances of producing a uniform compression are best. Consequently, such optical systems are classed as energy delivery rather than imaging and will therefore require different considerations to those of normal optical design. Apart from the number of beams, the orthogonality may be improved by the use of high aperture optics, typically about F/1. Lenses with apertures as high as F/0·6 have been made, but are close to the limit possible; much beyond this and marginal rays are internally reflected at a surface. With lens-mirror combinations the aperture can be increased to F/0·5 or more.

In order to minimize the nonlinear effects, the amount of glass traversed by the beam should be kept to a minimum. The dependence of refractive index on the intensity,

$$n_{NL} = n + n_I I \tag{1.7}$$

causes an amount of nonlinear phase to build up as the beam propagates through each optical element. As a result, the beam

tends to break up losing its spatial coherence. This nonlinear phase is expressed by the "B-integral". It is essentially the integral of the index change along the path of a ray.

$$B_{\int} = \frac{2\pi}{\lambda} \int \Delta n.ds \qquad (1.8)$$

$$B_{\int} = \frac{2\pi}{\lambda} n_I \int Ids \qquad (1.9)$$

or equivalently the integral of the intensity along the ray path. The contribution to the B-integral from the focussing optics (Chapter 7) should therefore be kept to a minimum. The obvious way to do this is to keep the number of refracting elements as low as possible. In practice, never more than two are used.

With only one or two elements and apertures of F/1, the aberrations can only be corrected by the use of aspheric surfaces. Fig. (1.2) illustrates why, by showing how large the spherical aberration can become by using only spherical surfaces. The figure shows a singlet of about F/1 which has been bent so that the aberration is a <u>minimum</u>. This can be reduced an order of magnitude by using a doublet, but certainly not to an acceptable level. With target sizes of 0·1mm diameter the degree of correction required should produce an image size significantly less than this. Aspheric surfaces must be used.

Generally, because the optics, target and beam should all be aligned to a common axis, spherical aberration is of far more concern than the off-axis aberrations. Of the latter, coma is the most important. If not corrected, then any slight misalignment with the beam may cause a serious image deterioration. Thus coma must not be allowed to get out of hand.

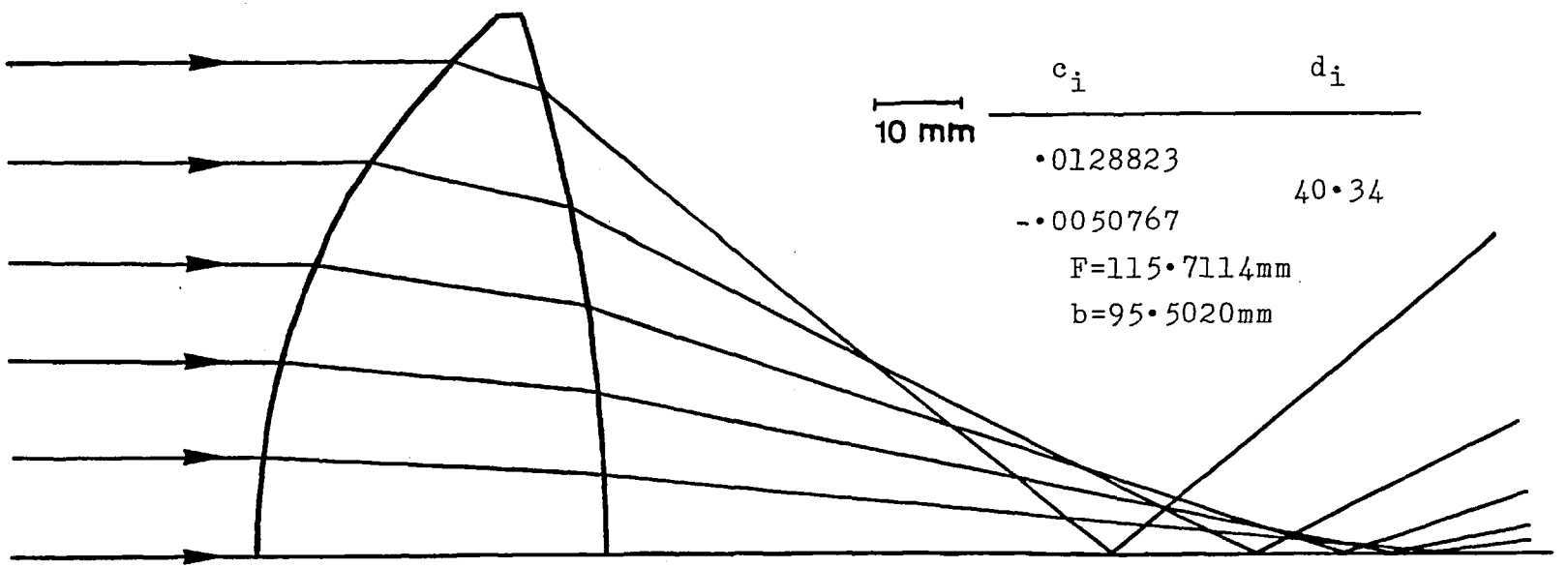


Fig. (1.2) Singlet with only spherical surfaces and having minimum aberration.

The other aberrations; namely astigmatism, field curvature and distortion are of little consequence. Their correction will go unnoticed unless the coma is very well corrected. Chromatic aberrations are of no concern either since we are dealing with monochromatic radiation. There is, therefore, no need to use more than one type of glass. BK7(516642) is the usual choice for the kind of system considered here. Its optical properties are summarized in Table (1.1).

According to a paper by Weber, Smith and Milan (2), low index and low dispersion glasses tend to also have low nonlinear indices, n_I. With BK7 fitting this category better than most, this is one good reason for its choice. In all the designs to be given, BK7 is used. The above paper also give nonlinear indices for Beryllium Fluoride glasses as being three to four times smaller but have the disadvantages of toxicity and being hygroscopic as well as more expensive.

Whatever glass is chosen, it must be of the highest quality; any defects, internal or on a surface are likely to result in damage from the high powered beam.

The considerations for the design of the focussing optics in this thesis are on the following lines. In the next chapter a method by which the target illumination may be calculated is given, and various conditions for uniformity are found. It is realized that the transformation of energy from the beam to the target is determined by the aberrations present and the shape of the principal surface.

Chapter 3 then provides a technique for designing lenses that are corrected for the constant optical path condition and have a specified principal surface. In other words, lenses

Refractive index and dispersion:

$$n_d = 1{\cdot}51633,\quad n_{1{\cdot}064} = 1{\cdot}50621,\quad n_{0{\cdot}532} = 1{\cdot}51900$$

$$V_d = \frac{n_d - 1}{n_f - n_c} = 64{\cdot}15$$

Dispersion formula constants:

$$n^2 = D_0 + D_1\lambda^2 + \frac{D_2}{\lambda^2} + \frac{D_3}{\lambda^4} + \frac{D_4}{\lambda^6} + \frac{D_5}{\lambda^8}$$

where λ is in microns.

$D_0 = 2{\cdot}2697665$

$D_1 = -9{\cdot}6395197 \times 10^{-3}$

$D_2 = 1{\cdot}1025458 \times 10^{-2}$

$D_3 = 7{\cdot}9465126 \times 10^{-5}$

$D_4 = 1{\cdot}0120957 \times 10^{-5}$

$D_5 = -4{\cdot}4096694 \times 10^{-7}$

Nonlinear index:

$$n_{NL} = n + n_I I$$

$$n_I = 3{\cdot}4 \times 10^{-16}\,cm^2/W$$

Table (1.1) Properties of BK7 (516642) glass.

that provide the desired energy transformation. This requires the use of two aspheric surfaces.

In order to reduce the cost of manufacture for the focussing optics, designs containing a single aspheric surface should no doubt be considered. This is done in the next two chapters. In Chapter 4 a series expansion method for designing the aspheric surface is described, either for correcting spherical aberration or for producing a selectively aberrated image (both of which are required). Although this could be done by more orthodox techniques the method used here is in anticipation of Chapter 6 and also has the advantage of applying equally well to any surface, whether it is the first, last or interior of a system and refracting or reflecting. It may therefore find use elsewhere because of its general applicability. Various examples are also given in Chapter 4.

Chapter 5 deals with wavelength tunable lenses. These are doublets whose aspheric can be null figured at a convenient visible wavelength and then recorrected (tuned) for use in the infrared, thus facilitating their manufacture.

Chapter 6 reflects the increasing interest of using harmonic generation to produce an output of shorter wavelength. These are dual-wavelength lenses, applying here to 1.064μm and 0•532μm. They contain two aspheric surfaces which correct the system at two wavelengths. The aspheric routine in Chapter 4 was also developed to cope with this problem. All the usual methods cannot easily be used.

Finally, in Chapter 7 a discussion of the problems that

may arise from the high intensities is given, and how they might be avoided. The location of ghost images for example.

Geometrical optics has been used throughout. Provided the target radius is much smaller than the input beam radius and substantially larger than the wavelength of radiation, diffraction effects can be ignored. Phipps, Bodner and Shearer (1) quote an error of less than 1% for the geometrical approach when the input beam is diffraction limited and the target radius within the range,

$$0{\cdot}07Y_m > R_T > 2\lambda \qquad (1.10)$$

With the beam radius, $Y_m = 50mm$ and $\lambda = 1\mu m$ the target radius, R_T is to lie in the interval,

$$3{\cdot}5mm > R_T > 0{\cdot}002mm \qquad (1.11)$$

Typical values for R_T are $50\mu m$ ($0{\cdot}050mm$) so geometrical optics should provide good accuracy.

1.4. Mirrors versus Lenses

In general the focussing optics may use both lenses and mirrors though there are certain advantages and disadvantages in using each. Proposed systems that employ mirrors and found in the literature (3), (4), (5), (1) are illustrated in Fig. (1.3). Of these only one makes no use of any lenses, ie is a purely reflecting system.

Table (1.2) summarizes the main advantages and disadvantages of using mirrors. Two obvious benefits are their wavelength invariant properties and the absence of any nonlinear refractive effects. The former aids the testing of the figure

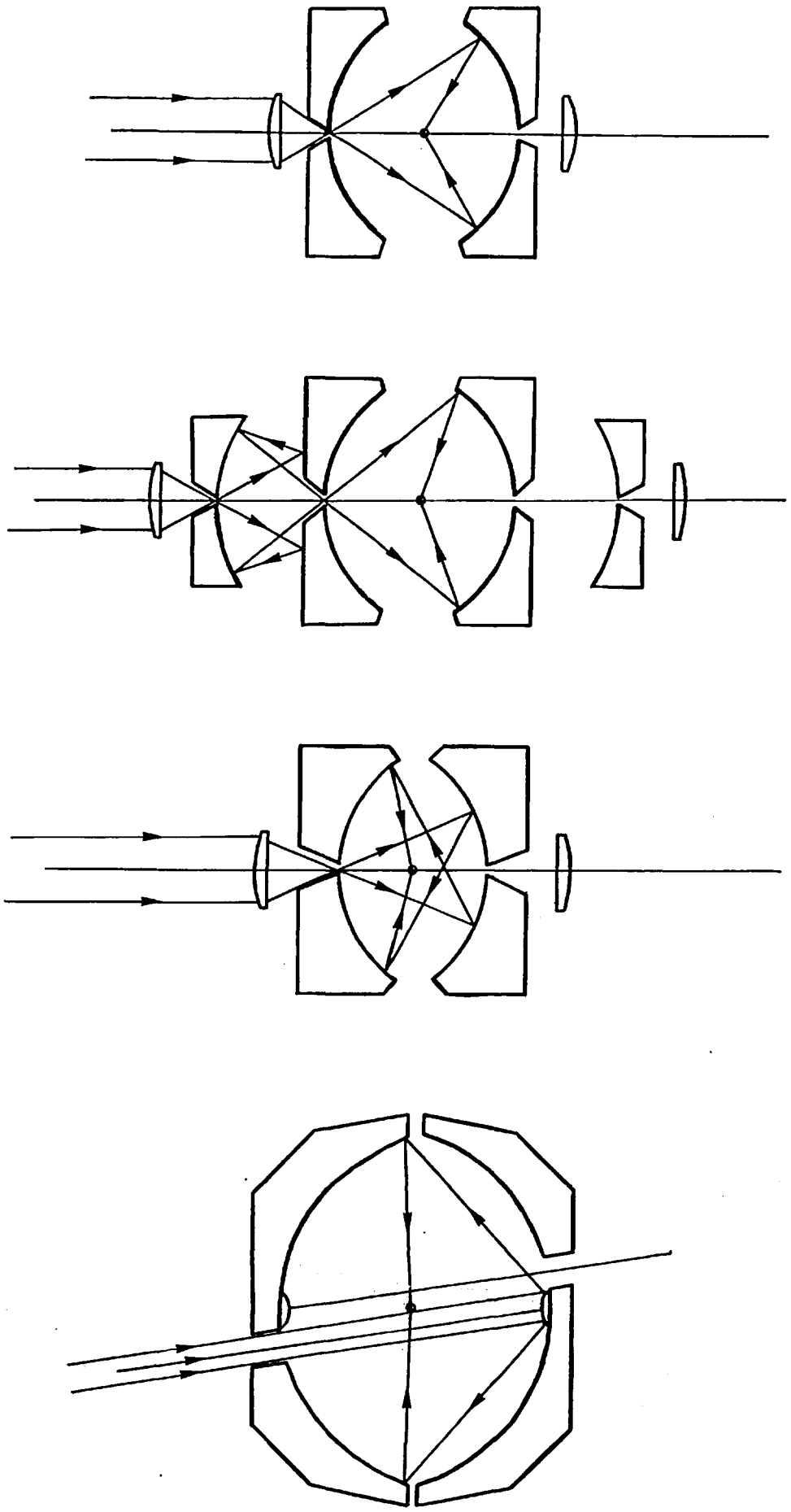


Fig. (1.3) Various systems using mirrors that may be found in the literature.(3),(4),(5),(1)

Advantages

1.) Wavelength invariant properties.

2.) No nonlinear refractive effects.

3.) Improve the illumination orthogonality when used with only a one or two beam laser.

Disadvantages

1.) Systems utilizing mirrors for a multibeam laser become very complicated.

2.) A purely reflecting system can only illuminate the target from a substantial amount of the 4π solid angle if the beam number is low.

3.) The reflective coating lowers the damage threshold for a surface; typically by $\frac{1}{2}$.

Table (1.2) Summary of the advantages and disadvantages of using mirrors.

during manufacture and allows their use at the laser harmonic frequencies.

Another reason for their use, and is certainly the one which is relevant to those designs in Fig. (1.3), is to improve the orthogonality of illumination . With only two beams, a pair of lenses would provide a maximum converging ray angle of about 30^0-40^0, thus giving a 50^0-60^0 angle of incidence on the target. Used in combination with mirrors, however, this ray angle may be increased to more than 60^0 with a corresponding reduction of the incidence angle to less than 30^0. The top three designs show different versions of these "angle amplifying optics".

If the beam number is larger than this, systems of this nature are no longer necessary since the reduced target coverage per beam can be achieved from lenses with a reasonable degree of orthogonality. Designing a mirror system for a large beam number tends to result in a complicated three-dimensional network of criss-crossing beams in order that incident and reflected beams are to be left unobstructed. Alignment problems no doubt increase under such situations. These reasons may have something to do with the fact that all the designs in the literature involving mirrors (to the authors knowledge) are for single or double beam lasers only.

Apart from the less practical setup that would result from using mirrors with the multibeam approach, another disadvantage for a purely reflecting system is that the target can only be illuminated from a substantial amount of the 4π solid angle if the beam number is only one or two.

Thus lenses are capable of providing better orthogonality for large beam numbers. The fourth design (1) in Fig. (1.3) shows a pair of beams incident on small convex parabolic mirrors whose foci coincide with one focus of a pair of larger ellipsoidal mirrors. The second focus of the latter also coincide at the target position. To allow this to work the beams and parabolic primaries are tilted with respect to the axis of the ellipsoidal mirrors. Apart from the shadow areas where the beams enter the cavity and at the primary positions, such a system can illuminate the target from most directions. However, extending such an optical system to cope with many beams would cause these areas of shadow to quickly mount up. Thus the orthogonality cannot really be improved upon.

In the multibeam approach, an approximate limit to the amount of solid angle from which the target could be illuminated is 2π. As viewed from the target, the entrance aperture for each beam into the target chamber may generally subtend as much solid angle as the mirror that focusses it onto the target. Since this has to be left unobstructed, roughly half of the available solid angle is unusable. Of course, this will vary from one configuration to another and with the use of multiple-mirror and off-axis optics.

Although the use of mirrors will reduce the total optical path length inside glass and thereby reduced the nonlinear refractive effects such as self focussing, this merit is counteracted by the fact that the reflective coating lowers the damage threshold to the optical surface; typically by half (2). It is this and the difficulty in applying a uniform dielectric film to a steeply curved surface that favours the

use of lenses. Ideally a lens itself should have an antireflection coating on each surface, but they are not absolutely necessary and have sometimes been dispensed with to avoid damage. The effect of the transmission losses on the target illumination and the more pronounced ghost images that would result from this can be dealt with in ways to be considered later.

To summarize therefore, the use of mirrors is not very convenient for a multibeam laser for the reasons of complicated design and lower orthogonality than a purely refracting system. For only one or two beam work they have been used to improve the orthogonality but even in this case lenses are usually employed to reduce the shadow areas. It should also be pointed out that in such systems, the mirrors shape is determined by the necessary position of the two foci; e.g. one focus at the centre of the cavity where the target is positioned and the second at the pole of the other mirror requires ellipsoidal mirrors of eccentricity $1/3$. Thus the mirror's energy transforming properties will be fixed. The lenses however, can be designed to control the overall energy transformation since they contain a few degrees of freedom. Even when mirrors are useful, the lenses tend to require greater consideration. In this thesis the emphasis is therefore on the refracting optics.

The wavelength invariant properties of a mirror would seem to be a distinct advantage, but in Chapters 5 and 6 some alternative refracting systems with useful properties are considered.

CHAPTER 2

TARGET ILLUMINATION

2.1. Introduction

In order to compress the spherical target symmetrically it seems reasonable that the illumination should be uniform and orthogonal over the entire surface. This would certainly seem to be true in the early stages; before the laser pulse has produced a plasma around the target. Sources of literature (6), (7) seem to agree that once this plasma has developed the uniformity becomes less critical; any non-uniformities tend to be smoothed out by electron transport.

The generally accepted method of illuminating the target is by the multibeam method. Several laser beams are directed onto the target usually in a symmetric pattern determined by the five Platonic solids, each focussed onto a portion of the target. When more than two beams are involved the illumination areas have to overlap slightly if the entire surface is to be covered. This is a disadvantage of the multibeam approach. Regions illuminated by more than one beam have a non-uniform intensity. Interference effects are also produced. These are of a small scale and will be smoothed out in the same manner as other irregularities. Only with two beams can complete coverage occur without beam overlap. Despite this problem and the trouble of alignment the multibeam approach is still the trusted method.

The optical problem is therefore to devise a refracting (and less often a reflecting or catadioptric) system that

transforms the intensity profile of the laser beam onto a portion of the target so as to achieve an acceptable target illumination. To do this we must first be able to determine the energy delivering properties of any focussing system.

2.2. Calculating the target intensity
strictly target illumination

Consider an optical system that focusses a plane wavefront of the incident beam onto a portion of the spherical target (Fig (2.1)). The focussing optics can be characterized by a principal surface, the aberrations and the focal length. With correctly aligned optics and target, spherical aberration is only of concern. The target is positioned in the beam so that the desired coverage is obtained. A ray incident on the system at a height Y is directed onto the target at a position defined by the "target angle" ϕ. The marginal ray at a height Y_m strikes the target at the maximum angle ϕ_m. ϕ_m defines the target "coverage" per beam.

An annular ring of the beam with area, $dA_B=2\pi YdY$ is mapped onto the target area, $dA_T=2\pi R_T^2 \sin\phi d\phi$ by a one-to-one relation. By "one-to-one" we mean that the energy flowing through the area dA_B does not fall onto more than one area dA_T and further, that the energy falling on dA_T does not originate from more than one annular area, dA_B. The former is impossible since it requires the non-uniqueness of a ray; two or more rays in image space would have to originate from a single incident ray. The latter is possible but undesirable. It would mean that the target surface intersects a caustic

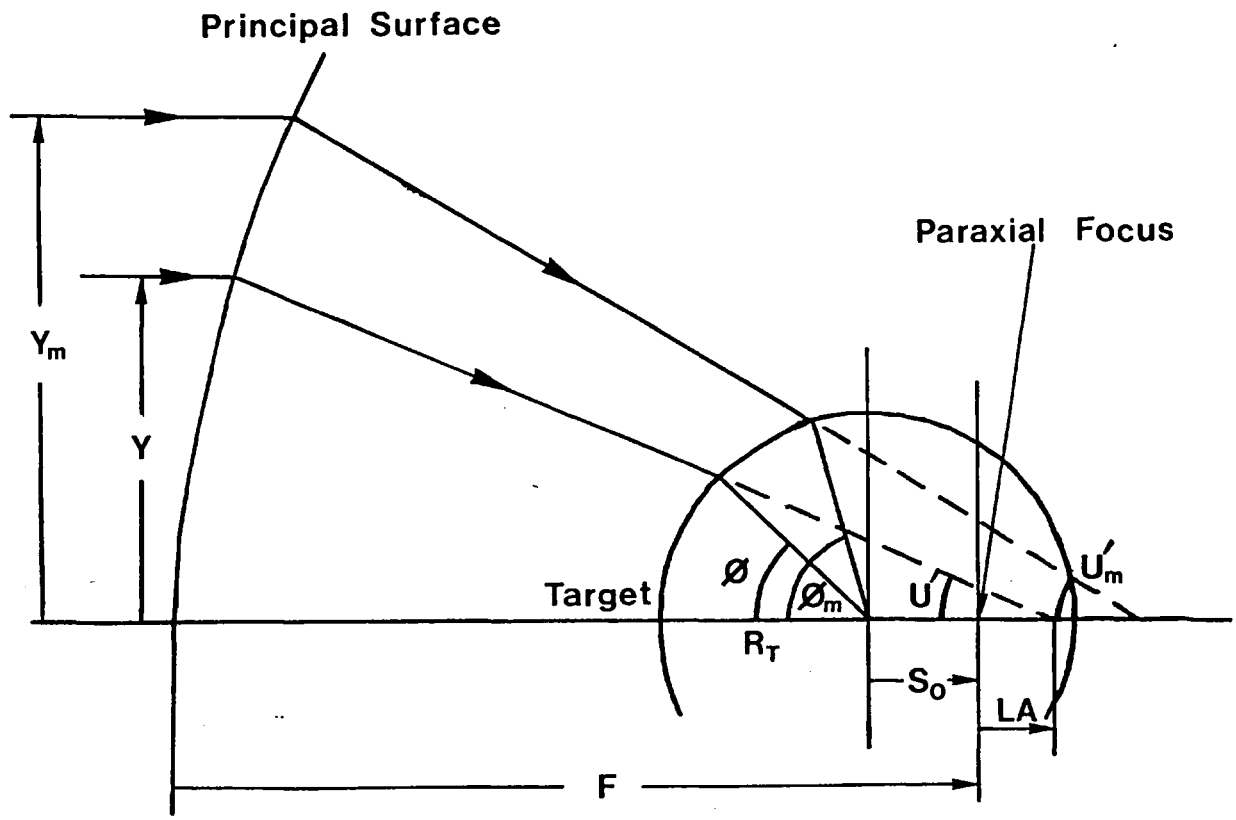


Fig. (2.1) Target illumination geometry. Not to scale.

surface, producing a ring of very high intenstiy on the pellet. Since the aim is to produce a reasonably uniform illumination this situation is to be avoided.

Following Thomas (3), let the intensity profiles across the beam and target be $I_B(Y)$ and $I_T(\phi)$ respectively. Then by the conservation of energy,

$$I_T(\phi)dA_T = I_B(Y)dA_B \tag{2.1}$$

or

$$I_T(\phi) = I_B(Y)\frac{Y}{R_T^2 \sin\phi}\frac{dY}{d\phi} \tag{2.2}$$

Also by geometry,

$$\tan U' = \frac{\sin\phi}{\cos\phi + \frac{S}{R_T}} \tag{2.3}$$

where S represents the distance from the target center to the point where the particular ray crosses the axis. It may be written as,

$$S = S_0 + LA \tag{2.4}$$

Here S_0 is the target center-paraxial focus distance and LA the longitudinal aberration.

For a particular coverage, the target position can be found using equations (2.3) and (2.4) for the marginal ray.

$$\frac{Sm}{R_T} = \frac{S_0 + LA_m}{R_T} = \frac{\sin\phi_m}{\tan U'_m} - \cos\phi_m \tag{2.5}$$

Raytracing can give the mapping $Y=Y(\phi)$ and $dY/d\phi$ in numerical form, from which $I_T(\phi)$ is easily found using (2.2). In some special cases of interest $I_T(\phi)$ may be evaluated

analytically. The inverse problem is also of importance; namely, to deduce the principal surface $Y=Y(U')$ that produces a desired target illumination.

Thomas assumed his lenses to have a flat principal surface,

$$\tan U' = -\frac{Y}{F} \qquad (2.6)$$

As pointed out by Kleinhans (8) this assumption is unrealistic in that the coma would be intolerable even for a slight misalignment. It would therefore be wise to look at the coma in each case of interest.

On moving away from the axis the aberration depending linearly on the field coordinate is the first to become significant. These are the "linear coma" terms whose wavefront error are of the form,

$$W_J = C_J(X_p^2 + Y_p^2)^J Y_p \eta' \qquad (J = 1,2,3,\ldots) \qquad (2.7)$$

The J=1 term constitutes the familiar primary coma. Since we are dealing with high apertures the terms of higher order in the pupil coordinates (Xp, Yp) will become appreciable. Thus all orders (J-values) need to be considered together.

A quantity which gives a measure of linear coma for all orders was given by Conrady (9) called the "offense against the sine condition"; and is denoted by OSC. It is quite a remarkable expression that gives the linear coma for an off-axis sagittal ray in terms of a ray imaging an axial object point.

$$OSC = \frac{u'}{\sin U'} \frac{\sin U}{u} \frac{R'}{(R'+LA)} - 1 \qquad (2.8)$$

Here R' is the radius of the reference sphere (distance from the exit pupil to the paraxial image plane). Thus as expected, the coma is dependent on the stop position. A derivation of this is not given here in the main body of the text since it would entail a too lengthy digression. A more detailed exposition is the subject of Appendix D.

In the limit of an infinite object point, the factor sin U/u tends to unity, and when spherical aberration is absent (LA=0),

$$OSC = \frac{u'}{\sin U'} - 1 \qquad (2.9)$$

Using (2.9) a system with a flat principal surface has an OSC given by,

$$OSC = \sqrt{1 + Y^2/F^2} - 1 \qquad (2.10)$$

and is plotted in Fig.(2.2) for an F/1 example along with other cases to be considered. Preferably the system should have some correction for coma. An aplanat has a spherical principal surface,

$$\sin U' = -\frac{Y}{F} \qquad (2.11)$$

This case will be looked at in detail later. Firstly, because we are striving for a uniform illumination it seems obvious that the next step should be to find the principal surface that does exactly that .

2.3. Conditions for uniform target illumination

If I_T is required to be constant then equation (2.2)

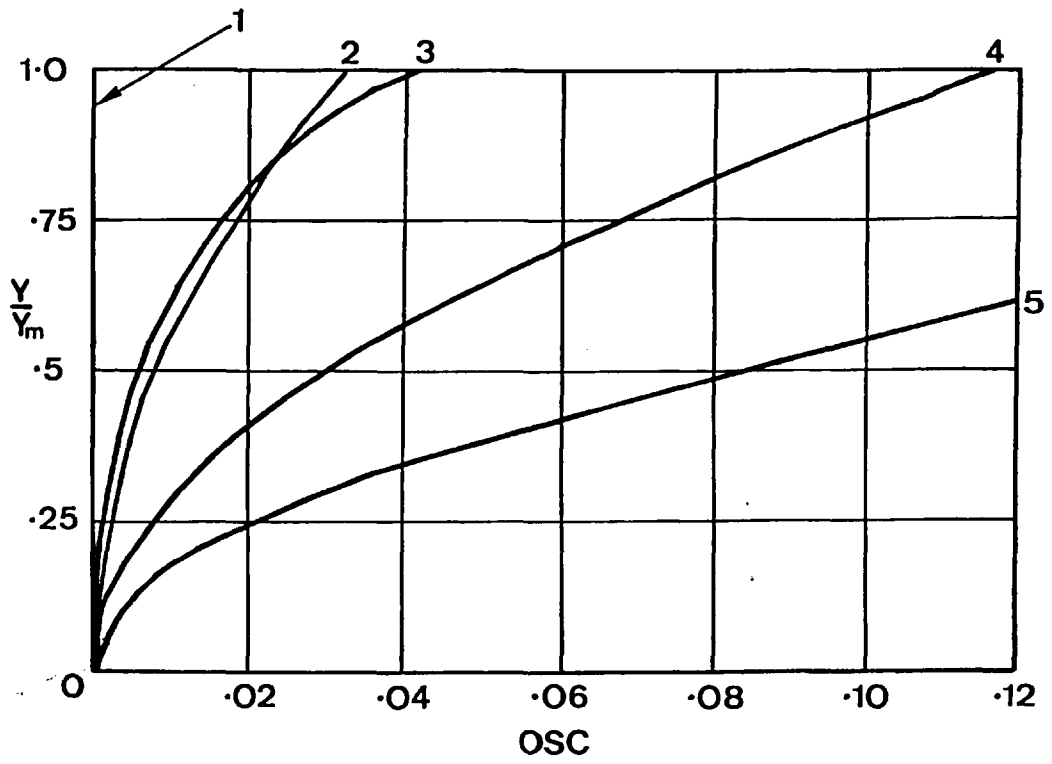


Fig. (2.2) OSC variation with aperture for various principal surfaces. Case of an F/1 lens.

1. Aplanat $\sin U' = -Y/F$
2. $\sin(U'/2) = -Y/2F$
3. $\tan U' = \dfrac{-Y\sqrt{1-Y^2\sin^2(\phi_m/2)/Y_m^2}}{F+Y^2\sin(\phi_m/2)/Y_m}$
4. $\tan U' = -Y/F$
5. F/1 lens designed to provide a uniform illumination from a Gaussian beam. $F=110\text{mm}, Y_m=55\text{mm}, \phi_m=-90^0, g=55\text{mm}$.

becomes, on separating the variables,

$$I_T R_T^2 \sin\phi d\phi = I_B(Y) Y dY \tag{2.12}$$

Integrating gives,

$$-I_T R_T^2 \cos\phi + \text{constant} = \int I_B(Y) Y dY \tag{2.13}$$

Unless otherwise stated, we shall assume that the intensity profile across the laser beam is also uniform. Usually some attempt at smoothing out the beam is made anyway. If a Gaussian or super-Gaussian profile is in fact involved it is easy to obtain an idea of the illumination by simply weighting $I_T(\phi)$ for constant I_B according to the actual $I_B(Y)$. This can be done simply by looking at a graph of $I_T(\phi)$.

Therefore, we have

$$-I_T R_T^2 \cos\phi + \text{constant} = \frac{I_B Y^2}{2} \tag{2.14}$$

When $Y = 0$, $\phi = 0$ so,

$$I_T R_T^2 (1-\cos\phi) = \frac{I_B Y^2}{2} \tag{2.15}$$

In particular, this is valid for the marginal zone,

$$I_T R_T^2 (1-\cos\phi_m) = \frac{I_B Y_m^2}{2} \tag{2.16}$$

Dividing the first by the second,

$$\frac{Y^2}{Y_m^2} = \frac{1-\cos\phi}{1-\cos\phi_m} \tag{2.17}$$

This can be written in another form by means of the identity, $\cos\phi = 1 - 2\sin^2(\phi/2)$

$$\frac{Y}{Y_m} = \frac{\sin(\phi/2)}{\sin(\phi_m/2)} \tag{2.18}$$

Rays that enter the optics at a height Y must hit the target at the angle ϕ (see Fig. (2.1)) given by this relation. Eliminating ϕ between this and equation (2.3) gives the principal surface shape,

$$\tan U' = \frac{\frac{2Y\sin(\phi_m/2)}{Y_m}\sqrt{1-\frac{Y^2}{Y_m^2}\sin^2(\frac{\phi_m}{2})}}{1-\frac{2Y^2}{Y_m^2}\sin^2(\frac{\phi_m}{2})+\frac{S}{R_T}} \tag{2.19}$$

The principal surface depends on the coverage angle, ϕ_m, so the system will only provide a uniform illumination when used for this amount of coverage

To find the correct target position relative to the focus, equation (2.19) is paraxialized; i.e let $\tan U' \to u'$, $S \to S_0$, $Y \to y$ and powers of y greater than one are ignored. This gives,

$$u' = \frac{2y\sin(\phi_m/2)}{Y_m(1+\frac{S_0}{R_T})} \tag{2.20}$$

and means that the focal length is given by,

$$-\frac{1}{F} = \frac{2\sin(\phi_m/2)}{Y_m(1+S_0/R_T)} \tag{2.21}$$

For the time being we will consider cases when no aberrations are present. So LA=0 and $S=S_0$. Rearranging then gives,

$$\frac{S_0}{R_T} = \frac{-2F\sin(\phi_m/2)}{Y_m} - 1 \tag{2.22}$$

Substituting this in (2.19),

$$\tan U' = \frac{-Y\sqrt{1-Y^2\sin^2(\phi_m/2)/Y_m^2}}{F + Y^2\sin(\phi_m/2)/Y_m} \qquad (2.23)$$

Since the principal surface shape is determined by the illumination profile it follows that the amount of coma cannot be controlled. Using (2.9) for OSC, sin U′ can be found from (2.23) while the paraxial ray angle, u′ = -Y/F. These then give the result,

$$OSC = \sqrt{\frac{1 + \frac{Y^2}{F^2}\left(\frac{2F}{Y_m}\sin\left(\frac{\phi_m}{2}\right) + 1\right)}{1 - \frac{Y^2}{Y_m^2}\sin^2\left(\frac{\phi_m}{2}\right)}} - 1 \qquad (2.24)$$

As an example consider an F/1 lens (e.g. one that focusses a beam of radius Ym = 55mm and therefore has a focal length, F = 110mm). If it is designed to illuminate a hemisphere of the target, $\phi_m = -90^0$. The OSC given by (2.24) is plotted in Fig. (2.2).

2.4. Uniform illumination from a Gaussian beam

If the beam intensity is Gaussian and written in the form,

$$I_B(Y) = I_o e^{-Y^2/g^2} \qquad (2.25)$$

so that the power in the beam is given by

$$2\pi \int_0^{Y_m} I_B(Y) Y dY = \pi g^2 I_o (1 - e^{-Y_m^2/g^2})$$

then substituting into equation (2.13), integrating and applying the same boundary conditions,

$$e^{-Y^2/g^2} = \frac{(1 - e^{-Y_m^2/g^2})\cos\phi - (\cos\phi_m - e^{-Y_m^2/g^2})}{(1 - \cos\phi_m)} \qquad (2.26)$$

This replaces equation (2.18) for a uniform beam intensity. Along with (2.3) this gives the principal surface and in particular,

$$\tan U_m' = \frac{\sin\phi_m}{\cos\phi_m - \frac{2F}{g}\sin(\phi_m/2) - 1} \qquad (2.27)$$
$$\sqrt{1 - e^{-Y_m^2/g^2}}$$

For Y_m = 55mm, F = 110mm, $\phi_m = -90^0$ and g = 55mm also, then $U_m' = -21{\cdot}36^0$ and $OSC_m = 0{\cdot}373$ (see Fig. (2.2)). This amount of coma is more than three times that produced by a flat principal surface. Here it is actually curved backwards. This explains why U_m' is so shallow; -30^0 is nearer the value one would expect. It seems unlikely that such a system would be of much use. Rather, it suggests that some attempt should be made at smoothing out the beam.

2.5. The Aplanat

As just illustrated, for the high apertures usually required the focussing optics may be totally useless if the amount of coma present is too large. It would be worthwhile, therefore, to look at the target illumination produced by an aplanat.

With OSC = 0 and $u' = -Y/F$, (2.9) gives

$$\sin U' = -\frac{Y}{F} \tag{2.28}$$

or

$$\frac{Y}{Y_m} = \frac{\sin U'}{\sin U'_m} \tag{2.29}$$

where the focal length, F, is

$$F = \frac{-Y_m}{\sin U'_m} \tag{2.30}$$

The principal surface is therefore a sphere of radius F. Now with,

$$Y = -F\sin U' = \frac{-F\tan U'}{\sqrt{1 + \tan^2 U'}} \tag{2.31}$$

then using equation (2.3),

$$Y = \frac{-F\sin\phi}{\sqrt{\left(\frac{S_o}{R_T}\right)^2 + 2\left(\frac{S_o}{R_T}\right)\cos\phi + 1}} \tag{2.32}$$

The required derivative is,

$$\frac{dY}{d\phi} = \frac{-F\cos\phi}{\sqrt{\left(\frac{S_o}{R_T}\right)^2 + 2\left(\frac{S_o}{R_T}\right)\cos\phi + 1}} - \frac{\left(\frac{S_o}{R_T}\right)F\sin^2\phi}{\left[\left(\frac{S_o}{R_T}\right)^2 + 2\left(\frac{S_o}{R_T}\right)\cos\phi + 1\right]^{3/2}} \tag{2.33}$$

$$\frac{dY}{d\phi} = \frac{-F\left(\frac{S_o}{R_T} + \cos\phi\right)\left(\frac{S_o}{R_T}\cos\phi + 1\right)}{\left[\left(\frac{S_o}{R_T}\right)^2 + 2\left(\frac{S_o}{R_T}\right)\cos\phi + 1\right]^{3/2}} \tag{2.34}$$

Substituting this and (2.32) into the expression for $I_T(\phi)$,

$$I_T(\phi) = \frac{I_B F^2}{R_T^2} \cdot \frac{\left(\frac{S_o}{R_T} + \cos\phi\right)\left(\frac{S_o}{R_T}\cos\phi + 1\right)}{\left[\left(\frac{S_o}{R_T}\right)^2 + 2\left(\frac{S_o}{R_T}\right)\cos\phi + 1\right]^2} \qquad (2.35)$$

The target location is obtained from (2.5) and (2.31) evaluated at the marginal zone,

$$\frac{S_o}{R_T} = \frac{\sin\phi_m}{\tan U'_m} - \cos\phi_m \qquad (2.36)$$

or

$$\frac{S_o}{R_T} = -\sin\phi_m \sqrt{\frac{F^2}{Y_m^2} - 1} - \cos\phi_m \qquad (2.37)$$

The quantity $R_T^2 I_T(\phi)/F^2 I_B$ is plotted against ϕ for various coverages, ϕ_m, in Fig. (2.3). Although ϕ is negative, the sign is left out of the labelling on these graphs. The particular example is for an F/1 aplanat. (F/Ym=2). It shows that an aplanat is capable of very good uniformity for most degrees of coverage. The intensity does tend to drop away for the outer zones however. This effect will be increased if the incident beam also drops away at the edge, as is usually the case.

For the one case of $\phi_m=-60^0$ the illumination is perfectly uniform. If the target is positioned so that the beam is focussed on its second pole (Fig.(2.4)), then $S_o/R_T=1$, and it can be seen by geometry that for all rays, $\phi=2U'$. Putting this in equation (2.18) immediately gives the condition for aplanatism, equation (2.29). The target intensity reduces to,

$$I_T = \frac{I_B F^2}{4R_T^2} \qquad (2.38)$$

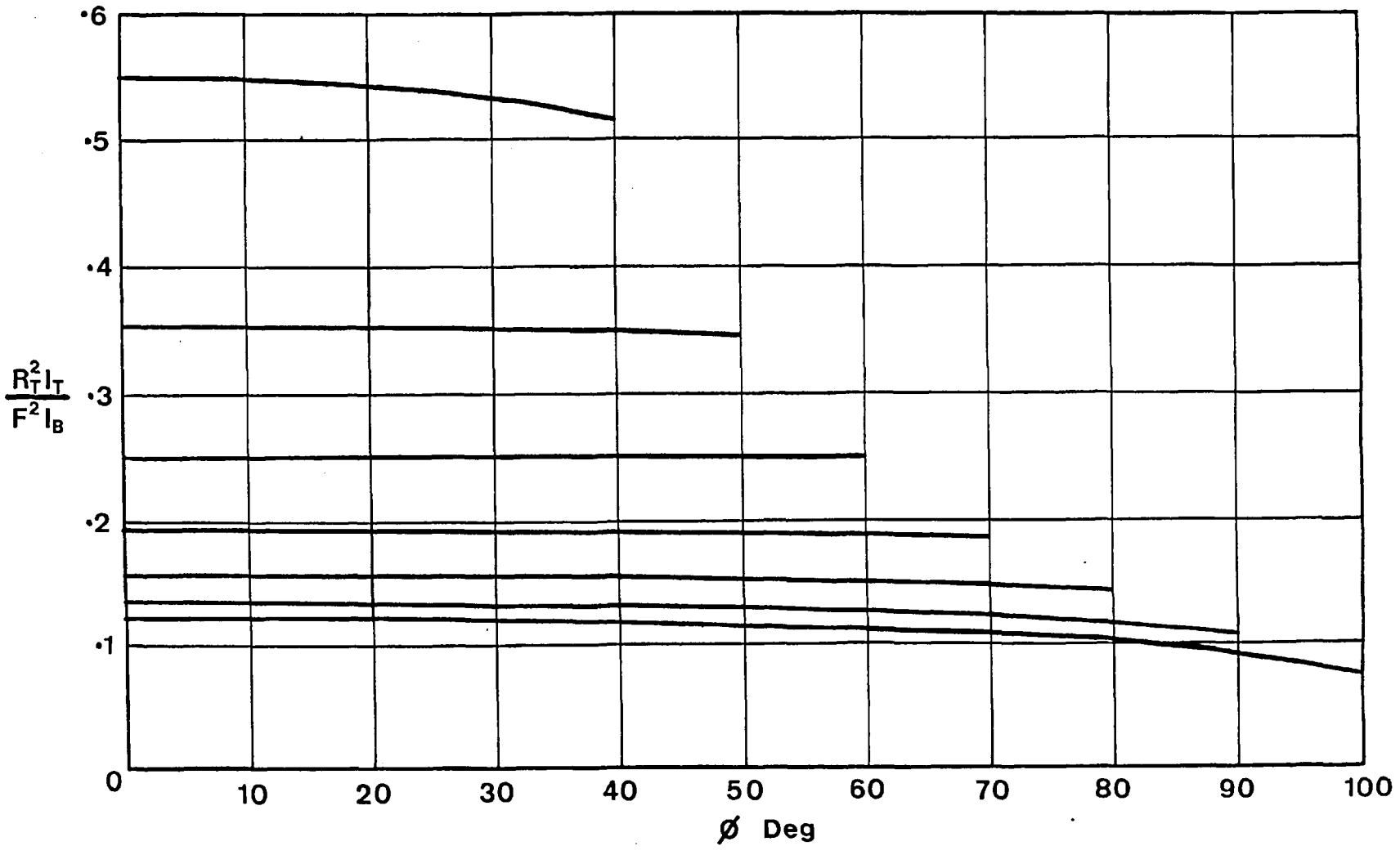


Fig. (2.3) Target illumination provided by an aplanat of F/1.

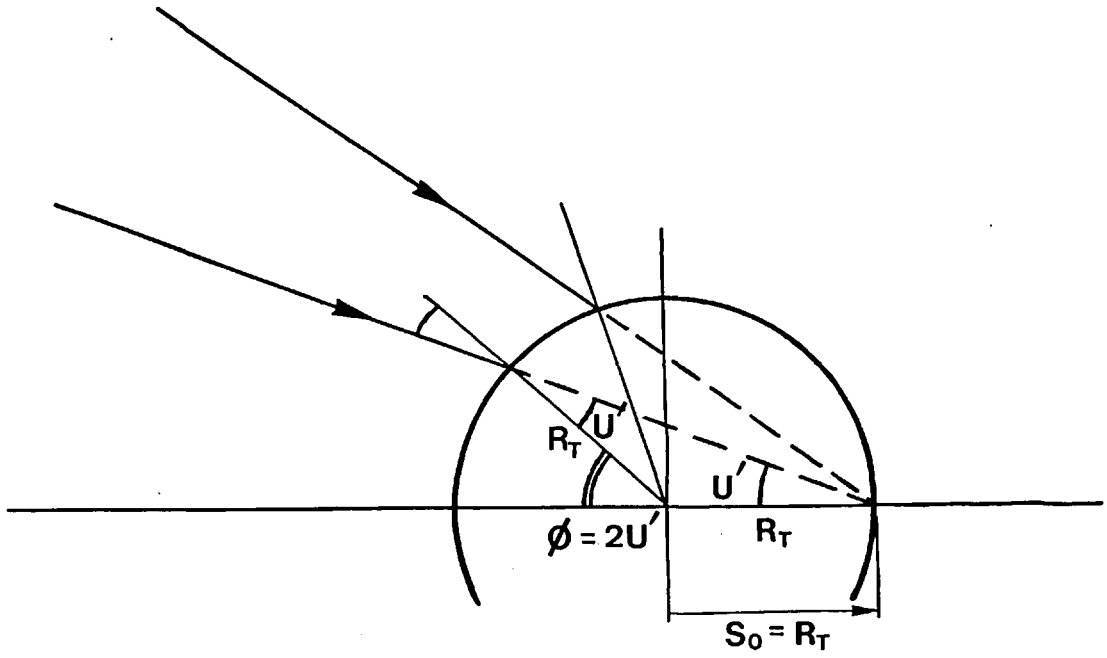


Fig. (2.4) Condition for a uniform illumination from an aplanat. The focus is positioned at the far pole of the target.

An aplanat with a maximum ray angle of U'_m can uniformly illuminate a portion of the target defined by $\Phi m = 2U'_m$. A price to pay for this is that the angle of incidence is equal to the ray angle, U'. As the lens aperture increases the orthogonality will be reduced.

2.6. Uniform and Orthogonal Illumination

If many beams are used then each only needs to cover a small portion of the target. Consequently the orthogonality can be greatly improved over a two-beam laser using similar optics, for example.

In order to improve the orthogonality for the latter various "angle-amplifying" systems have been devised (3),(4), (5). The simplest consists of a lens that converges the incoming beam down to one focus of an ellipsoidal mirror (Fig. (2.5)). The target is therefore positioned at the second focus of the ellipse. After reflection the final ray angle may be 60° or more, depending on the lens aperture.

The condition for perfectly uniform and orthogonal illumination follows from equation (2.3) as a special case. In order to achieve orthogonality the target center and focus must coincide, i.e. $S_0 = 0$. Therefore $\Phi = U'$ and so the principal surface is,

$$\frac{Y}{Y_m} = \frac{\sin(U'/2)}{\sin(U'_m/2)} \qquad (2.39)$$

This is Herschel's condition. It is more commonly known as the requirement of a system in order that the aberrations

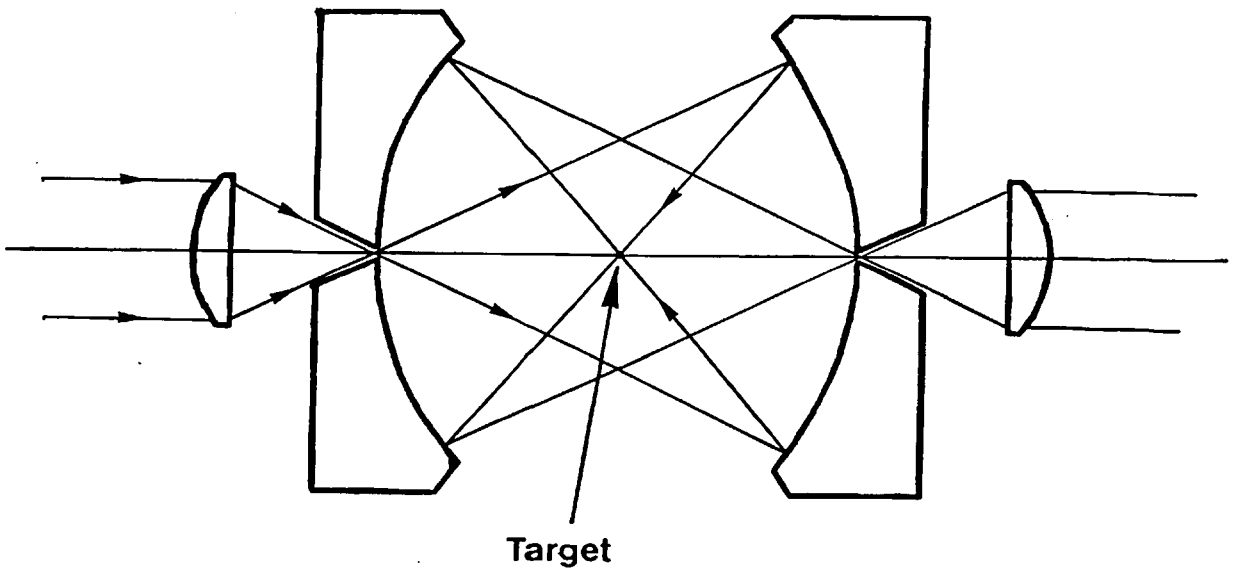


Fig. (2.5) The Thomas ellipsoidal mirror configuration.

remain unchanged for a small longitudinal displacement of the object; a "longitudinal" equivalent of the sine condition.

The focal length is given from above as,

$$F = \frac{Y_m}{2\sin(U'_m/2)} \tag{2.40}$$

Herschel's condition can then be written,

$$Y = -2F\sin(U'/2) \tag{2.41}$$

It then follows that,

$$\frac{dY}{dU'} = -F\cos(U'/2) \tag{2.42}$$

$$Y\frac{dY}{dU'} = 2F^2\sin(U'/2)\cos(U'/2) \tag{2.43}$$

$$= F^2\sin U' \tag{2.44}$$

$$= \frac{F^2\tan U'}{\sqrt{1 + \tan^2 U'}} \tag{2.45}$$

Using (2.3),

$$Y\frac{dY}{dU'} = \frac{F^2\sin\phi}{\sqrt{\left(\frac{S_0}{R_T}\right)^2 + 2\left(\frac{S_0}{R_T}\right)\cos\phi + 1}} \tag{2.46}$$

Differentiating (2.3) with respect to ϕ will result in,

$$\frac{dU'}{d\phi} = \frac{\left(\frac{S_0}{R_T}\right)\cos\phi + 1}{\left[\left(\frac{S_0}{R_T}\right)^2 + 2\left(\frac{S_0}{R_T}\right)\cos\phi + 1\right]} \tag{2.47}$$

Finally therefore,

$$I_T(\phi) = \frac{I_B F^2 \left(\frac{S_0}{R_T}\cos\phi + 1\right)}{R_T^2 \left[\left(\frac{S_0}{R_T}\right)^2 + 2\left(\frac{S_0}{R_T}\right)\cos\phi + 1\right]^{3/2}} \tag{2.48}$$

When $S_0 = 0$, this gives $I_T = I_B F^2/R_T^2$, a constant as expected.

The target position is given from (2.5),

$$\frac{S_0}{R_T} = \frac{\sin\phi_m}{\tan U_m'} - \cos\phi_m \tag{2.49}$$

$$\text{where } \tan U_m' = \frac{2\sin(U_m'/2)\cos(U_m'/2)}{1 - 2\sin^2(U_m'/2)} \tag{2.50}$$

$$\text{and } \sin(U_m'/2) = -Y_m/2F \tag{2.51}$$

so,

$$\frac{S_0}{R_T} = \frac{\left(\frac{Y_m^2}{2F^2} - 1\right)\sin\phi_m}{\frac{Y_m}{F}\sqrt{1 - \frac{Y_m^2}{4F^2}}} - \cos\phi_m \tag{2.52}$$

The amount of linear coma is easily found to be,

$$OSC = \frac{1}{\sqrt{1 - Y^2/4F^2}} - 1 \tag{2.53}$$

Fig.(2.6) illustrates how the target intensity varies with different coverages, ϕ_m. Again, for comparison, the F/1 example was considered. As in the case of an aplanat the uniformity is good for most coverages. Fig. (2.2) shows the OSC variation with aperture.

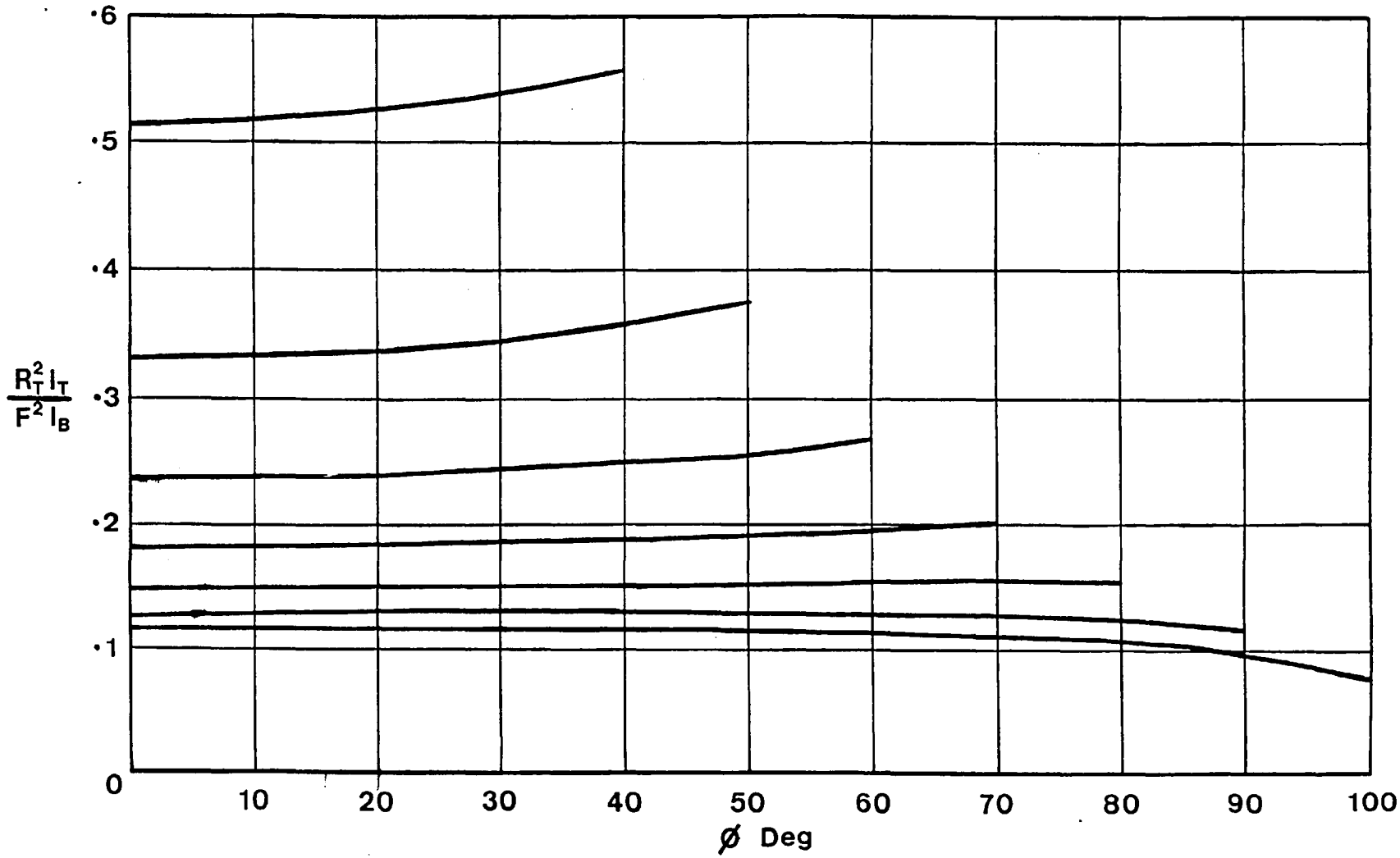


Fig. (2.6) Target illumination provided by a F/1 lens that satisfies Herschel's condition.

2.7. Lens-mirror combinations

As mentioned in Chapter 1 the mirror elements that may be used to improve orthogonality will usually have their shape determined by geometry. For instance, in the Thomas ellipsoidal mirror system of Fig. (2.5) the necessary position of the foci means that the mirrors must have an eccentricity of 1/3. The two lenses, on the other hand, may be used to control the target intensity.

Suppose that we require the system in Fig. (2.5) to uniformly illuminate the whole target; then $\Phi_m = -90^0$. If the incident ray height is Y and the ray angles are U' and U'' before and after reflection respectively, then (2.23) means that the system as a whole must satisfy,

$$\tan U'' = \frac{-Y\sqrt{1 - Y^2/2Y_m^2}}{F - Y^2/Y_m\sqrt{2}} \qquad (2.54)$$

But it may be shown that the angles U' and U'' are related by,

$$\tan U'' = \frac{4 \sin U'}{5\cos U' - 3} \qquad (2.55)$$

So the lens is required to have a principal surface given by,

$$\frac{4\sin U'}{5\cos U' - 3} = \frac{-Y\sqrt{1 - Y^2/2Y_m^2}}{F + Y^2/Y_m\sqrt{2}} \qquad (2.56)$$

In the paraxial limit this reduces to,

$$2u' = -y/F \qquad (2.57)$$

or

$$u' = -y/2F \qquad (2.58)$$

The focal length of the lens is therefore twice that of the whole system. i.e. the power is doubled by the mirror.

$$F_{lens} = 2F \tag{2.59}$$

If F_{lens} = 110mm, Y_m = 55mm then $U''_m = -67{\cdot}5^0$ and $U'_m = -36{\cdot}948^0$, so that the angle of incidence is halved. The amount of coma present is large, however; $OSC_m = 0{\cdot}0824$.

2.8. Spherical aberration required to uniformly illuminate the target

In all the cases previously discussed we have made the assumption that the focussing optics satisfy the constant optical path condition. This is no doubt a natural assumption to make, but is by no means necessary. When made, the intensity on the target is independent of target size; if the illumination is found to be uniform for one target size it will also be uniform for any other size, provided the angle ϕ_m is the same. If the constant optical path condition is not satisfied this is no longer true.

Thomas (3) illustrated by way of example, that given a system with a known principal surface, which in general will not provide uniform illumination, uniformity can be achieved by selectively aberrating the system.

The target position relative to the focus was given by,

$$\frac{S}{R_T} = \frac{\sin \phi}{\tan U'} - \cos \phi \tag{2.60}$$

or

$$\frac{S}{R_T} = \frac{2\sin(\phi/2)\cos(\phi/2)}{\tan U'} - (1 - 2\sin^2(\phi/2)) \tag{2.61}$$

S is no longer a constant if aberrations are present, but a function of ray height Y.

$$S(Y) = S_0 + LA(Y) \tag{2.62}$$

The condition for uniform target intensity was,

$$\frac{Y}{Y_m} = \frac{\sin(\phi/2)}{\sin(\phi_m/2)} \tag{2.63}$$

Writing $\rho = Y/Y_m$, the normalized ray height then,

$$\sin(\phi/2) = \rho \sin(\phi_m/2) \tag{2.64}$$

Substituting (2.62) and (2.64) into (2.61),

$$\frac{S_0 + LA(\rho)}{R_T} = \frac{2\rho\sin(\phi_m/2)\sqrt{1 - \rho^2\sin^2(\phi_m/2)}}{\tan U'} + 2\rho^2\sin^2(\phi_m/2) - 1 \tag{2.65}$$

This is a generalization of the result obtained by Thomas. His result follows by setting $\tan U' = -\rho Y_m/F$ and $\phi_m = -90^\circ$. In the paraxial limit, $\rho \to 0$, $LA \to 0$ and $\tan U' \to -\frac{\rho Y_m}{F}$, this becomes,

$$\frac{S_0}{R_T} = -\frac{2F}{Y_m}\sin(\phi_m/2) - 1 \tag{2.66}$$

Therefore,

$$\frac{LA(\rho)}{R_T} = \frac{2\rho\sin(\phi_m/2)\sqrt{1-\rho^2\sin^2(\phi_m/2)}}{\tan U'} + 2(\rho^2\sin(\phi_m/2) + F/Y_m)\sin(\phi_m/2) \tag{2.67}$$

or in terms of transverse ray aberrations, $T(\rho) = -LA(\rho)\tan U'$,

$$\frac{T(\rho)}{R_T} = -2\rho\sin(\phi_m/2)\sqrt{1-\rho^2\sin^2(\phi_m/2)} - 2(\rho^2\sin(\phi_m/2) + F/Y_m) \cdot \sin(\phi_m/2)\tan U' \tag{2.68}$$

Thus if the aberrations, $T(\rho)$, and the principal surface tan $U'(\rho)$, satisfy this relation for all zones, the focussing optics will uniformly illuminate the target. When $T(\rho) = 0$ and we solve for tan U', the result is expression (2.23) as expected.

The aberrations can be expanded as a power series in ρ to find the Seidel and higher order contributions. Note that the amount required is proportional to the target size. Only if the aberrations vanish does the R_T also disappear. Satisfying the constant optical path has this distinct advantage.

Thomas has suggested that the aberration given by (2.67) or (2.68) can be introduced by using corrector plates in the parallel part of the beam, or alternatively by a small relative displacement of the focussing elements, (Brueckner and Howard (10)). This latter method only approximates the aberration required.

A method by which these corrector plates may be designed will be left until later. In view of what has been said, the obvious disadvantage is that a new plate is required for each target size. The second method has the advantage that different target sizes and coverages may be accommodated by varying the element separation. This will also be investigated.

Equation (2.68) also supports the idea of a tunable lens: one that may be conveniently tested at a visible wavelength and then can be re-optimized at the laser wavelength for any target size, R_T and coverage, ϕ_m. These will be discussed in Chapter 5.

CHAPTER 3

DESIGN OF LENSES WITH A SPECIFIED PRINCIPAL SURFACE

3.1. Introduction

In the previous chapter several types of system were considered for the illumination they provide. The optics are characterized by the principal surface (a relation between Y and U') and also satisfied the constant optical path condition. Indeed, for any desired target intensity $I_T(\phi)$ Chapter 2 shows how the required principal surface shape may be found. The purpose of this chapter is to design a set of optical surfaces that have this principal surface and obey the constant optical path condition. To achieve these requirements two aspheric surfaces are necessary. This "generalized" sine condition problem has been solved by Vaskas (11) for applying to any optical system, though here such a general treatment is not necessary.

Since it is desired to keep the optical path in glass to a minimum, the number of elements is generally never greater than two. Therefore we only need to look at the cases of a singlet and doublet.

Of these two choices the doublet is preferable, for two reasons. Firstly, the final surface of the system is exposed to plasma damage when used to focus directly onto the target and will require re-polishing from time to time. For a singlet this surface has to be aspheric, making the polishing a costly process. Secondly, the manufacture of the aspheric surfaces will be easier if they are on separate elements. Aligning the surfaces of an optical system to a common axis is more critical

if a surface is aspheric. For the singlet this has to be done with the grinding and polishing. A doublet allows this to be left until mounting. Other undesirable characteristics of the singlet come to light when examples are actually designed. These will be mentioned in section 3.3.

The examples of the last chapter will be used in the designs. Their principal surfaces are given by equations (2.23), (2.28), (2.41) and (2.26). They are summarized in table (3.1).

3.2. The singlet

A method for designing the singlet was given by Greenbaum, Glass and Trenholme (12). It will be essentially their method we will give here.

Fig.(3.1) shows the singlet and the definition of the various symbols. Careful attention has to be given to the sign of each quantity.

A plane wavefront with maximum ray height, Y_m, is transformed to a spherical wavefront with a maximum ray angle of U'_m. Ray height and ray angle are related by a known expression $U' = U'(Y)$. The aspheric surfaces to be found are defined by the coordinates (Y,Z) and (Y',Z') as shown. Note that Z is measured from the rim.

To keep the glass thickness to a minimum, the surfaces are designed so that they meet at the maximum ray height, Y_m. In practice, the value of Y_m is then chosen to be slightly larger than the beam radius; this allows for mounting.

This minimum thickness, d, can be found as follows; if b is the required back focal length and Z_a the value of Z for the axial point, then the geometrical distance d + b is

Transformation	Principal surface $U' = U'(Y)$
Uniform illumination of target (Equal beam areas mapped onto equal target areas)	$\tan U' = \dfrac{-Y\sqrt{1-Y^2\sin^2(\phi_m/2)/Y_m^2}}{F + Y^2\sin(\phi_m/2)/Y_m}$
Aplanat	$\sin U' = -Y/F$
Uniform and orthogonal illumination (Equal beam areas mapped onto equal solid angles)	$\sin(U'/2) = -Y/2F$
Uniform illumination from a Gaussian input beam	$\tan U' = \dfrac{\sin\phi}{S_0/R_T + \cos\phi}$ where, $\cos\phi = \dfrac{(1-\cos\phi_m)e^{-Y^2/g^2} + (\cos\phi_m - e^{-Y_m^2/g^2})}{(1-e^{-Y_m^2/g^2})}$ $S_0/R_T = \dfrac{\sin\phi_m}{\tan U'_m} - \cos\phi_m$

Table (3.1) List of principal surfaces of particular interest.

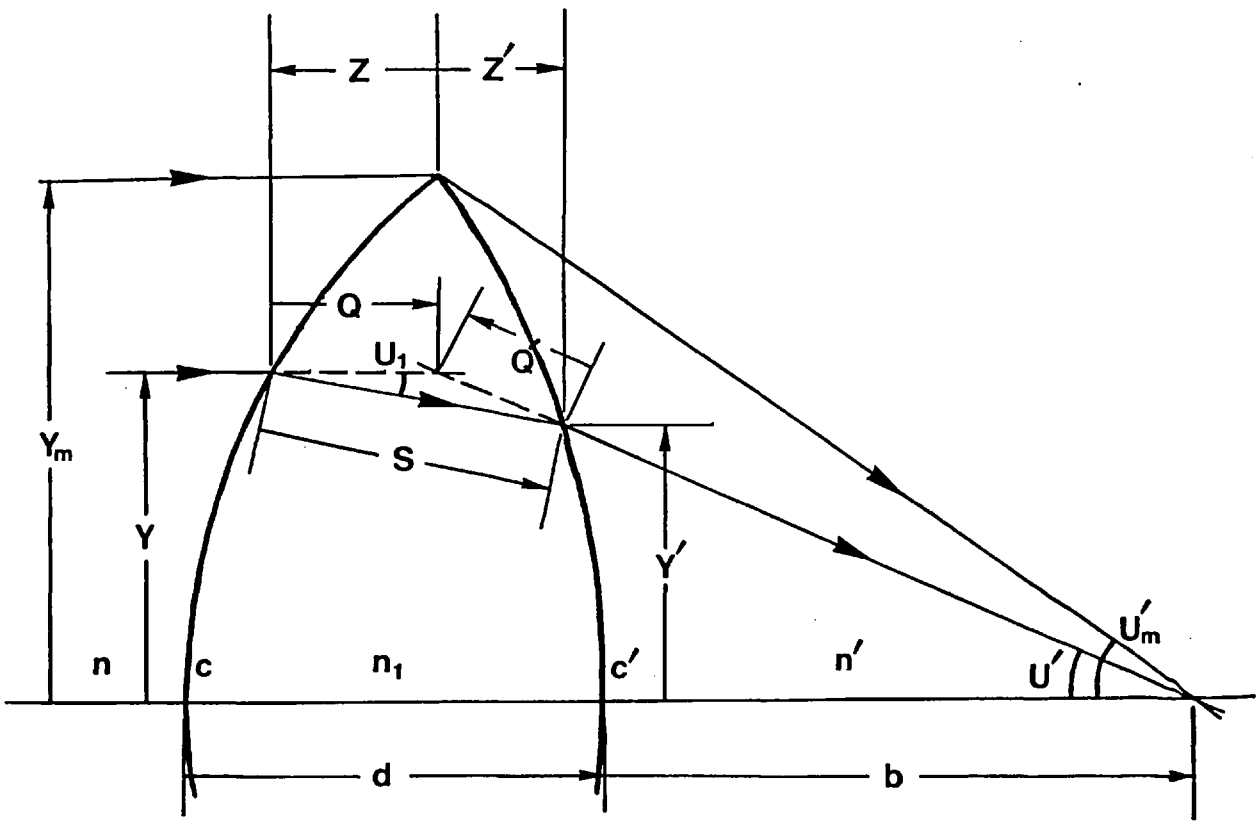


Fig. (3.1) The Bi-aspheric singlet.

given by,

$$d + b = -Z_a - \frac{Y_m}{\tan U'_m} \tag{3.1}$$

The optical path length along marginal and axial rays must be equal, so

$$n_1 d + n' b = -nZ_a - \frac{n' Y_m}{\sin U'_m} \tag{3.2}$$

Eliminating Z_a between these and solving for d,

$$d = \frac{Y_m}{(n_1-n)} \left[\frac{n}{\tan U'_m} - \frac{n'}{\sin U'_m} \right] + b \left[\frac{n-n'}{n_1-n} \right] \tag{3.3}$$

The paraxial curvatures may then be calculated from,

$$c = \frac{n_1(F-b)}{Fd(n_1-n)} \tag{3.4}$$

$$c' = \frac{n_1(b-F) + n' d}{bd(n'-n_1)} \tag{3.5}$$

Now refraction at the first and second surfaces give for the normal gradients,

$$\frac{dZ}{dY} = \frac{n_1 \sin U_1}{n - n_1 \cos U_1} \tag{3.6}$$

$$\frac{dZ'}{dY'} = - \frac{n_1 \sin U_1 - n' \sin U'}{n_1 \cos U_1 - n' \cos U'} \tag{3.7}$$

From the triangle with sides Q, Q', S (Fig. (3.1)),

$$\frac{Q}{\sin(U'-U_1)} = \frac{-Q'}{\sin U_1} = \frac{S}{\sin U'} \tag{3.8}$$

where Q, $S > 0$ and $Q' < 0$. The constant optical path condition gives,

$$\frac{n' Y_m}{\sin U'_m} + nZ = n' \left(\frac{Y}{\sin U'} - Q'\right) - n_1 S \tag{3.9}$$

Substituting for Q' and S in terms of Q using (3.8)

$$\frac{n' Y_m}{\sin U'_m} - \frac{n' Y}{\sin U} + nZ = \frac{Q}{\sin (U-U_1)} (n' \sin U_1 - n_1 \sin U') \tag{3.10}$$

Further, the geometrical distance $\frac{Y_m}{\tan U'_m} + Z$ for any ray is

$$\frac{Y_m}{\tan U'_m} + Z = \frac{Y}{\tan U'} - Q \tag{3.11}$$

so,

$$Q = \frac{Y}{\tan U'} - \frac{Y_m}{\tan U'_m} - Z \tag{3.12}$$

Rearranging equation (3.10) so that terms containing U_1 are on one side and terms containing Z and Y are on the other,

$$\frac{n' \sin U_1 - n_1 \sin U'}{\sin(U' - U_1)} = \frac{n' \left(\frac{Y_m}{\sin U'_m} - \frac{Y}{\sin U'}\right) + nZ}{Q} \tag{3.13}$$

We then define the function $h(Z,Y,U')$ as,

$$h(Z,Y,U') = \frac{n' \left(\frac{Y_m}{\sin U'_m} - \frac{Y}{\sin U'}\right) + nZ}{Q} \tag{3.14}$$

so that,

$$n' \sin U_1 - n_1 \sin U = h \sin(U' - U_1) \tag{3.15}$$

$$= h(\sin U' \cos U_1 - \cos U' \sin U_1) \tag{3.16}$$

or,

$$(n' + h\cos U')\sin U_1 - h\sin U' \cos U_1 - n_1 \sin U' = 0 \tag{3.17}$$

This equation is solved for U_1, from which dZ/dY and dZ'/dY' may be calculated using (3.6) and (3.7). dZ/dY can therefore be numerically integrated. It is not necessary to integrate dZ'/dY' because the coordinates can be found analytically.

$$Y' = Y + S \sin U_1 \tag{3.18}$$

$$\text{and} \quad Z' = Z + S \cos U_1 \tag{3.19}$$

or in terms of Q

$$Y' = Y + \frac{Q \sin U' \sin U_1}{\sin(U' - U_1)} \tag{3.20}$$

$$Z' = Z + \frac{Q \sin U' \cos U_1}{\sin(U' - U_1)} \tag{3.21}$$

Starting from the axis with initial conditions,

$$Z_a = \frac{-Y_m}{\tan U'_m} - d - b \tag{3.22}$$

$$Y_a = 0 \tag{3.23}$$

$$\left.\frac{dZ}{dY}\right|_a = 0 \tag{3.24}$$

and using a fourth-order Runge-Kutta method, the surfaces can be generated to any desired degree of accuracy (Appendix B).

The authors of the reference started their integration from the rim. The initial conditions here are $Z=0$, $Y=Y_m$ while dZ/dY must be found by using L'Hospital's rule on equation (3.14) which becomes indeterminate at the rim. By starting from the axis we avoid this. Even at the end of the integration it is not necessary since the rim point is not actually realized in practice.

To clarify the process a summary of the relevant equations in the correct sequence would be in order.

Summary of procedure

The specifications of the lens are to be as follows-

F	Focal length
b	Back focal length
Y_m	Maximum ray height
$U' = U'(Y)$	Principal surface
n, n_1, n'	Refractive indices

Calculate,

$U'_m = U'(Y_m)$, the maximum ray angle.

$$d = \frac{Y_m}{(n_1-n)}\left[\frac{n}{\tan U'_m} - \frac{n'}{\sin U'_m}\right] + b\left[\frac{n-n'}{n_1-n}\right]$$ the lens thickness

$$c = \frac{n_1(F-b)}{Fd(n_1-n)}$$

$$c' = \frac{n_1(b-F)+n'd}{bd(n'-n_1)}$$ the paraxial curvatures

Set the initial conditions for the axial points.

$$Z_a = \frac{-Y_m}{\tan U'_m} - d - b$$

$$Z'_a = d + Z_a$$

$$Y_a = Y'_a = 0$$

$$\left.\frac{dZ}{dY}\right|_a = \left.\frac{dZ'}{dY'}\right|_a = 0$$

Then integrate using a Runge-Kutta method,

$$U' = U'(Y)$$

$$Q = \frac{Y}{\tan U'} - \frac{Y_m}{\tan U'_m} - Z$$

where $\lim_{Y,U'\to 0}\left[\frac{Y}{\tan U'}\right] = -F$

$$h(Z,Y,U') = \frac{n'\left(\frac{Y_m}{\sin U'_m} - \frac{Y}{\sin U'}\right) + nZ}{Q}$$

Solve for U_1 from,

$$(n' + h\cos U')\sin U_1 - h\sin U'\cos U_1 - n_1\sin U' = 0$$

$$\frac{dZ}{dY} = \frac{n_1\sin U_1}{n - n_1\cos U_1}$$

When each new point on surface 1 is found, the corresponding point on the second surface is then given by,

$$Y' = Y + \frac{Q\sin U'\sin U_1}{\sin(U'-U_1)}$$

$$Z' = Z + \frac{Q \sin U' \cos U_1}{\sin(U' - U_1)}$$

$$\frac{dZ'}{dY'} = - \frac{n_1 \sin U_1 - n' \sin U'}{n_1 \cos U_1 - n' \cos U'}$$

The number of integration steps should be increased until the results remain constant to within the accuracy required. The program for performing the above calculation should also check at each integration step for total internal reflection, and the integration terminated if this occurs.

3.3. Design examples

As an example an F/1 lens was chosen as being typical of the aperture required. A beam diameter of 110mm was selected. The focal length is therefore also 110mm. To allow for mounting, Y_m was set to 58·0mm rather than 55·0mm. The glass was taken to be BK7 with an index of $n_1 = 1{\cdot}50621$ at 1·064μm. The surrounding media were given indices of unity; $n = n' = 1$.

We chose two cases for examination: when

$$\sin U' = -\frac{Y}{F} \tag{3.25}$$

and

$$\sin\left(\frac{U'}{2}\right) = -\frac{Y}{2F} \tag{3.26}$$

With F=110 and Y_m=58, the angles U'_m are -31·821° and -30·572° respectively. Equation (3.3) then gives the thickness 32·661mm and 31·314mm in each case.

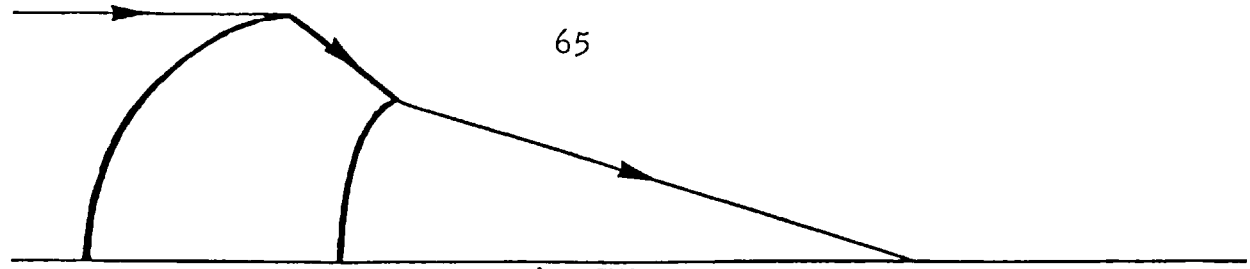
The number of integration steps was 116 (every 0·5mm). The only parameter not specified yet is the back focal length.

This was given 9 different values in each example and the resulting lens drawn. The lenses generated are shown in Figs. (3.2) and (3.3).

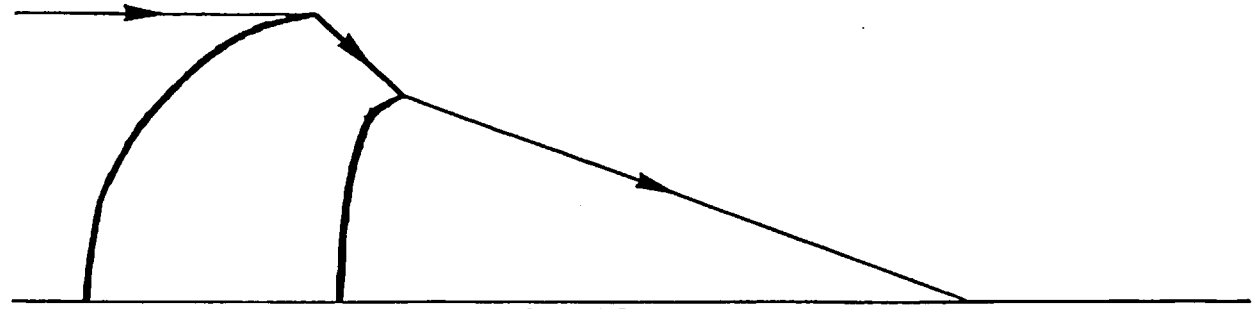
In each case it is found that only one of the lenses actually achieved the aperture originally aimed for. All the others became limited by total internal reflection at the second surface or grazing incidence on the first surface. This means that the lens also has the undesirable property of its transmission dropping down to zero at the marginal zone. A plot showing the approximate aperture achieved against the back focal length has been included in the figures. The one case which does achieve the aperture initially set, closely resemble those given in the reference.

These results also agree with those obtained by Head (13). He points out that if the back focal length is too small the aperture is limited by the front surface, but by the rear surface if it becomes too large. The maximum aperture is obtained when the two surfaces meet; this occurs for only one back focal length. Only this case is useful.

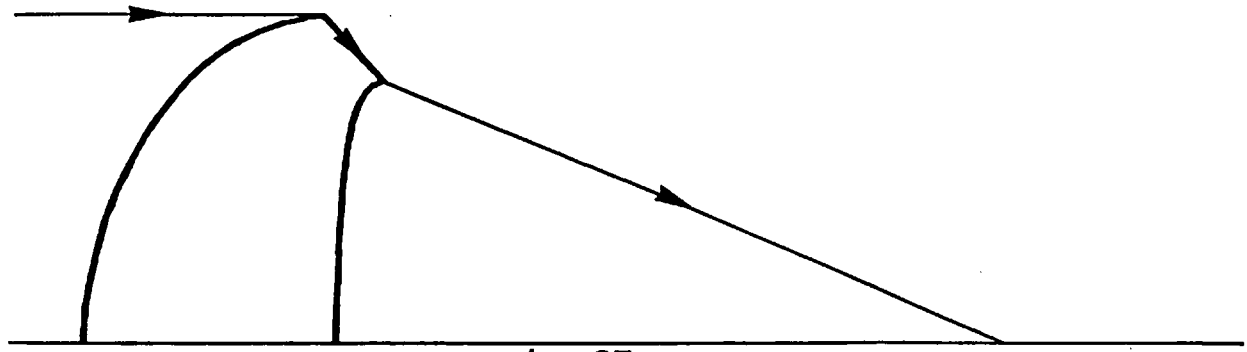
Luckily, this is also the lens with the most reasonably shaped surfaces. All the others would be a nightmare to fabricate and test even if they were suitable. Martin (14) also produced a paper concerned with high aperture aplanatic singlets. He used a method of generating the aspherics by series expansions rather than by integrating differential equations. Again, the results appear to agree with those in Fig. (3.2).



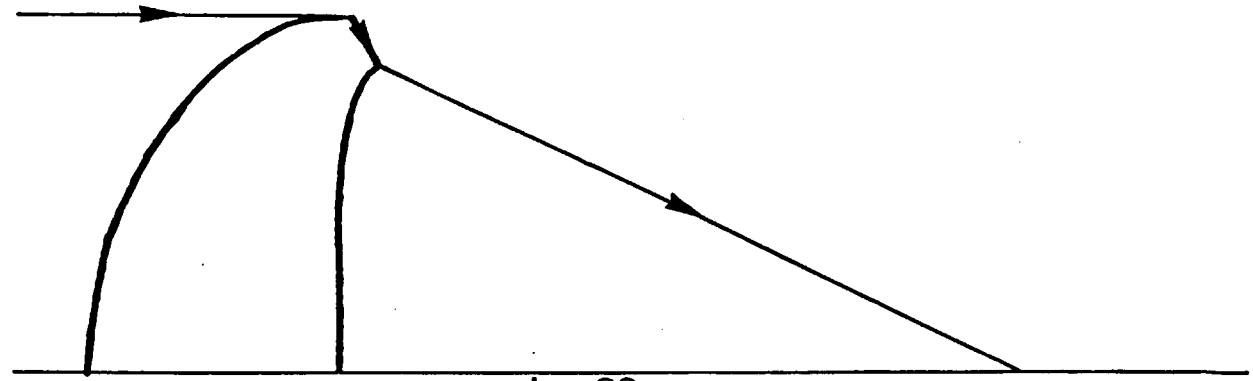
b = 75



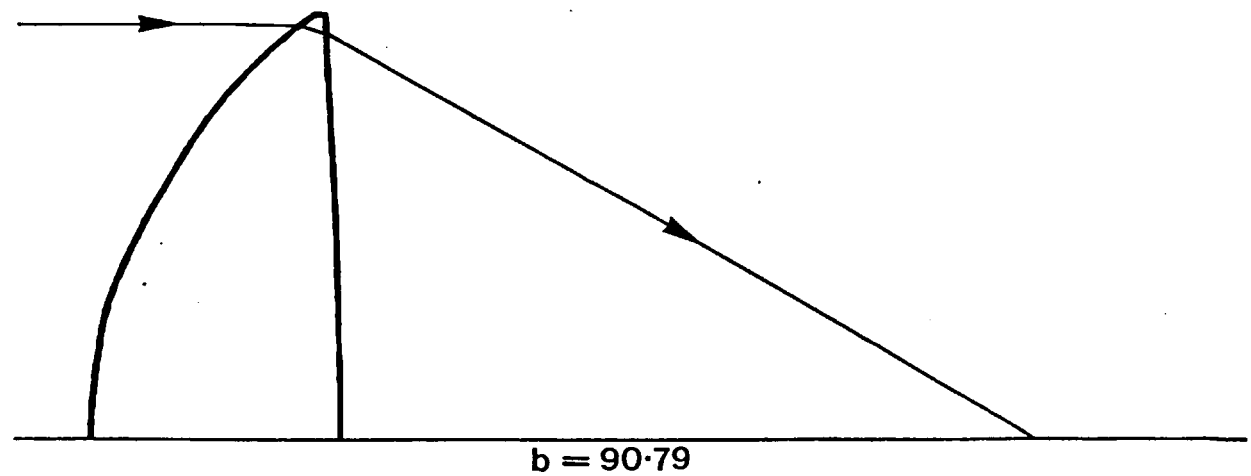
b = 82



b = 87

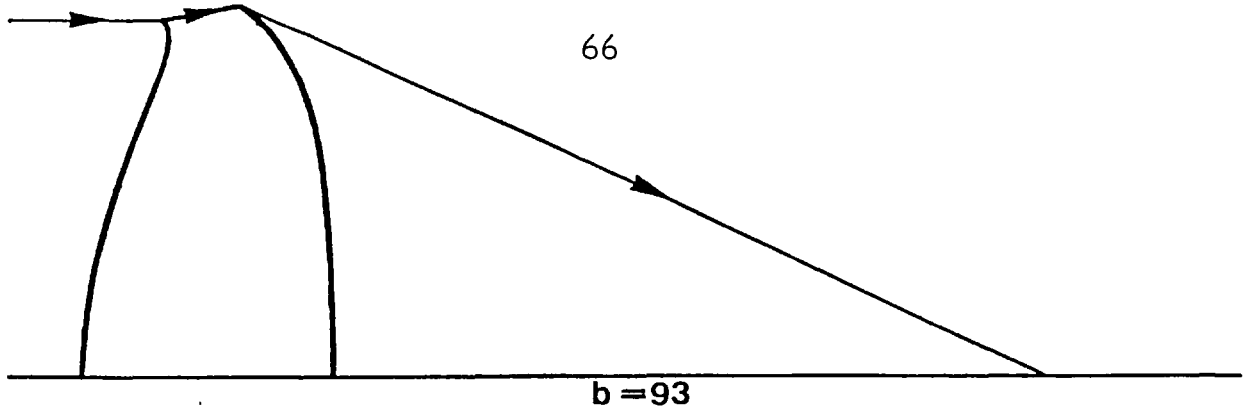


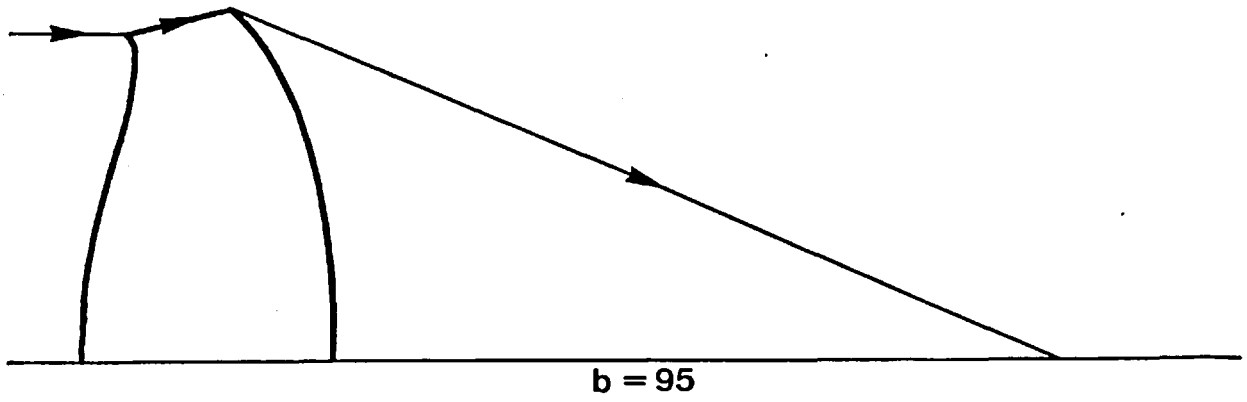
b = 89

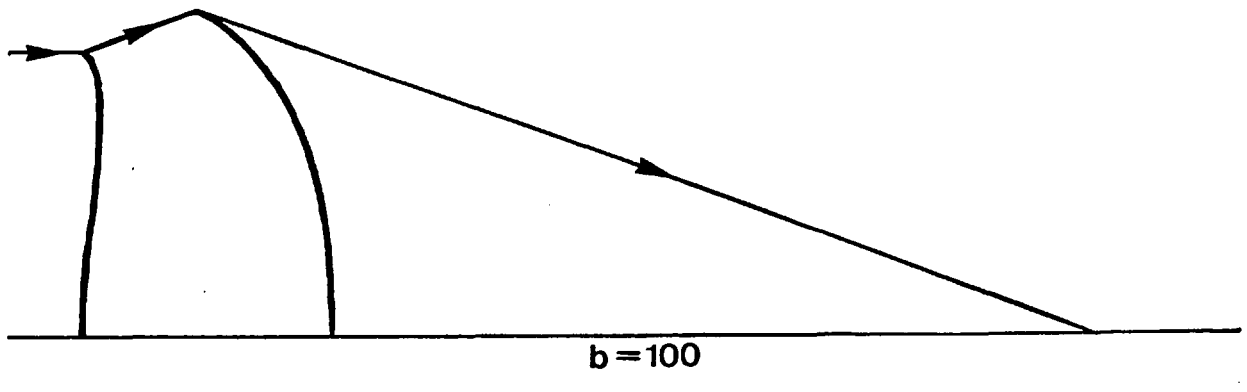


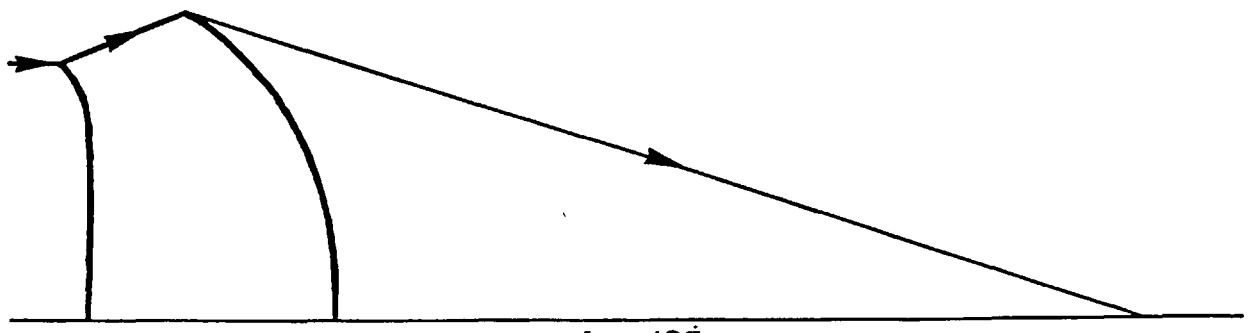
b = 90·79

Fig. (3.2)









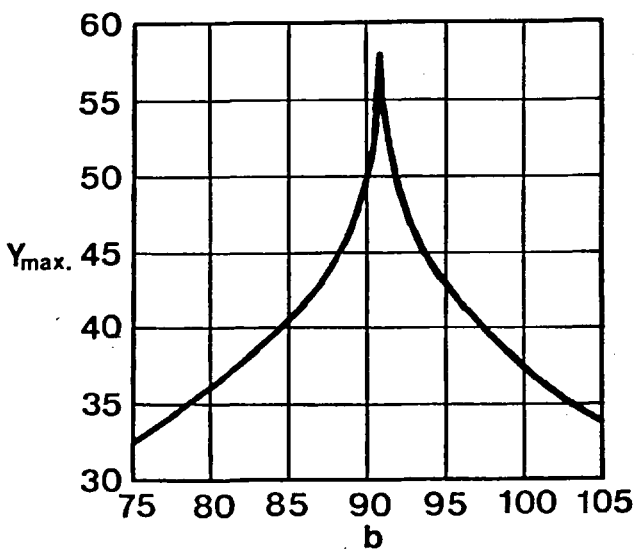
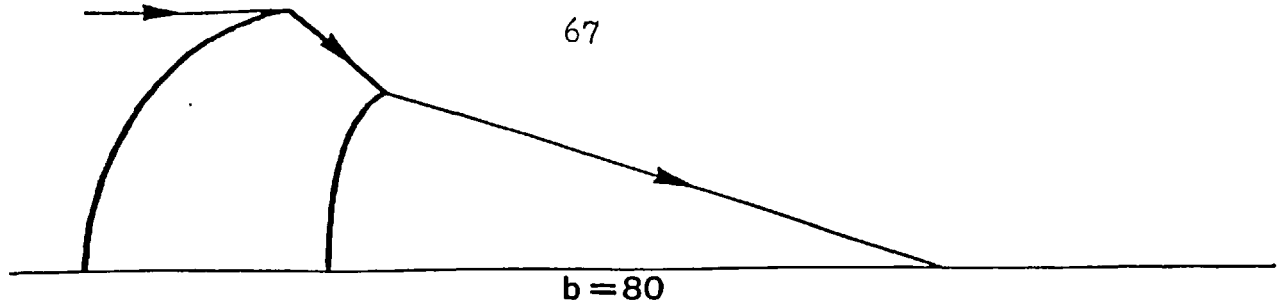
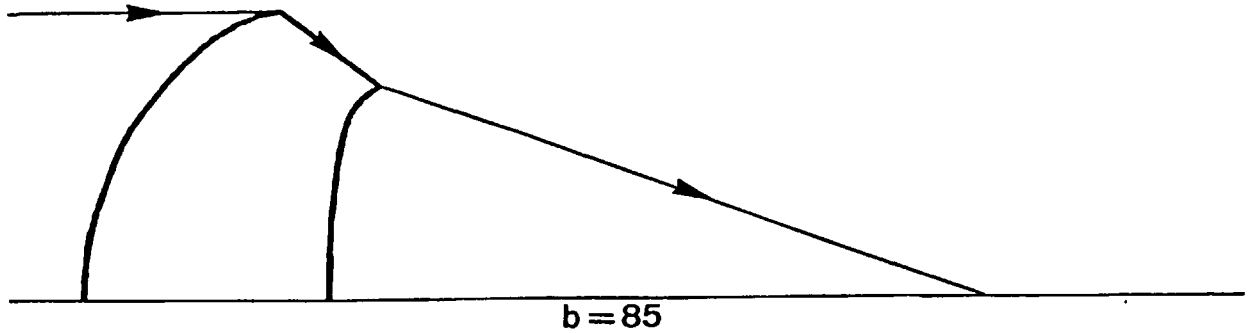


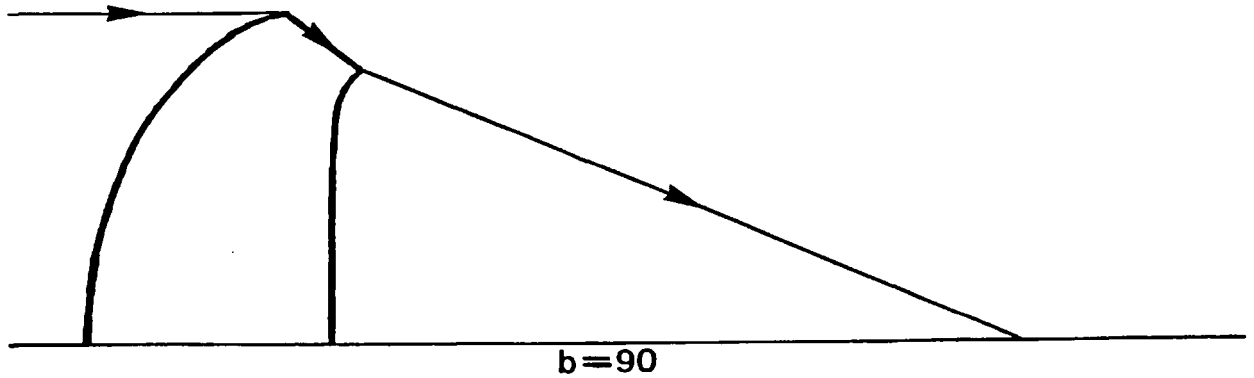
Fig. (3.2) Aplanatic singlets F=110mm, b=75 to 105mm and max. aperture is 58mm. Plot of $Y_{max.}$ against b is shown.



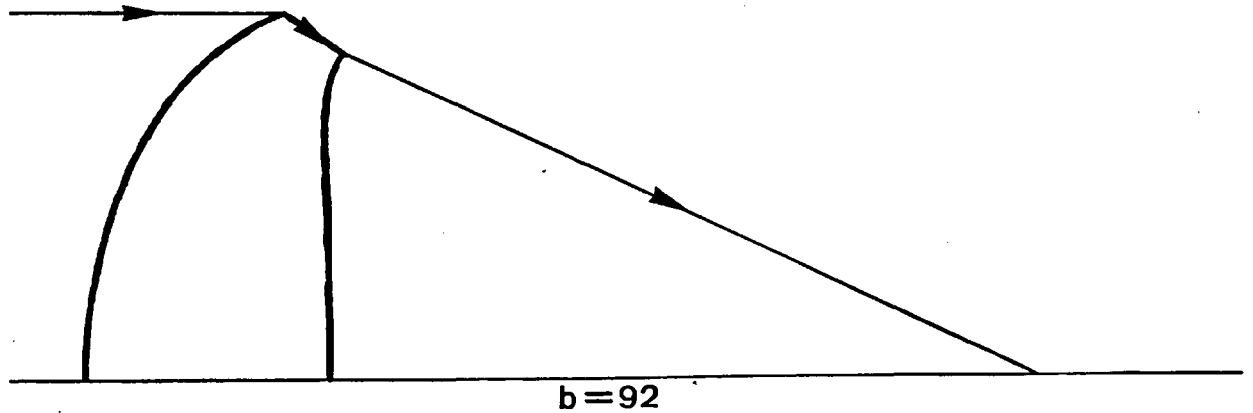
b = 80



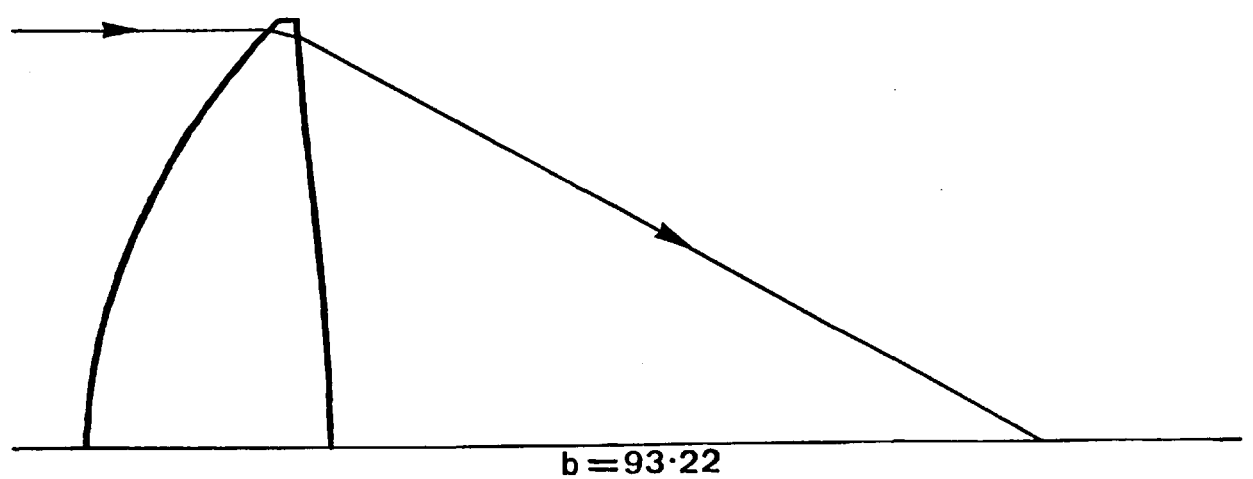
b = 85



b = 90

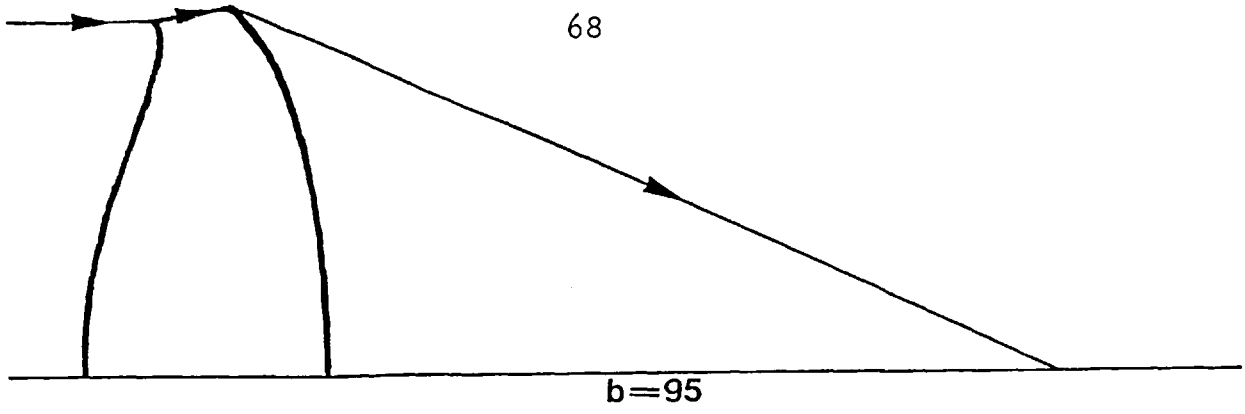


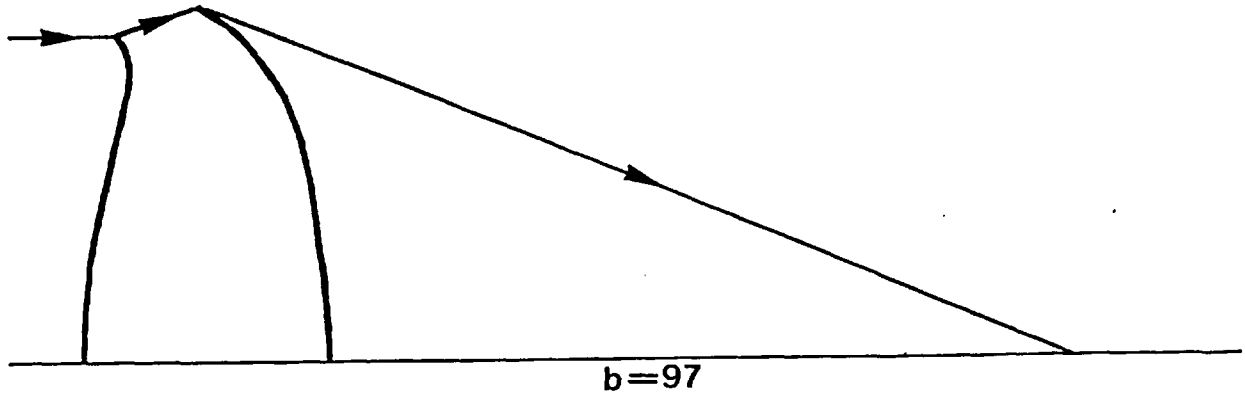
b = 92

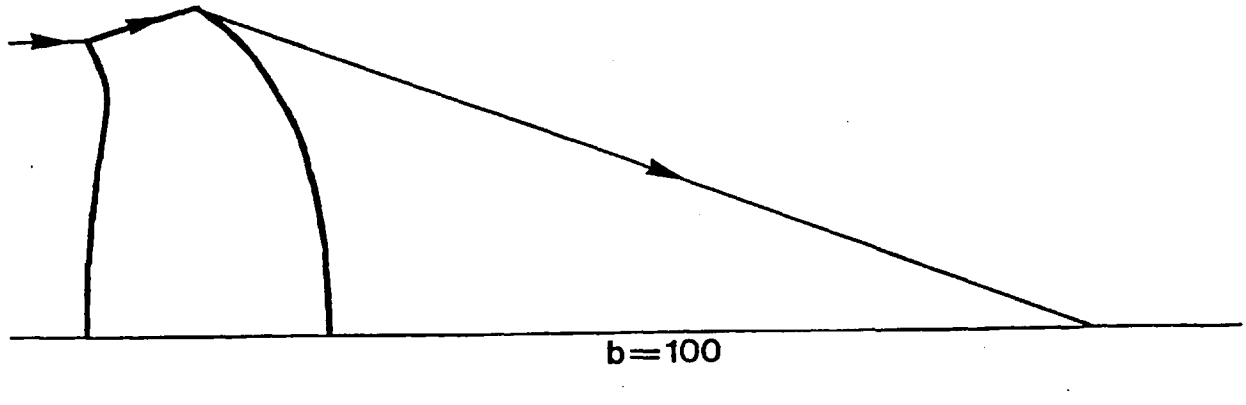


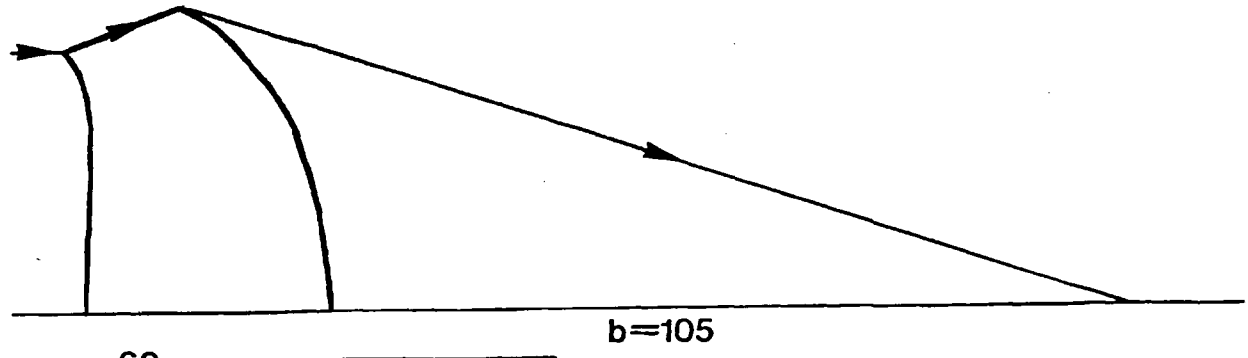
b = 93·22

Fig. (3.3)









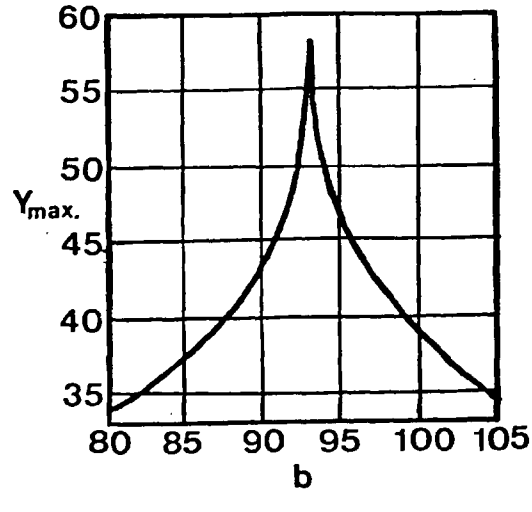


Fig. (3.3) Singlets that satisfy Herschel's condition. F=110mm, b=80 to 105mm and max. aperture is 58mm. Plot of $Y_{max.}$ against b is shown.

3.4. The doublet

In order to decide what use the extra parameters may be put to, it is worth remembering that the spherical aberration and coma are determined by the aspheric surfaces. They would normally be adjusted by altering the component shapes. Since this is now unnecessary we can therefore set the elements to be plano-aspheric without loss of generality. This will greatly simplify the calculations. The other aberrations such as astigmatism and field curvature are unaffected by bending. They depend on lens power. Therefore, their correction will not be possible with both elements positive and a small separation. In cases, where the lens is not aplanatic it is unlikely that their correction would be noticed anyway.

Fig. (3.4) shows the doublet and definition of the symbols. Again, in order to keep the path length inside the glass to a minimum, the elements' front and rear surfaces will be designed to meet at the marginal zone, defined by Y_m, U_{2m} and U'_m.

Refraction at the first surface gives,

$$\frac{dZ}{dY} = \frac{n_1 \sin U_1}{n - n_1 \cos U_1} \tag{3.27}$$

$$= \frac{n_1 \sin U_1}{n - n_1 \sqrt{1 - \sin^2 U_1}} \tag{3.28}$$

At the planar second surface,

$$n_1 \sin U_1 = n_2 \sin U_2 \tag{3.29}$$

Therefore,

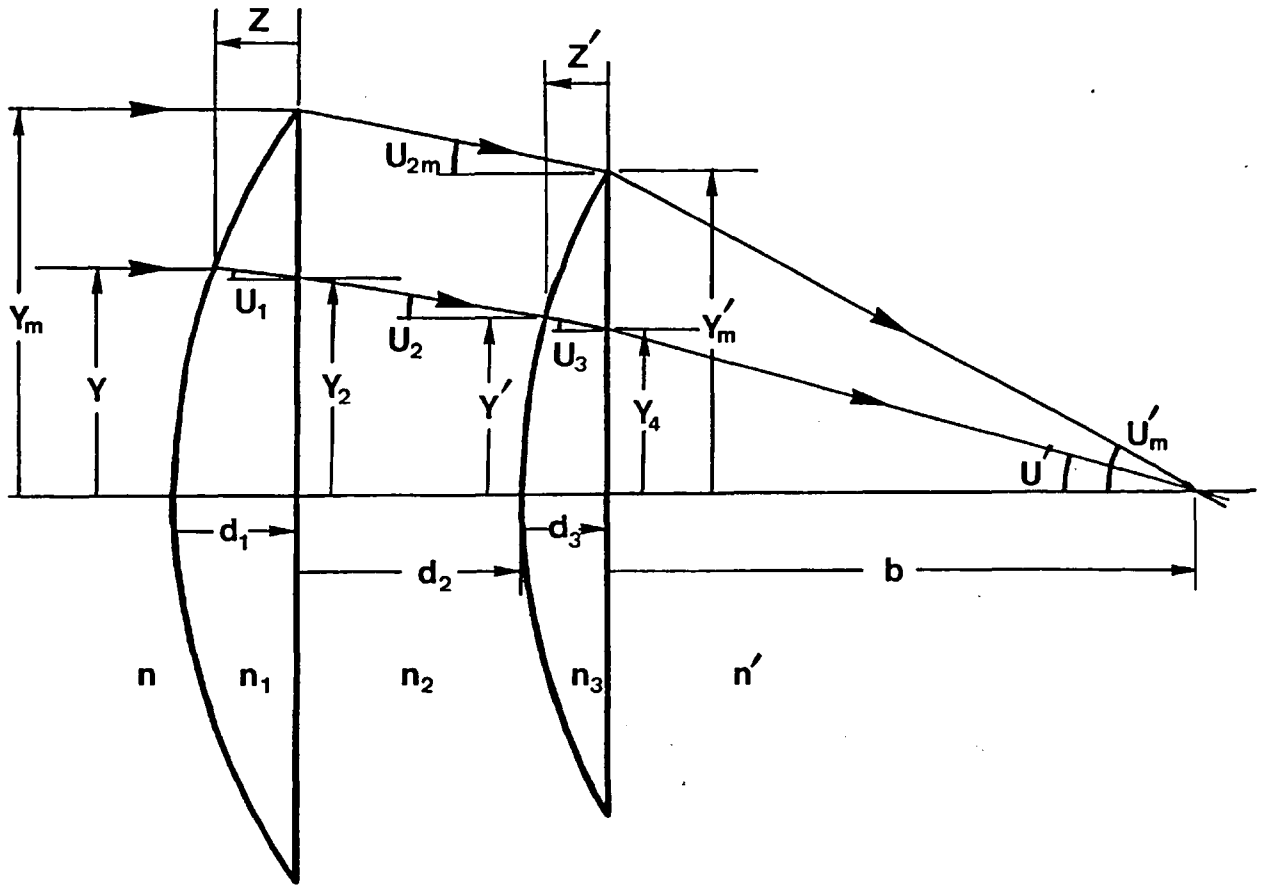


Fig. (3.4) The Plano-aspheric doublet.

$$\frac{dZ}{dY} = \frac{n_2 \sin U_2}{n - \sqrt{n_1^2 - n_2^2 \sin^2 U_2}} \tag{3.30}$$

Similarly, for refraction at the third surface,

$$\frac{dZ'}{dY'} = - \frac{n_2 \sin U_2 - n_3 \sin U_3}{n_2 \cos U_2 - n_3 \cos U_3} \tag{3.31}$$

and at the fourth,

$$n_3 \sin U_3 = n' \sin U' \tag{3.32}$$

Eliminating U_3,

$$\frac{dZ'}{dY'} = - \frac{n_2 \sin U_2 - n' \sin U'}{n_2 \cos U_2 - \sqrt{n_3^2 - n'^2 \sin^2 U'}} \tag{3.33}$$

The equations of transfer are easily seen to be,

$$Y_4 = -b \tan U' \tag{3.34}$$

$$Y' = Y_4 + Z' \tan U_3 \tag{3.35}$$

$$Y_2 = Y' - (d_2 + d_3 + Z') \tan U_2 \tag{3.36}$$

$$Y = Y_2 + Z \tan U_1 \tag{3.37}$$

Substituting one into the other to eliminate Y_2, Y' and Y_4,

$$Y = Z \tan U_1 - (d_2 + d_3 + Z') \tan U_2 + Z' \tan U_3 - b \tan U' \tag{3.38}$$

But,

$$\tan U_1 = \frac{\sin U_1}{\sqrt{1 - \sin^2 U_1}} \tag{3.39}$$

$$= \frac{n_2 \sin U_2}{\sqrt{n_1^2 - n_2^2 \sin^2 U_2}} \tag{3.40}$$

$$\text{and } \tan U_3 = \frac{n' \sin U'}{\sqrt{n_3^2 - n'^2 \sin^2 U'}} \tag{3.41}$$

Therefore,

$$\tan U_2 = \frac{\dfrac{Zn_2 \sin U_2}{\sqrt{n_1^2 - n_2^2 \sin^2 U_2}} + \dfrac{Z' n' \sin U'}{\sqrt{n_3^2 - n'^2 \sin^2 U'}} - b\tan U' - Y}{d_2 + d_3 + Z'} \qquad (3.42)$$

When Y, U', Z and Z' are known this expression is solved for U_2 the ray angle between the components. It can be turned into a quartic equation in $\sin U_2$ or $\tan U_2$ or solved numerically as it stands.

The value of d_2+d_3 is found from transfer equation (3.36),

$$Y_2 = Y' - (d_2 + d_3 + Z')\tan U_2 \qquad (3.43)$$

For the marginal zone,

$$Y_{2m} = Y'_m - (d_2 + d_3 + Z'_m)\tan U_{2m} \qquad (3.44)$$

But,

$$Z'_m = 0, \quad Y_{2m} = Y_m \text{ and } Y'_m = Y_{4m} = -b\tan U'_m$$

Therefore,

$$d_2 + d_3 = \frac{-b\tan U'_m - Y_m}{\tan U_{2m}} \qquad (3.45)$$

It is therefore necessary to specify U_{2m}, the marginal ray angle between the elements.

Unlike the case of the singlet, both dZ/dY and dZ'/dY' have to be integrated simultaneously. To do this it is necessary to integrate with respect to the same independent variable, chosen here to be Y. We therefore have to find dZ'/dY.

Now,

$$\frac{dZ'}{dY} = \frac{dZ'}{dY'}\frac{dY'}{dY} = \frac{dZ'}{dY'}\frac{dY'}{dU'}\frac{dU'}{dY} \qquad (3.46)$$

$\frac{dU'}{dY}$ is found from the known relation $U' = U'(Y)$.

From the transfer equations (3.34) and (3.35),

$$Y' = -b\tan U' + Z'\tan U_3 \tag{3.47}$$

$$\frac{dY'}{dU'} = -b\sec^2 U' + \frac{dZ'}{dU'}\tan U_3 + Z'\frac{d\tan U_3}{dU'} \tag{3.48}$$

But,

$$\frac{dZ'}{dU'} = \frac{dZ'}{dY'}\frac{dY'}{dU'} \tag{3.49}$$

$$\frac{dY'}{dU'} = -b\sec^2 U' + \frac{dZ'}{dY'}\frac{dY'}{dU'}\tan U_3 + Z'\frac{d\tan U_3}{dU'} \tag{3.50}$$

Solving for $\frac{dY'}{dU'}$

$$\frac{dY'}{dU'} = \left[\frac{Z'\dfrac{d\tan U_3}{dU'} - b\sec^2 U'}{1 - \dfrac{dZ'}{dY'}\tan U_3}\right] \tag{3.51}$$

Therefore from (3.46)

$$\frac{dZ'}{dY} = \frac{dZ'}{dY'}\frac{dU'}{dY}\left[\frac{Z'\dfrac{d\tan U_3}{dU'} - b\sec^2 U'}{1 - \dfrac{dZ'}{dY'}\tan U_3}\right] \tag{3.52}$$

Finally since,

$$\tan U_3 = \frac{n'\sin U'}{\sqrt{n_3^2 - n'^2\sin^2 U'}} \tag{3.53}$$

so,

$$\frac{d\tan U_3}{dU'} = \frac{n' n_3^2 \cos U'}{(n_3^2 - n'^2 \sin^2 U')^{3/2}} \qquad (3.54)$$

dZ/dY and dZ′/dY can be integrated using a simultaneous Runge-Kutta method. (Appendix B). By decreasing the step length any accuracy can be achieved. The integration is started at the marginal zone rather than from the axis. In the doublet's case we do not know the element thicknesses to start with, so we cannot find Z_a and Z'_a. On the otherhand, the marginal ray angle U_{2m} is given and enables $\left.\frac{dZ}{dY}\right|_m$ and $\left.\frac{dZ'}{dY'}\right|_m$ to be calculated.

Again a summary would be appropriate.

Summary of procedure

The lens specifications are,

b	Back focal length
Y_m	Maximum ray height
U'_m	Maximum ray angle
U_{2m}	Maximum ray angle between the elements
$U' = U'(Y)$	Principal surface
n, n_1, n_2, n_3, n'	Refractive indices

Given these, the initial conditions for the integration are,

$$Z_m = Z'_m = 0$$

$$Y'_m = -b\tan U'_m$$

$$\left.\frac{dZ}{dY}\right|_m = \frac{n_2 \sin U_{2m}}{n - \sqrt{n_1^2 - n_2^2 \sin^2 U_{2m}}}$$

$$\left.\frac{dZ'}{dY'}\right|_m = -\frac{n_2 \sin U_{2m} - n' \sin U'_m}{n_2 \cos U_{2m} - \sqrt{n_3^2 - n'^2 \sin^2 U'_m}}$$

$$d_2 + d_3 = \frac{-Y_m - b\tan U'_m}{\tan U_{2m}}$$

At the beginning of each integration step Y, Z, Z' are known.

$U' = U'(Y)$ gives U' and $\frac{dU'}{dY}$ for the next Y value.

Solve for U_2 using,

$$\tan U_2 = \frac{\frac{Zn_2 \sin U_2}{\sqrt{n_1^2 - n_2^2 \sin^2 U_2}} + \frac{Z' n' \sin U'}{\sqrt{n_3^2 - n'^2 \sin^2 U'}} - b\tan U' - Y}{d_2 + d_3 + Z'}$$

Then calculate,

$$\frac{dZ'}{dY'} = - \frac{n_2 \sin U_2 - n' \sin U'}{n_2 \cos U_2 - \sqrt{n_3^2 - n'^2 \sin^2 U'}}$$

$$\frac{d\tan U_3}{dU'} = \frac{n' n_3^2 \cos U'}{(n_3^2 - n'^2 \sin^2 U')^{3/2}}$$

$$\tan U_3 = \frac{n' \sin U'}{\sqrt{n_3^2 - n'^2 \sin^2 U'}}$$

The derivatives required for the integration are,

$$\frac{dZ}{dY} = \frac{n_2 \sin U_2}{n - \sqrt{n_1^2 - n_2^2 \sin^2 U_2}}$$

$$\frac{dZ'}{dY} = \left(\frac{dZ'}{dY'}\right)\left(\frac{dU'}{dY}\right)\left[\frac{Z' \frac{d\tan U_3}{dU'} - b\sec^2 U'}{1 - \left(\frac{dZ'}{dY'}\right)\tan U_3}\right]$$

The axial values of Z and Z' are equal to $-d_1$ and $-d_3$ respectively. The component spacing d_2 is then found from the value of $d_2 + d_3$. Obviously these must not turn out to be negative. Changing U_{2m} can help avoid this. Although it is clear that $d_2 + d_3$ will increase if U_{2m} decreases the thickness d_3 may also increase, since more of the ray bending is left to the second element. Thus how d_2 behaves is not obvious.

Once the spacings are known the paraxial curvatures may be calculated along with any other paraxial data.

The choice of back focal length should be made carefully. Since it is desirable for the beam to converge on passing through the first element, then,

$$Y'_m < Y_m \tag{3.55}$$

but

$$Y'_m = -b\tan U'_m \tag{3.56}$$

$$b < -\frac{Y_m}{\tan U'_m} \tag{3.57}$$

i.e. there is an upper limit for reasonable solutions.

3.5. Design examples

For comparison, the same basic specifications were used to design the doublets as for the singlet; namely, $Y_m = 58$ and $F = 110$. We will give an example for each of the principal surfaces mentioned in section 1. Where suitable solutions resulted U_{2m} was set at -15^0, about half the final ray angle for the examples. The amount of ray bending is therefore approximately shared between the components.

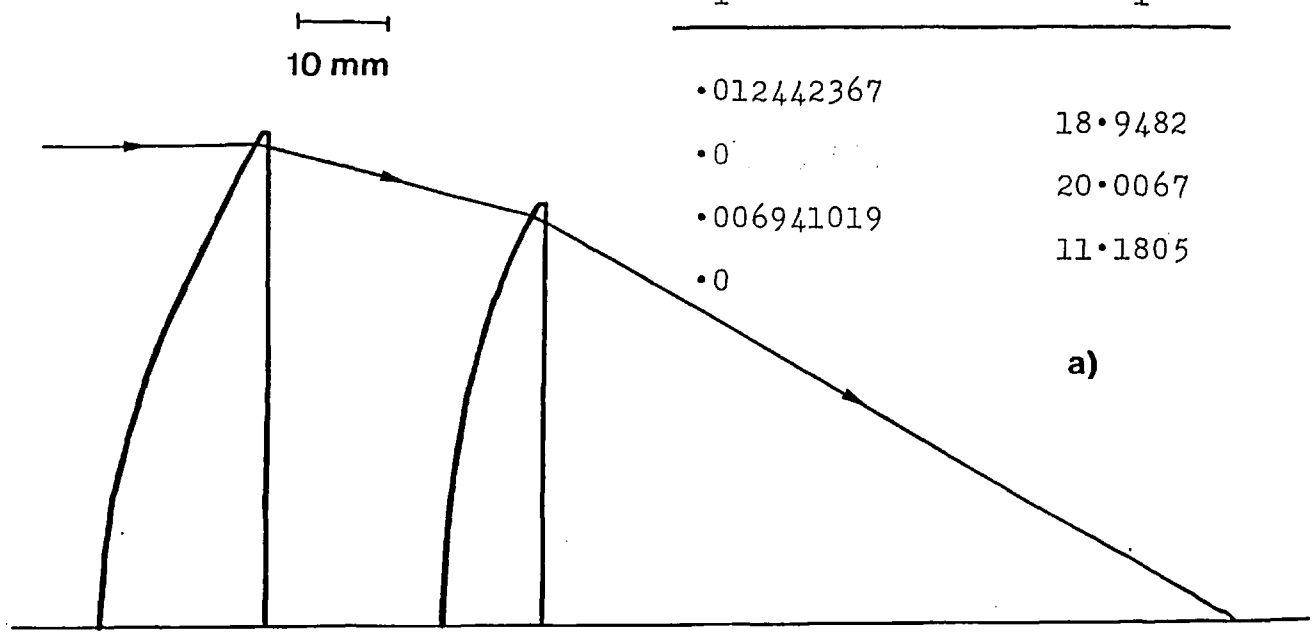
Again 116 integration steps were taken and the aspheric surface coordinates interpolated using a polynomial containing ten terms;

$$Z = A_2Y^2 + A_4Y^4 + \ldots + A_{20}Y^{20} \qquad (3.58)$$

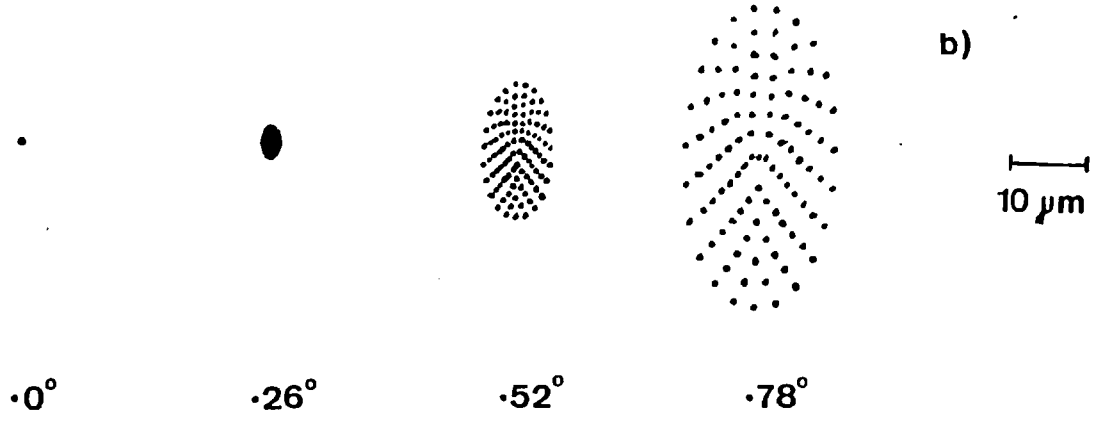
where Z is now measured from the pole rather than the rim. The method using orthogonal polynomials was found to give much better results than a least square method, although even the former tended to give poor results for highly aspheric and oddly shaped surfaces.

Each lens was then raytraced at four field angles to test off-axis performance and to give an idea of image sizes corresponding to the OSC curves of Fig.(2.2). These particular field angles will be used throughout this thesis; they are $0{\cdot}78^0$, $0{\cdot}52^0$, $0{\cdot}26^0$ and $0{\cdot}0^0$ (on axis) and correspond to an ideal image 1·5, 1·0, 0·5 and 0·0mm from the axis when the focal length is 110·0mm.

Figs. (3.5) to (3.8) shows the lenses and data. The aplanat and lens for dealing with a Gaussian beam have larger element spacings than would normally be set. In the case of the aplanat the back focal length is 80mm. If it is increased to 90mm the resulting lens has the appearance in Fig (3.9). The spacing has been reduced but the first element is four times thicker than the second and has more than five times the power. The aspheric on the front surface has an inflection point at a ray height of about 49mm. This and its highly aspheric nature were no doubt responsible for the poor polynomial fit. Raytracing was correspondingly inaccurate. Consequently, it was decided to try and achieve

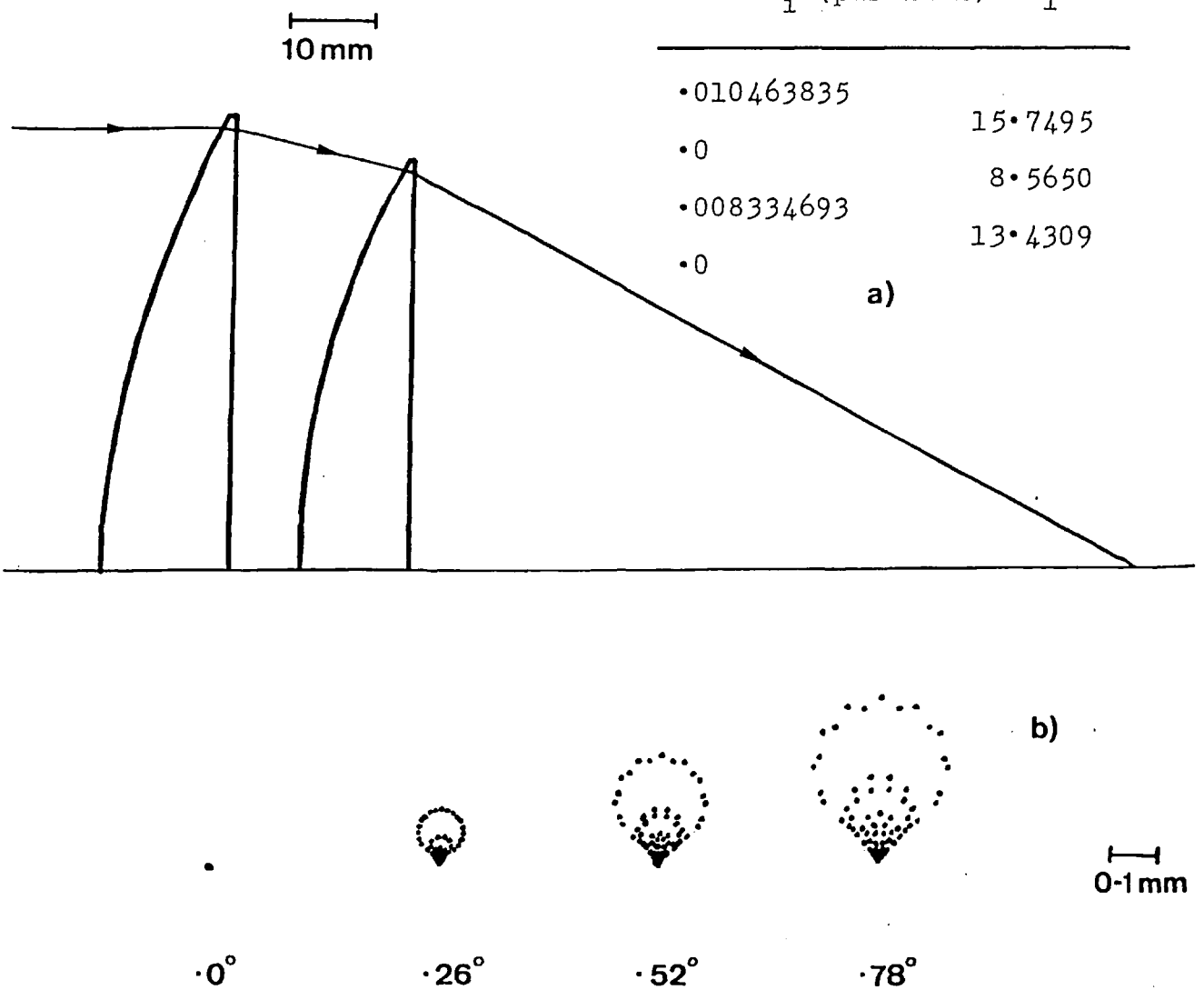


c_i (paraxial)	d_i
·012442367	
	18·9482
·0	
	20·0067
·006941019	
	11·1805
·0	



j	A_{j1}	A_{j3}
2	·6221183157 x10^{-2}	·3470509435 x10^{-2}
4	-·5600952628 x10^{-7}	·2516233657 x10^{-6}
6	-·2760091082 x10^{-10}	·6526774663 x10^{-10}
8	-·2894119491 x10^{-14}	·4399719600 x10^{-14}
10	-·7238212073 x10^{-19}	-·2661538753 x10^{-18}
12	·7390036168 x10^{-22}	-·7298539097 x10^{-22}
14	-·9675040484 x10^{-26}	·4653036222 x10^{-26}
16	·6179676185 x10^{-29}	-·3267666410 x10^{-29}
18	-·1241398128 x10^{-32}	·8787945436 x10^{-33}
20	·7402372468 x10^{-37}	-·5480356806 x10^{-37}

Fig. (3.5) a.) F/1 aplanat. F=110mm, b=80mm, U_{2m}=-15°, U'_m=-30°.
b.) Spot diagrams.

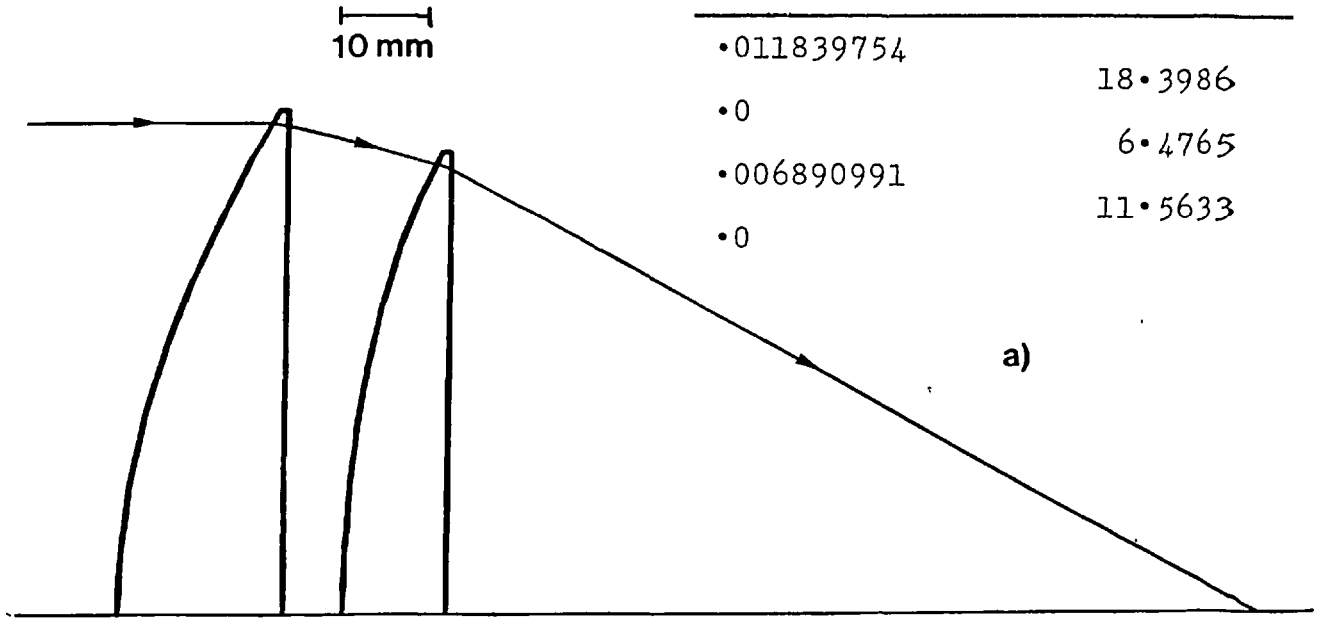


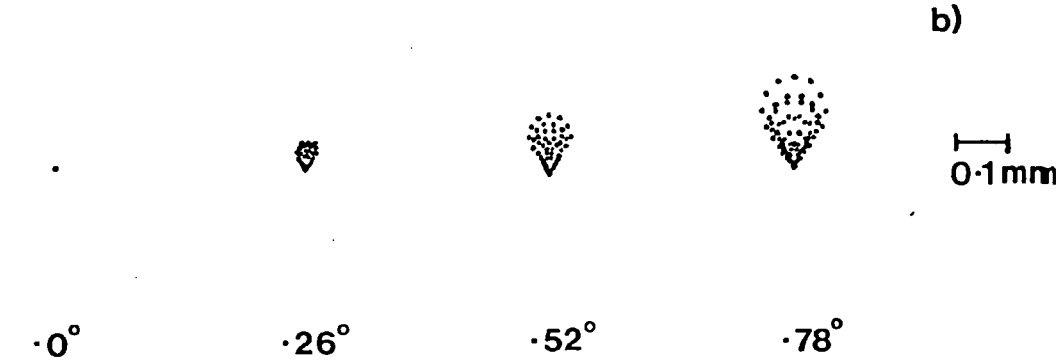
Field angle

j	A_{j1}	A_{j3}
2	•5231911636 x10^{-2}	•4167389920 x10^{-2}
4	-•1854807573 x10^{-6}	•3205532709 x10^{-6}
6	-•1801160230 x10^{-10}	•2817685410 x10^{-10}
8	•5049255122 x10^{-14}	-•1912816028 x10^{-13}
10	-•8022015978 x10^{-19}	•1753253453 x10^{-16}
12	•8239466469 x10^{-21}	-•1882046619 x10^{-19}
14	-•4310704778 x10^{-24}	•1135628562 x10^{-22}
16	•1239893202 x10^{-27}	-•4184973612 x10^{-26}
18	-•1946872149 x10^{-31}	•8558432146 x10^{-30}
20	•1455215024 x10^{-35}	-•7689402137 x10^{-34}

Fig. (3.6) a.) F/1 lens that illuminates a section of a spherical target uniformly. F=110mm, b=90mm, U_{2m}=-15°, U'_m=-28•675°, ϕ_m=-90°.
b.) Spot diagrams.

c_i (paraxial)	d_i
•011839754	
	18•3986
•0	
	6•4765
•006890991	
	11•5633
•0	

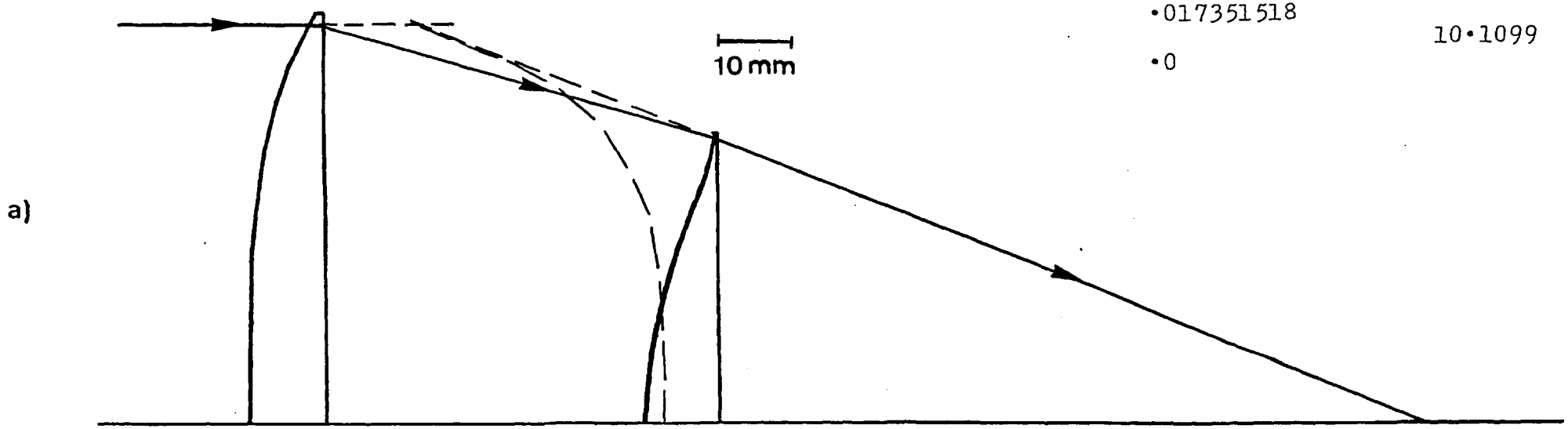




j	A_{j1}	A_{j3}
2	•5919876614 x10^{-2}	•3445495390 x10^{-2}
4	-•2309504349 x10^{-7}	•8382455798 x10^{-7}
6	-•2414238337 x10^{-10}	•3915853203 x10^{-10}
8	-•3098640885 x10^{-14}	•4951931482 x10^{-14}
10	-•2242570944 x10^{-18}	•9737337962 x10^{-19}
12	•1331441084 x10^{-21}	-•1123269835 x10^{-21}
14	-•3739435024 x10^{-25}	•6127516970 x10^{-26}
16	•1818704809 x10^{-28}	-•9615794058 x10^{-29}
18	-•3469935145 x10^{-32}	•2612569161 x10^{-32}
20	•2181759951 x10^{-36}	-•1910803788 x10^{-36}

Fig. (3.7) a.)F/1 lens that maps equal beam areas into equal solid angles: $\sin(U'/2)=-Y/2F$, F=110mm, b=90mm, $U_{2m}=-15^0$, $U'_m=-28•955^0$.
b.) Spot diagrams.

c_i (paraxial)	d_i
·001125880	10·7357
·0	45·3163
·017351518	10·1099
·0	



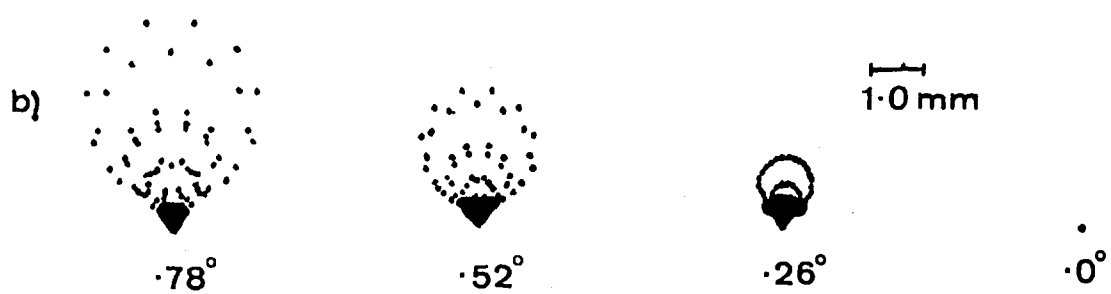


Fig. (3.8) a.) F/1 lens that transforms a Gaussian beam profile to a uniform illumination over the target. F=110mm, b=100mm, U_{2m}=-18°, U'_m=-21·356°, ϕ_m=-90°, g=55mm.

b.) Spot diagrams

j	A_{j1}	A_{j3}
2	·5629528410 x10^{-3}	·8676098550 x10^{-2}
4	·1041805963 x10^{-5}	-·8853995843 x10^{-6}
6	-·1950911062 x10^{-9}	-·3017146297 x10^{-10}
8	·7342098785 x10^{-13}	-·5398454138 x10^{-12}
10	-·2448328193 x10^{-16}	·1231841093 x10^{-14}
12	·7537724209 x10^{-20}	-·2063705554 x10^{-17}
14	-·1828867161 x10^{-23}	·2090054920 x10^{-20}
16	·3127961754 x10^{-27}	-·1289713484 x10^{-23}
18	-·3261034274 x10^{-31}	·4415698632 x10^{-27}
20	·1539187407 x10^{-35}	-·6541356782 x10^{-31}

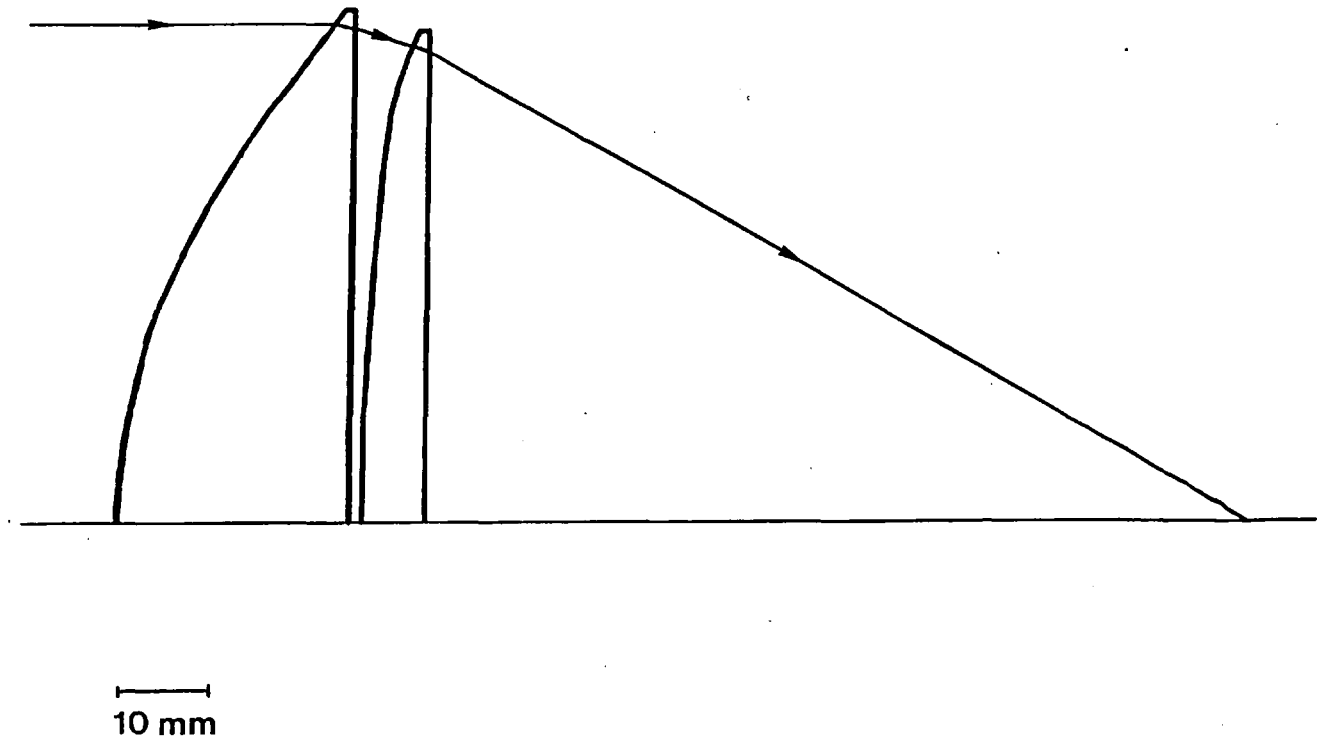


Fig. (3.9) F/1 aplanat.Same specifications as in fig.(3.5) except that b=90mm.

a more reasonable design.

If the lens of Fig. (3.8) has its spacing reduced any further, the front aspheric begins to take on a strangely shaped surface. The outer zones remain positively curved while the central region becomes negatively curved. The surface becomes "doughnut" shaped. Rays passing through the center of the lens thus become divergent in the air gap. This leaves the second element to do all the focussing. Its power and asphericity increase dramatically and the inflection point becomes very much more pronounced.

By referring to the diagram of the lens we can understand how the Gaussian profile is transformed to a uniform one over the target. The incident beam has a higher energy flux in the center than at the edge. Central rays pass through the first element virtually undeviated while the outer rays are strongly refracted. The effect of this is to increase the energy flux at the outer zones while leaving the central flux unchanged. On passing through the second component the central rays are those refracted.

By decreasing the element spacing further it is evident that marginal rays are not given enough space to converge sufficiently. The energy flux here is therefore not raised so much. To compensate for this the central flux is decreased by diverging the central pencils. Thus the first aspheric takes on the doughnut shape.

The principal surface for such a lens is therefore actually curved backwards (as illustrated in Fig.(3.8).) Consequently the coma is enormous, as shown by the spot diagrams. In fact, the blur circle diameter is much larger

than the ideal image displacement from the axis! Very accurate target alignment would be necessary for such a lens.

In complete contrast the aplanat is very well corrected. The other two cases are drawn to the same scale for comparison (Fig. (3.6) and (3.7)). The sudden increase in image size at the marginal region shown in Fig.(3.6) matches the more rapid rise in the corresponding curve of Fig.(2.2). The amount of coma actually present is still very large.

3.6. Conclusions

In order to allow a reasonable tolerance when aligning the optics with the target, the coma must not be allowed to become too large. An aplanat would therefore seem the ideal choice from this point of view. But as shown already, the aplanat can only provide uniform illumination on the target if the incident beam is also approximately uniform. On the other hand, a lens designed to provide uniformity from a Gaussian beam or any profile which drops away towards the edge, requires a principal surface that is considerably flatter than the aplanat's or even curved backwards. Consequently, such a lens suffers from large amounts of coma and may be totally useless because of this.

It appears therefore, that the incident beam should be smoothed out or "rectified" prior to focussing. The most efficient way of doing this was given by Frieden (15) and by Rhodes and Shealy (16). Two aspheric lenses are used to transform a collimated beam with a non-uniform profile into another of uniform intensity. Such systems are of the same

type as those just discussed and may be designed by similar methods. For afocal systems, however, a principal surface does not have any meaning.

CHAPTER 4

DESIGNS CONTAINING A SINGLE ASPHERIC (I)

4.1. Introduction

In the examples discussed so far it was seen that two aspheric surfaces are necessary to provide the desired target illumination and satisfy the constant optical path condition. The cost of producing such specialized optics, of which only a few copies are needed will be considerable. The main factor contributing to the expense is the manufacture of the aspheric surfaces. Involved in their production is the testing. No doubt in the initial stages of figuring each aspheric will be tested individually. Such a test could either be a direct measure of the surface shape or an interferometric test to find the aberrations present. Finally, as a check, both elements could be tested together. It is only when tested as a pair and at the laser wavelength, however, that a null test can be performed, i.e. one in which fringes are absent from an interferogram indicating the absence of any aberrations. The extra equipment required to carry out an infrared testing procedure complicates matters and adds to the cost of the whole process.

The work and cost can be somewhat reduced by use of only one aspheric surface. The system can then be designed to satisfy either one of the two conditions mentioned above. If the constant optical path condition is satisfied there is no real guarantee as to what the target illumination will turn out to be, but this can be controlled to some extent by bending the various elements. Alternatively, the aspheric can

be used to achieve the desired target illumination. The image will then be aberrated of course; with either of these choices we are still left with a null test at a non-visible wavelength or a non-null test to perform.

The reduction in work can be taken a stage further by demanding that the aspheric should be capable of being null tested at a convenient visible wavelength. This means that the system will not be corrected for the laser wavelength. Therefore, to make the lens useful a recorrection in the infrared is necessary.

A paper by Nicholas, Pataky and Welford (17) has suggested several methods of recorrecting the lens. Among them include regrinding one of the spherical surfaces after figuring the aspheric and the use of auxiliary systems, either for the testing or recorrecting stage.

The final method adopted was that of the wavelength tunable lens. Here a doublet can be null tested in the visible and the element spacing adjusted to recorrect at the laser wavelength.

Recall equation (2.68), the condition for uniform illumination of the target,

$$\frac{T(\rho)}{R_T} = -2\rho\sin(\tfrac{\phi_m}{2}) \sqrt{1 - \rho^2\sin^2(\tfrac{\phi_m}{2})} - 2(\rho^2\sin(\tfrac{\phi_m}{2}) + \frac{F}{Y_m})\sin(\tfrac{\phi_m}{2})\tan U' \quad (4.1)$$

It seems reasonable that by adjusting the component spacing, the aberrations, $T(\rho)$, and the principal surface $\tan U'$, can be made to "approximately satisfy" this equation. In other words, by adjusting the spacing we should not only be able to recorrect the aberrations but also improve the

illumination uniformity for any R_T and Φ_m. The spacing required being different in each case. In this and the next chapter we shall look at such systems.

The main problem is to generate the single aspheric surface required to correct spherical aberration either at the laser or visible wavelengths. The technique used here is also ideal for designing the corrector plates mentioned by Thomas.

4.2. Correcting spherical aberration by the use of an aspheric surface.

The calculation of a surface shape that achieves exact axial stigmatism for an optical system is generally easier if the surface is the first or last. When the final surface is to be aspheric, for example, the usual method involves tracing a large number of finite rays from the object point to the penultimate surface, and for each evaluating the optical path length up to this surface. The aspheric surface coordinates (Y,Z) and gradient $\frac{dZ}{dY}$ can then be determined from the constant optical path between known conjugate points and using Snell's law.

If the first surface is to be aspherized the raytracing is carried back through the system, starting from the image point.

By this point-characteristic method a set of coordinates and gradients are generated which are usually then interpolated using a sphere plus polynomial expression,

$$Z = \frac{cY^2}{1 + \sqrt{1 - c^2Y^2}} + a_4Y^4 + a_6Y^6 + a_8Y^8 + \ldots \qquad (4.2)$$

Knowing the aspheric in such a form allows skew rays to be traced enabling an assessment of off-axis imaging. A disadvantage with the point-characteristic method is that the two stage process is uneconomic.

When the aspheric is an interior surface a different procedure has to be followed. The computation is much more laborious. A complete description of the methods and references is collected by Jurek (18). They all involve iterative techniques. Some are based on the point-characteristic method while others determine the polynomial expansion directly. One such example of the latter involves using N terms in the polynomial of equation (4.2) and an equal number of rays. Small changes are made to each of the a_i in turn and the aberration changes for each ray evaluated. If the transverse ray aberration of the initial system and for the Jth ray is T_J and the change due to δa_i is δT_J then a set of partial derivatives, $\frac{\partial T_J}{\partial a_i}$, can be approximated by $\frac{\delta T_J}{\delta a_i}$. Correction for the polynomial coefficients, Δa_i, can then be found by solving the N linear equations,

$$-T_J = \sum_{i=1}^{N} \frac{\delta T_J}{\delta a_i} \Delta a_i \qquad (J = 1,2,\ldots, N) \qquad (4.3)$$

This method assumes a linear variation of T_J with respect to a_i and may result in a slow convergence or even divergence if the starting system is too far from the final result required.

The method used here will be an extension of the technique used in Seidel theory. It is particularly well suited for designing a selectively aberrated system, which will be

required later.

It is well known that the contribution of the term a_4 to the Seidel transverse ray aberration, $T^{(3)}$, is given by,

$$\frac{4a_4y^4\Delta n}{n'u'} \tag{4.4}$$

where y is the paraxial ray height, Δn the index change about the surface considered and n', u' the index and paraxial ray angle in image space respectively.

If $T^{(3)}$ is the aberration due to the spherical surfaces and any other aspherics, then correction can be achieved by choosing a_4 to equal,

$$a_4 = -\frac{n'u'T^{(3)}}{4y^4\Delta n} \tag{4.5}$$

The simple way in which a_4 contributes to $T^{(3)}$ will be shown to apply to higher order aberrations. By choosing a_6 appropriately it is possible to make $T^{(5)}$ vanish. In general $T^{(i)}$ may be corrected by using the coefficient a_{i+1}. In fact the contribution a_{i+1} makes to $T^{(i)}$ is,

$$\frac{(i+1)a_{i+1}y^{i+1}\Delta n}{n'u'} \tag{4.6}$$

This method has several distinct advantages over other techniques. Firstly, it determines the polynomial expansion directly without calculating a set of surface coordinates and then interpolating them. The polynomial terms are found exactly and their computation may be continued until the surface is represented to sufficient accuracy. Also, the method can be applied to any surface with equal ease. It may be the first,

last or interior and can be refracting and reflecting. The surface number is just specified and the appropriate y and Δn used in (4.6).

Perhaps the apparent problem of designing the aspheric this way is the task of calculating the aberration coefficients. To find a polynomial which represents the aspheric accurately, it is common to use as many as ten terms in the expansion. This suggests that using this method we may have to calculate the first ten spherical aberration orders; something which may appear to be a formidable task.

In the next section, however, a surprisingly concise method for doing just this will be given. In fact it allows any number to be found in principle; the limiting factor is the amount of computer storage available. The derivation of the relevant equations is quite lengthy but the end result is easily programmed in about 100 lines.

4.3. Calculation of the spherical aberration coefficients

Since we are dealing with an axially symmetric system and an axial object point we can confine our attention to a single plane; the meridional plane. For a definition of the notation refer to fig. (4.1).

The number of surfaces in the optical system is assumed to be K (numbered 1 to K). When referring to the object plane a 0 will be used. The media are so numbered from 0 to K. A subscript by a quantity denotes the surface or medium to which it applies.

Each surface is defined by the equations,

$$c_i(Y_i^2 + Z_i^2) - 2Z_i + f_i = 0 \qquad (4.7)$$

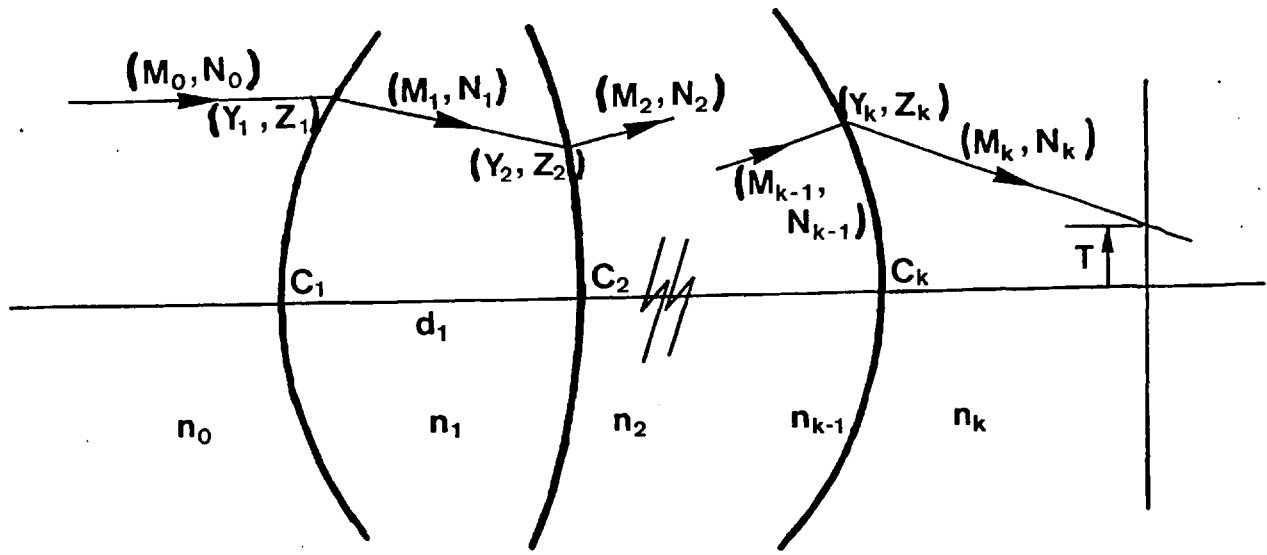


Fig. (4.1) The notation for a general optical system.

$$Z_i = A_{2i}Y_i^2 + A_{4i}Y_i^4 + \ldots = \sum_{J=2,4,.} A_{Ji}Y_i^J \qquad (4.8)$$

where (Y_i, Z_i) defines a point on the surface relative to an origin at the pole. c_i is the paraxial curvature and f_i an even function of Y_i which defines the aspheric nature of the surface.

If the spherical term in equation (4.2) is expanded as a power series in Y_i then an expansion of the form (4.8) results. Both (4.7) and (4.8) will be useful.

We now define β_i and γ_i to be,

$$\beta_i = -2c_iY_i - f_{Yi} \qquad (4.9)$$

$$\gamma_i = 2(1 - c_iZ_i) \qquad (4.10)$$

where $f_{Yi} \equiv \frac{df_i}{dY_i}$. On multipling β_i and γ_i by the factor

$$\frac{1}{\sqrt{f_{Yi}^2 + 4c_iY_if_{Yi} + 4 - 4c_if_i}}$$

they become the Y and Z direction cosines of the surface normal. This factor is dispensed with since it can be cancelled out of the raytracing equation given below. These are,

$$(Y_i - Y_{i-1})N_{i-1} = M_{i-1}(Z_i - Z_{i-1} + d_{i-1}) \qquad (4.11)$$

$$(n_iM_i - n_{i-1}M_{i-1})\gamma_i = (n_iN_i - n_{i-1}N_{i-1})\beta_i \qquad (4.12)$$

$$N_i^2 = 1 - M_i^2 \qquad (4.13)$$

Here, n_i are the refractive indices and d_i the surface spacing. (M_i, N_i) are the direction cosines of a ray.

The final basic equation is known as the Aldis theorem. It is a total aberration expression which gives the transverse

ray aberration of a finite ray relative to a paraxial principal ray. When applied to an axial object point it becomes,

$$n_K u_K N_K T = \sum_{i=1}^{K} B_i\{Y_i(N_i-N_{i-1})-Z_i(M_i-M_{i-1})\}$$

$$+ \frac{y_i f_{Y_i}}{2}(n_i N_i - n_{i-1}N_{i-1}) \qquad (4.14)$$

$B_i = n_i(y_i c_i + u_i) = n_{i-1}(y_i c_i + u_{i-1})$ is called the refraction invariant.

A derivation of the Aldis theorem is given by Cox (19), who then puts it to good use by deriving the Seidel aberrations as a third order approximation along with a few higher order expressions. A description of the derivation is given in Appendix A. In fact the method to be used here is an extension of the technique Cox used.

If the aberration T is written as a sum of surface contributions,

$$T = \sum_{i=1}^{K} T_i$$

then we can regard the terms of (4.14) inside the summation as individual surface contributions. So,

$$n_K u_K N_K T_i = B_i\{Y_i(N_i-N_{i-1})-Z_i(M_i-M_{i-1})\} \qquad (4.15)$$

$$+ \frac{y_i f_{Y_i}}{2}(n_i N_i - n_{i-1}N_{i-1})$$

We now proceed to write each variable as a power series expansion in the variable, $\rho = \frac{Y_1}{Y_{1\,m}}$, the normalized incident ray height (or more precisely, the normalized pupil height). Now due to the symmetry of the system under consideration, if ρ changes sign all variables related to Y_i (i.e. Y_i, M_i, β_i, T_i

and f_{Yi}) will also change sign. Thus we write them as a series in only odd powers of ρ.

$$\left.\begin{aligned} Y_i &= \Sigma\, Y_i^{(J)} \rho^J \qquad J = 1,3,5,\ldots \\ M_i &= \Sigma\, M_i^{(J)} \rho^J \qquad J = 1,3,5,\ldots \\ \beta_i &= \Sigma\, \beta_i^{(J)} \rho^J \qquad J = 1,3,5,\ldots \\ T_i &= \Sigma\, T_i^{(J)} \rho^J \qquad J = 1,3,5,\ldots \\ f_{Yi} &= \Sigma\, f_{Yi}^{(J)} \rho^J \qquad J = 1,3,5,\ldots \end{aligned}\right\} \quad (4.16)$$

The superscript in brackets, (J), denotes the "order" of a particular term. J = 1 are the first order terms and apply to the paraxial region. The paraxial ray height and angles are therefore,

$$\left.\begin{aligned} y_i &= Y_i^{(1)} \rho \\ U_i &= M_i^{(1)} \rho \end{aligned}\right\} \quad (4.17)$$

Similarly, any variable related to Z_i (i.e. Z_i, N_i and γ_i) will not change sign when ρ changes sign. They are written as a sum of terms containing only even powers of ρ.

$$Z_i = \Sigma\, Z_i^{(J)} \rho^J \qquad J = 0,2,4,\ldots \quad (4.18)$$

$$N_i = \Sigma\, N_i^{(J)} \rho^J \qquad J = 0,2,4,\ldots$$
$$\gamma_i = \Sigma\, \gamma_i^{(J)} \rho^J \qquad J = 0,2,4,\ldots \qquad (4.18)$$

The zeroth-order terms are known to be,

$$Z_i^{(0)} = 0$$
$$N_i^{(0)} = 1 \qquad (4.19)$$
$$\gamma_i^{(0)} = 2$$

and the first-order terms,

$$T_i^{(1)} = 0$$
$$f_{Yi}^{(1)} = 0 \qquad (4.20)$$

The first of the latter two express the fact that there are no paraxial aberrations. The second can be deduced by noting that an aspheric differs from the paraxial sphere by terms of fourth-order. That is, f_i is of fourth-order or f_{Yi} of third-order. Thus $f_{Yi}^{(3)}$ is the first non-zero term.

The conditions under which the series will converge are not absolutely clear. Two obvious cases where convergence will not occur are when rays undergo total internal reflection or when they miss a surface altogether. Apart from these, there does not seem to be any reason why the series should diverge. We shall therefore assume that they are valid except for those two exceptions.

Now by substituting the series expansions into equation (4.15) an expression for the $T_i^{(J)}$ can be found.

$$n_K u_K \{\Sigma N_K^{(J)} \rho^J\}\{\Sigma T_i^{(J)} \rho^J\} = B_i\{\Sigma Y_i^{(J)} \rho^J\}\{\Sigma (N_i^{(J)} - N_{i-1}^{(J)}) \rho^J\}$$
$$J=0,2,4,. \quad J=1,3,5,. \quad J=1,3,5,. \quad J=0,2,4,..$$

$$-B_i\{\Sigma Z_i^{(J)} \rho^J\}\{\Sigma (M_i^{(J)} - M_{i-1}^{(J)}) \rho^J\}$$
$$J=0,2,4,. \quad J=1,3,5,..$$

$$+\frac{y_i}{2}\{\Sigma f_{Yi}^{(J)} \rho^J\}\{\Sigma (n_i N_i^{(J)} - n_{i-1} N_{i-1}^{(J)}) \rho^J\} \qquad (4.21)$$
$$J=1,3,5,. \quad J=0,2,4,.$$

Equating the terms of total power ρ^q where q is odd,

$$n_K u_K \sum_{J=1,3,.}^{q} T_i^{(J)} N_K^{(q-J)} = \sum_{J=1,3,.}^{q} B_i\{Y_i^{(J)} (N_i^{(q-J)} - N_{i-1}^{(q-J)})$$

$$-(M_i^{(J)} - M_{i-1}^{(J)}) Z_i^{(q-J)}\}$$

$$+ \frac{y_i}{2} \sum_{J=1,3,.}^{q} f_{Yi}^{(J)} (n_i N_i^{(q-J)} - n_{i-1} N_{i-1}^{(q-J)}) \qquad (4.22)$$

Writing the terms with J=q separately we get,

$$n_K u_K T_i^{(q)} N_K^{(0)} + n_K u_K \sum_{J=1,3,..}^{q-2} T_i^{(J)} N_K^{(q-J)} = B_i\{Y_i^{(q)} (N_i^{(0)} - N_{i-1}^{(0)})$$

$$-(M_i^{(q)} - M_{i-1}^{(q)}) Z_i^{(0)}\} + \frac{y_i}{2} f_{Yi}^{(q)} (n_i N_i^{(0)} - n_{i-1} N_{i-1}^{(0)})$$

$$+ \sum_{J=1,3,.}^{q-2} B_i\{Y_i^{(J)} (N_i^{(q-J)} - N_{i-1}^{(q-J)}) - (M_i^{(J)} - M_{i-1}^{(J)}) Z_i^{(q-J)}\}$$

$$+ \frac{y_i}{2} f_{Yi}^{(J)} (n_i N_i^{(q-J)} - n_{i-1} N_{i-1}^{(q-J)}) \qquad (4.23)$$

Using equations (4.19) to substitute for the zeroth order

terms, and rearranging for $T_i^{(q)}$ finally gives,

$$n_K u_K T_i^{(q)} = \frac{y_i}{2} f_{Yi}^{(q)} (n_i - n_{i-1}) + \sum_{J=1,3,..}^{q-2} B_i \{ Y_i^{(J)} (N_i^{(q-J)} - N_{i-1}^{(q-J)}) - (M_i^{(J)} - M_{i-1}^{(J)}) Z_i^{(q-J)} \} + \frac{y_i}{2} f_{Yi}^{(J)} (n_i N_i^{(q-J)} - n_{i-1} N_{i-1}^{(q-J)}) - n_K u_K T_i^{(J)} N_K^{(q-J)} \quad (4.24)$$

To find all the quantities on the right-hand side we apply a similar treatment to the raytracing and surface equations. Firstly the transfer equation (4.11) gives,

$$\{ \sum_{J=1,3,5,.} (Y_i^{(J)} - Y_{i-1}^{(J)}) \rho^J \} \{ \Sigma N_{i-1}^{(J)} \rho^J \} = \{ \Sigma M_{i-1}^{(J)} \rho^J \} \{ \Sigma (Z_i^{(J)} - Z_{i-1}^{(J)}) \rho^J \}$$
$$J=0,2,4,.. \quad J=1,3,5,. \quad J=0,2,4,..$$

$$+ d_{i-1} \sum_{J=1,3,5,.} M_{i-1}^{(J)} \rho^J \quad (4.25)$$

The term in ρ^q is,

$$\sum_{J=1,3,..}^{q} (Y_i^{(J)} - Y_{i-1}^{(J)}) N_{i-1}^{(q-J)} = \sum_{J=1,3,..}^{q} M_{i-1}^{(J)} (Z_i^{(q-J)} - Z_{i-1}^{(q-J)})$$

$$+ d_{i-1} M_{i-1}^{(q)} \quad (4.26)$$

$$(Y_i^{(q)} - Y_{i-1}^{(q)}) N_{i-1}^{(0)} + \sum_{J=1,3,..}^{q-2} (Y_i^{(J)} - Y_{i-1}^{(J)}) N_{i-1}^{(q-J)} = d_{i-1} M_{i-1}^{(q)}$$

$$+ M_{i-1}^{(q)} (Z_i^{(0)} - Z_{i-1}^{(0)})$$

$$+ \sum_{J=1,3,..}^{q-2} M_{i-1}^{(J)} (Z_i^{(q-J)} - Z_{i-1}^{(q-J)}) \quad (4.27)$$

$$Y_i^{(q)} = Y_{i-1}^{(q)} + d_{i-1}M_{i-1}^{(q)} + \sum_{J=1,3,.}^{q-2} M_{i-1}^{(J)}\left(Z_i^{(q-J)} - Z_{i-1}^{(q-J)}\right) - \left(Y_i^{(J)} - Y_{i-1}^{(J)}\right)N_{i-1}^{(q-J)} \qquad (4.28)$$

The equation of refraction (4.12) becomes

$$\left\{\sum_{J=1,3,.}\left(n_iM_i^{(J)} - n_{i-1}M_{i-1}^{(J)}\right)\rho^J\right\}\left\{\sum_{J=0,2,4.}\gamma_i^{(J)}\rho^J\right\} =$$

$$\left\{\sum_{J=0,2,4,.}\left(n_iN_i^{(J)} - n_{i-1}N_{i-1}^{(J)}\right)\rho^J\right\}\left\{\sum_{J=1,3,5,.}\beta_i^{(J)}\rho^J\right\} \qquad (4.29)$$

So,

$$\sum_{J=1,3,.}^{q}\left(n_iM_i^{(J)} - n_{i-1}M_{i-1}^{(J)}\right)\gamma_i^{(q-J)} = \sum_{J=1,3,.}^{q}\beta_i^{(J)}\left(n_iN_i^{(q-J)} - n_{i-1}N_{i-1}^{(q-J)}\right) \qquad (4.30)$$

$$\left(n_iM_i^{(q)} - n_{i-1}M_{i-1}^{(q)}\right)\gamma_i^{(0)} = \beta_i^{(q)}\left(n_iN_i^{(0)} - n_{i-1}N_{i-1}^{(0)}\right) + \sum_{J=1,3,.}^{q-2}\beta_i^{(J)}\left(n_iN_i^{(q-J)} - n_{i-1}N_{i-1}^{(q-J)}\right) - \left(n_iM_i^{(J)} - n_{i-1}M_{i-1}^{(J)}\right)\gamma_i^{(q-J)} \qquad (4.31)$$

$$2n_iM_i^{(q)} = 2n_{i-1}M_{i-1}^{(q)} + \beta_i^{(q)}\left(n_i - n_{i-1}\right) + \sum_{J=1,3..}^{q-2}\beta_i^{(J)}\left(n_iN_i^{(q-J)} - n_{i-1}N_{i-1}^{(q-J)}\right) - \left(n_iM_i^{(J)} - n_{i-1}M_{i-1}^{(J)}\right)\gamma_i^{(q-J)} \qquad (4.32)$$

To find $N_i^{(q+1)}$ we made use of equation (4.13),

$$\{ \sum_{J=0,2,4,.} N_i^{(J)} \rho^J \}^2 = 1 - \{ \sum_{J=1,3,5.} M_i^{(J)} \rho^J \}^2 \qquad (4.33)$$

The term in ρ^{q+1} is,

$$\sum_{J=0,2,4,.}^{q+1} N_i^{(J)} N_i^{(q+1-J)} = - \sum_{J=1,3,5,.}^{q} M_i^{(J)} M_i^{(q+1-J)} \qquad (4.34)$$

$$N_i^{(0)} N_i^{(q+1)} + N_i^{(q+1)} N_i^{(0)} = - \sum_{J=1,3,.}^{q} M_i^{(J)} M_i^{(q+1-J)} - \sum_{J=2,4,.}^{q-1} N_i^{(J)} N_i^{(q+1-J)} \qquad (4.35)$$

$$N_i^{(q+1)} = -\tfrac{1}{2}\{ \sum_{J=1,3,5,.}^{q} M_i^{(J)} M_i^{(q+1-J)} + \sum_{J=2,4,.}^{q-1} N_i^{(J)} N_i^{(q+1-J)} \} \qquad (4.36)$$

Turning now to the surface equations, $f_{Yi}^{(q)}$ is calculated as follows,

$$f_i = 2Z_i - c_i(Y_i^2 + Z_i^2) \qquad (4.37)$$

$$f_{Yi} = 2Z_{Yi} - 2c_i(Y_i + Z_i Z_{Yi}) \qquad (4.38)$$

Setting $Z_{Yi} = \sum_{J=1,3,.} Z_{Yi}^{(J)} \rho^J$ we arrive at,

$$f_{Yi}^{(q)} = 2\{ Z_{Yi}^{(q)} - c_i Y_i^{(q)} - c_i \sum_{J=1,3,.}^{q-2} Z_{Yi}^{(J)} Z_i^{(q-J)} \} \qquad (4.39)$$

It has now become necessary to evaluate the $Z_{Yi}^{(J)}$ as well. First, from equations (4.9) and (4.10) we immediately see that,

$$\beta_i^{(q)} = -2c_iY_i^{(q)} - f_{Y_i}^{(q)} \tag{4.40}$$

$$\gamma_i^{(q+1)} = -2c_iZ_i^{(q+1)} \tag{4.41}$$

The remaining variables $Z_i^{(q+1)}$ and $Z_{Y_i}^{(q)}$ need a little more attention. This is where the surface equation in polynomial form has to be used, (4.8),

$$Z_i = A_{2i}Y_i^2 + A_{4i}Y_i^4 + A_{6i}Y_i^6 + \ldots. \tag{4.42}$$

$$= \sum_{J=2,4,6,.} A_{Ji}Y_i^J \tag{4.43}$$

Therefore,

$$Z_i^{(q+1)} = \sum_{J=2,4,6,.}^{q+1} A_{Ji}(Y_i^J)^{(q+1)} \tag{4.44}$$

Note that the series terminates at $(Y_i^{q+1})^{(q+1)}$; powers of Y_i greater than q+1 do not contain terms of order q+1. The equivalent expression for $Z_{Y_i}^{(q)}$ is,

$$Z_{Y_i}^{(q)} = \sum_{J=2,4,6,.}^{q+1} JA_{Ji}(Y_i^{J-1})^{(q)} \tag{4.45}$$

We now have the problem of evaluating the terms $(Y_i^J)^{(p)}$. The power J and the order index (p) are both even or odd. If we write Y_i^J as a series in ρ,

$$Y_i^J = \sum_{\iota=J,J+2,.} (Y_i^J)^{(\iota)}\rho^\iota \tag{4.46}$$

then since,

$$Y_i^J = Y_i \cdot Y_i^{J-1} \tag{4.47}$$

We obtain,

$$\sum_{l=J,J+2,..} (Y_i^J)^{(l)} \rho^l = \{ \sum_{l=1,3,..} Y_i^{(l)} \rho^l \} \{ \sum_{l=J-1,J+1,..} (Y_i^{J-1})^{(l)} \rho^l \} \tag{4.48}$$

The term in ρ^p is,

$$(Y_i^J)^{(p)} = \sum_{l=1,3,..}^{p-J+1} Y_i^{(l)} (Y_i^{J-1})^{(p-l)} \tag{4.49}$$

The upper limit to the summation index, l, is when $J-1 = p-l$ or $l=p-J+1$, which is always an odd number. There are two forms in which this is needed; to cope with the even and odd powers of Y_i in the expressions for $Z_i^{(q+1)}$ and $Z_{Yi}^{(q)}$

$$(Y_i^J)^{(q+1)} = \sum_{l=1,3,..}^{q-J+2} Y_i^{(l)} (Y_i^{J-1})^{(q-l+1)} \qquad (J=2,4,..,q+1) \tag{4.50}$$

$$(Y_i^J)^{(q)} = \sum_{l=1,3,..}^{q-J+1} Y_i^{(l)} (Y_i^{J-1})^{(q-l)} \qquad (J=3,5,..,q) \tag{4.51}$$

The final problem is to work out the coefficients in the polynomial expression for Z_i; the A_{Ji}. These are simply given by,

$$A_{Ji} = a_{Ji} + \sigma_J c_i^{J-1} \tag{4.52}$$

The σ_J are constants and may be generated by the recurrence relation,

$$\sigma_{J+2} = \left(\frac{J-1}{J+2}\right) \sigma_J \qquad (4.53)$$

with $\sigma_2 = \frac{1}{2}$

They originate from the binomial expansion of Z_i for a spherical surface,*

$$Z_i = \frac{1}{c_i} - \frac{1}{c_i}\sqrt{1 - c_i^2 Y_i^2} = \sum_{J=2,4,..} \sigma_J c_i^{J-1} Y_i^J \qquad (4.54)$$

These equations are those required to evaluate the aberration coefficients and must be used in the correct sequence (see next section). The order of use can most easily be deduced by writing them out for the case of q=3. By making use of two dimensional arrays and one three dimensional array for $(Y_i^J)^{(q)}$, they can be compactly programmed. First and second order terms are more conveniently calculated in a separate subroutine.

If the equations are programmed exactly as written, the total storage is twice what it needs to be. For instance, $M_i^{(J)}$ is only defined for J=1,3,5,7,...; only every other location in the J-dimension is needed. To halve the memory, they should be stored in the array locations 1,2,3,4,.. respectively. In the (J+1)/2 location, generally.

If the array is even ordered like $Z_i^{(J)}$ the location is J/2. The three-dimensional array $(Y_i^J)^{(q)}$ can only have one index treated in this way, however.

*The binomial expansion for $\sqrt{1 - x^2}$ converges for $|x| < 1$ which covers the full range required.

The next section shows a summary of the procedure together with the initial conditions to be set.

4.4. Initial conditions and procedure

Since we are dealing with an axial and infinitely remote object point, $M_0 = 0$ and $N_0 = 1$. Therefore,

$$M_0^{(J)} = 0 \text{ for } J = 1,3,5,\ldots$$
$$N_0^{(J)} = 0 \text{ for } J = 2,4,6,\ldots$$

The incident ray height is chosen to be Y_{1m} so,

$$Y_0^{(1)} = Y_{1m}, \text{ the maximum ray height}$$
$$Y_0^{(J)} = 0 \text{ for } J = 3,5,7,\ldots$$
$$Z_0^{(J)} = 0 \text{ for } J = 2,4,6,\ldots$$
$$d_0 = 0$$

The aspheric polynomial coefficients, a_{Ji}, should be set to the appropriate values; zero if the surface is spherical.

Generate the σ_J,

$$\sigma_2 = \tfrac{1}{2}$$

$$\sigma_{J+2} = \left(\frac{J-1}{J+2}\right)\sigma_J \quad \text{for} \quad J=2,4,6,..$$

Calculate for each surface ($i = 1,2,\ldots,k$),

$$A_{Ji} = a_{Ji} + \sigma_J c_i^{J-1} \quad \text{for } J = 2,4,6,\ldots$$

For each surface evaluate all the first and second order quantities.

$$Y_i^{(1)} = Y_{i-1}^{(1)} + d_{i-1}M_{i-1}^{(1)}$$

$$(Y_i^2)^{(2)} = (Y_i^{(1)})^2$$

$$Z_i^{(2)} = \frac{c_i(Y_i^2)^{(2)}}{2}$$

$$Z_{Yi}^{(1)} = c_i Y_i^{(1)}$$

$$\beta_i^{(1)} = -2c_i Y_i^{(1)}$$

$$M_i^{(1)} = (2n_{i-1}M_{i-1}^{(1)} + \beta_i^{(1)}(n_i - n_{i-1}))/2n_i$$

$$\gamma_i^{(2)} = -2c_i Z_i^{(2)}$$

$$N_i^{(2)} = -\tfrac{1}{2}M_i^{(1)2}$$

$$T_i^{(1)} = f_{Yi}^{(1)} = 0$$

$$B_i = n_i(c_i Y_i^{(1)} + M_i^{(1)})$$

The higher orders may now be found. For each q=3,5,7,... calculate the following for each surface (i = 1,2,..,K)

$$Y_i^{(q)} = Y_{i-1}^{(q)} + d_{i-1}M_{i-1}^{(q)} + \sum_{J=1,3,..}^{q-2} M_{i-1}^{(J)}(Z_i^{(q-J)} - Z_{i-1}^{(q-J)}) - (Y_i^{(J)} - Y_{i-1}^{(J)})N_{i-1}^{(q-J)}$$

$$(Y_i^J)^{(q)} = \sum_{l=1,3,..}^{q-J+1} Y_i^{(l)}(Y_i^{J-1})^{(q-l)} \quad \text{for } J=3,5,\ldots,q$$

$$(Y_i^J)^{(q+1)} = \sum_{l=1,3,.}^{q-J+2} Y_i^{(l)}(Y_i^{J-1})^{(q-l+1)} \quad \text{for } J=2,4,..,q+1$$

$$Z_{Yi}^{(q)} = \sum_{J=2,4,.}^{q+1} JA_{Ji}(Y_i^{J-1})^{(q)}$$

$$f_{Yi}^{(q)} = 2\{Z_{Yi}^{(q)} - c_i Y_i^{(q)} - c_i \sum_{J=1,3,.}^{q-2} Z_{Yi}^{(J)} Z_i^{(q-J)}\}$$

$$\beta_i^{(q)} = -2c_i Y_i^{(q)} - f_{Yi}^{(q)}$$

$$M_i^{(q)} = \frac{n_{i-1}}{n_i} M_{i-1}^{(q)} + \frac{n_i - n_{i-1}}{2n_i} \beta_i^{(q)}$$

$$+ \frac{1}{2n_i} \sum_{J=1,3,.}^{q-2} \beta_i^{(J)} (n_i N_i^{(q-J)} - n_{i-1} N_{i-1}^{(q-J)})$$

$$-(n_i M_i^{(J)} - n_{i-1} M_{i-1}^{(J)}) \gamma_i^{(q-J)}$$

$$Z_i^{(q+1)} = \sum_{J=2,4,.}^{q+1} A_{Ji}(Y_i^{J})^{(q+1)}$$

$$\gamma_i^{(q+1)} = -2c_i Z_i^{(q+1)}$$

$$N_i^{(q+1)} = -\tfrac{1}{2}\{ \sum_{J=1,3,.}^{q} M_i^{(J)} M_i^{(q-J+1)} + \sum_{J=2,4,.}^{q-1} N_i^{(J)} N_i^{(q-J+1)} \}$$

and,

$$n_K M_K^{(1)} T_i^{(q)} = \frac{Y_i^{(1)}}{2} f_{Yi}^{(q)} (n_i - n_{i-1}) + \sum_{J=1,3,.}^{q-2} B_i\{Y_i^{(J)} (N_i^{(q-J)} - N_{i-1}^{(q-J)})$$

$$-(M_i^{(J)} - M_{i-1}^{(J)}) Z_i^{(q-J)}\} + \frac{Y_i^{(1)}}{2} f_{Yi}^{(J)} (n_i N_i^{(q-J)} - n_{i-1} N_{i-1}^{(q-J)})$$

$$-n_K M_K^{(1)} T_i^{(J)} N_K^{(q-J)}$$

$$T^{(q)} = \sum_{i=1,2,..}^{K} T_i^{(q)}$$

The wavefront aberration can always be derived from these by integration.

Appendix C shows a FORTRAN listing of subroutines, INTI, FIRST and HIGHER which program the initial conditions, first and second order terms, and higher order quantities respectively.

4.5. Calculating the shape of the aspheric surface.

The original purpose behind calculating the various aberration orders was to find the polynomial for an aspheric in the form,

$$Z_i = \frac{c_i Y_i^2}{1 + \sqrt{1 - c_i^2 Y_i^2}} + a_{4i} Y_i^4 + a_{6i} Y_i^6 + a_{8i} Y_i^8 + \ldots \qquad (4.55)$$

that will correct spherical aberration. This we now proceed to do.

Recall the expression for $T_i^{(q)}$. Summing over the surfaces and setting the total transverse ray aberration,

$$T^{(q)} = \sum_{i=1,2,.}^{K} T_i^{(q)}$$ we obtain

$$n_K M_K^{(1)} T^{(q)} = \sum_{i=1,2,.}^{K} \left[\frac{Y_i^{(1)}}{2} f_{Yi}^{(q)} (n_i - n_{i-1}) + \sum_{J=1,3,.}^{q-2} B_i \{ Y_i^{(J)} (N_i^{(q-J)} - N_{i-1}^{(q-J)}) \right.$$

$$- (M_i^{(J)} - M_{i-1}^{(J)}) Z_i^{(q-J)} \} + \frac{Y_i^{(1)}}{2} f_{Yi}^{(J)} (n_i N_i^{(q-J)} - n_{i-1} N_{i-1}^{(q-J)})$$

$$\left. - n_K M_K^{(1)} T_i^{(J)} N_K^{(q-J)} \right] \qquad (4.56)$$

Since $T^{(q)}$ can be adjusted by changing the a_{q+1}, we need to sort out the terms that contain it. The a_{q+1} is actually the coefficient of Y_i^{q+1} and therefore would only normally appear in terms of order q+1 and higher. The highest order term in the above expression is only q, so a_{q+1i} can only appear via a derivative. Thus $f_{Yi}^{(q)}$ is responsible. Now

$$f_{Yi}^{(q)} = 2\{Z_{Yi}^{(q)} - c_i Y_i^{(q)} - c_i \sum_{J=1,3,.}^{q-2} Z_{Yi}^{(J)} Z_i^{(q-J)}\} \qquad (4.57)$$

and

$$Z_{Yi}^{(q)} = \sum_{J=2,4,.}^{q+1} J A_{Ji} (Y_i^{(J-1)})^{(q)} \qquad (4.58)$$

$$Z_{Yi}^{(q)} = \sum_{J=2,4,.}^{q+1} J(a_{Ji} + \sigma_J c_i^{J-1})(Y_i^{J-1})^{(q)} \qquad (4.59)$$

$$= 2(a_{2i} + \sigma_2 c_i) Y_i^{(q)} + \sum_{J=4,6,.}^{q-1} J(a_{Ji} + \sigma_J c_i^{J-1})(Y_i^{J-1})^{(q)}$$

$$+ (q+1)(a_{q+1i} + \sigma_{q+1} c_i^q)(Y_i^q)^{(q)} \qquad (4.60)$$

But $a_{2i} = 0$, $\sigma_2 = \frac{1}{2}$ and $(Y_i^q)^{(q)} = Y_i^{(1)q}$,

$$Z_{Yi}^{(q)} = c_i Y_i^{(q)} + \sum_{J=4,6,.}^{q-1} J(a_{Ji} + \sigma_J c_i^{J-1})(Y_i^{J-1})^{(q)}$$

$$+ (q+1)(a_{q+1i} + \sigma_{q+1} c_i^q) Y_i^{(1)q} \qquad (4.61)$$

and therefore,

$$f_{Yi}^{(q)} = 2\{(q+1)(a_{q+1i}+\sigma_{q+1}c_i^{q})Y_i^{(1)q} + \sum_{J=4,6,.}^{q-1} J(a_{Ji}+\sigma_J c_i^{J-1})(Y_i^{J-1})^{(q)}$$

$$- c_i \sum_{J=1,3,.}^{q-2} Z_{Yi}^{(J)} Z_i^{(q-J)}\} \qquad (4.62)$$

The contribution to $f_{Yi}^{(q)}$ from a_{q+1i} is therefore,

$$2(q+1)a_{q+1i}Y_i^{(1)q} \qquad (4.63)$$

and to $n_K M_K^{(1)}T^{(q)}$ from $f_{Yi}^{(q)}$,

$$\sum_{i=1,2,..}^{K} \frac{Y_i^{(1)}}{2} f_{Yi}^{(q)}(n_i-n_{i-1}) \qquad (4.64)$$

Combining these gives the contribution to $n_K M_K^{(1)}T^{(q)}$ from the a_{q+1i} as,

$$\sum_{i=1,2,.}^{K} (q+1)a_{q+1i}(n_i-n_{i-1})Y_i^{(1)q+1} \qquad (4.65)$$

Only one aspheric is required to correct spherical aberration, the ith say. Now if $T^{(q)}$ is the aberration present when $a_{q+1i}=0$ and $T_{req.}^{(q)}$ is the required value for the complete system then,

$$n_K M_K^{(1)}T^{(q)} + (q+1)a_{q+1i}(n_i-n_{i-1})Y_i^{(1)q+1} = n_K M_K^{(1)}T_{req.}^{(q)} \qquad (4.66)$$

Hence,

$$a_{q+1i} = \frac{n_K M_K^{(1)} (T_{req}^{(q)} - T^{(q)})}{(q+1)(n_i - n_{i-1}) Y_i^{(1)q+1}} \qquad (4.67)$$

To satisfy the constant optical path condition, $T_{req}^{(q)}=0$, otherwise an aspheric can be designed which selectively aberrates the system.

Initially all the a_{Ji} should be set to zero. Each is evaluated in turn from (4.67) and replaces the zero in store. After calculating each a_{q+1i} it is necessary to recalculate quantities of order q before proceeding to the next term, because some of them are changed by the new value of a_{q+1i}. Once this is done it will be unnecessary to repeat it again since terms of this order are unaffected by changes in higher order variables yet to be calculated. Fig.(4.2) shows the flow chart for the process.

This method can be applied to any surface in the optical system and the process continued until the polynomial represents the aspheric to sufficient accuracy; the number of polynomial coefficients need not be fixed.

4.6. Target intensity from the polynomial expansions.

The imaging properties of an optical system can be determined in two different ways. One way is to trace a large number of finite rays at different apertures while the other is to trace a single ray but to many degrees of accuracy, i.e. to first order (as a paraxial ray), to third, fifth

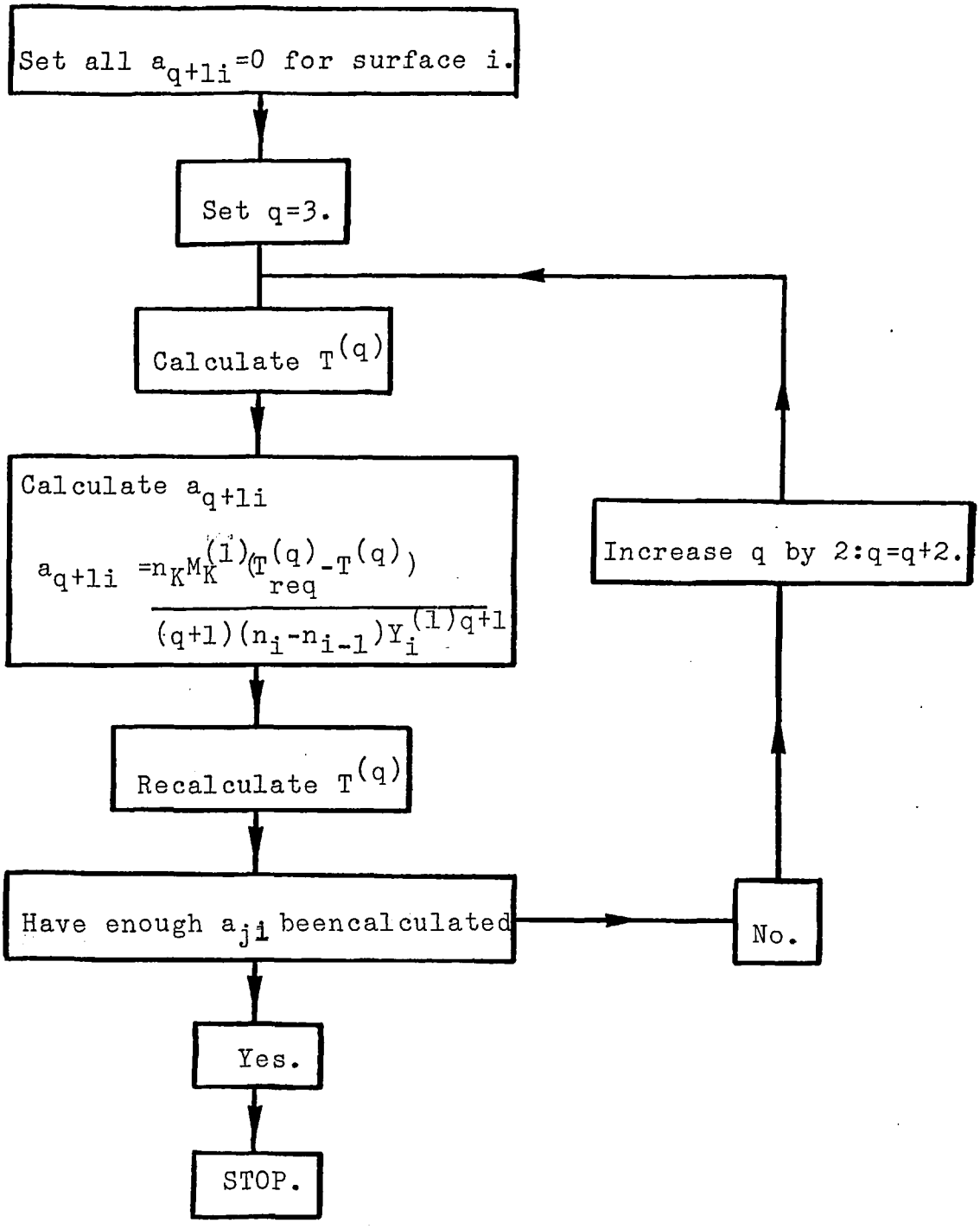


Fig. (4.2) Flow chart of the aspherizing process.

orders etc. The same information is obtained by both methods. There is no reason therefore, why the target intensity cannot be found from the polynomial expansions. The relevant equations are (2.2) and (2.3).

$$\frac{I_T}{I_B} = \frac{Y}{R_T^2 \sin \phi} \cdot \frac{dY}{d\phi} \tag{4.68}$$

and

$$\tan U_K = \frac{-T}{LA} = \frac{R_T \sin \phi}{R_T \cos \phi + LA + S_0} \tag{4.69}$$

Rearranging the second and differentiating with respect to Y,

$$(R_T \cos \phi + LA + S_0) \tan U_K = R_T \sin \phi \tag{4.70}$$

$$(-R_T \sin \phi \frac{d\phi}{dY} + \frac{dLA}{dY}) \tan U_K + (R_T \cos \phi + LA + S_0) \frac{d \tan U_K}{dY} = R_T \cos \phi \frac{d\phi}{dY} \tag{4.71}$$

Solving for $\frac{dY}{d\phi}$,

$$\frac{dY}{d\phi} = \frac{R_T(\cos \phi + \sin \phi \tan U_K)}{\frac{dLA}{dY} \tan U_K + (R_T \cos \phi + LA + S_0) \frac{d \tan U_K}{dY}} \tag{4.72}$$

where $$LA = \frac{-T}{\tan U_K} \tag{4.73}$$

$$\frac{dLA}{dY} = -\frac{1}{\tan U_K} \frac{dT}{dY} + \frac{T}{\tan^2 U_K} \frac{d \tan U_K}{dY} \tag{4.74}$$

$$\frac{dLA}{dY} = \frac{1}{\tan U_K} \left(\frac{T}{\tan U_K} \frac{d \tan U_K}{dY} - \frac{dT}{dY} \right) \tag{4.75}$$

and

$$\tan U_K = \frac{M}{N} \tag{4.76}$$

$$\frac{d\tan U_K}{dY} = \frac{1}{N}\frac{dM}{dY} - \frac{M}{N^2}\frac{dN}{dY} \tag{4.77}$$

$$\frac{d\tan U_K}{dY} = \frac{1}{N}\left(\frac{dM}{dY} - \tan U_K \frac{dN}{dY}\right) \tag{4.78}$$

Finally, since we know M,N and T in series form,

$$M = M^{(1)}\rho + M^{(3)}\rho^3 + \ldots \tag{4.79}$$

$$N = 1 + N^{(2)}\rho^2 + N^{(4)}\rho^4 + \ldots \tag{4.80}$$

and $$T = T^{(3)}\rho^3 + T^{(5)}\rho^5 + \ldots. \tag{4.81}$$

$$\rho = \frac{Y}{Y_m} \tag{4.82}$$

then

$$\frac{dM}{dY} = \frac{1}{Y_m}\frac{dM}{d\rho} \tag{4.83}$$

$$\frac{dN}{dY} = \frac{1}{Y_m}\frac{dN}{d\rho} \tag{4.84}$$

$$\frac{dT}{dY} = \frac{1}{Y_m}\frac{dT}{d\rho} \tag{4.85}$$

Some of these expressions become indeterminate on the axis. In the paraxial limit equation (4.69) becomes,

$$u_K = \frac{R_T\phi}{R_T + S_0} \tag{4.86}$$

and since $u_K = -y/F$ also, equating gives

$$y = \frac{-FR_T\phi}{R_T + S_0} \tag{4.87}$$

$$\frac{dy}{d\phi} = \left.\frac{dY}{d\phi}\right|_{axis} = \frac{-FR_T}{R_T + S_0} \tag{4.88}$$

Therefore,

$$\left.\frac{I_T}{I_B}\right|_{axis} = \frac{1}{R_T^2}\left[\frac{-FR_T}{R_T + S_0}\right]^2 \tag{4.89}$$

$$\left.\frac{I_T}{I_B}\right|_{axis} = \frac{F^2}{(R_T + S_0)^2} \tag{4.90}$$

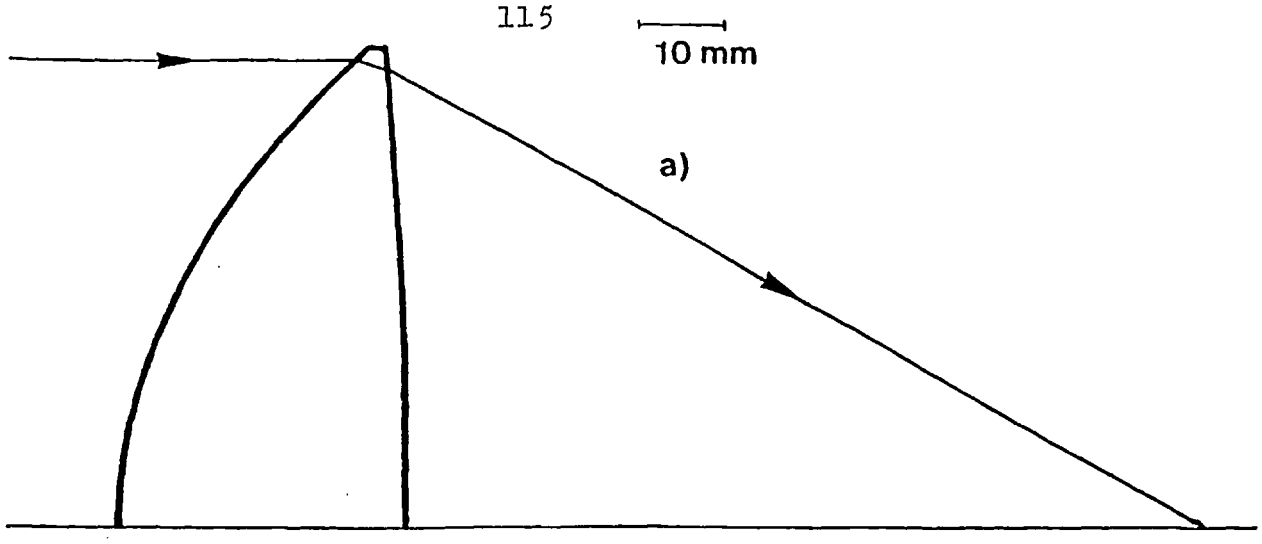
This result obviously follows from the inverse square law.

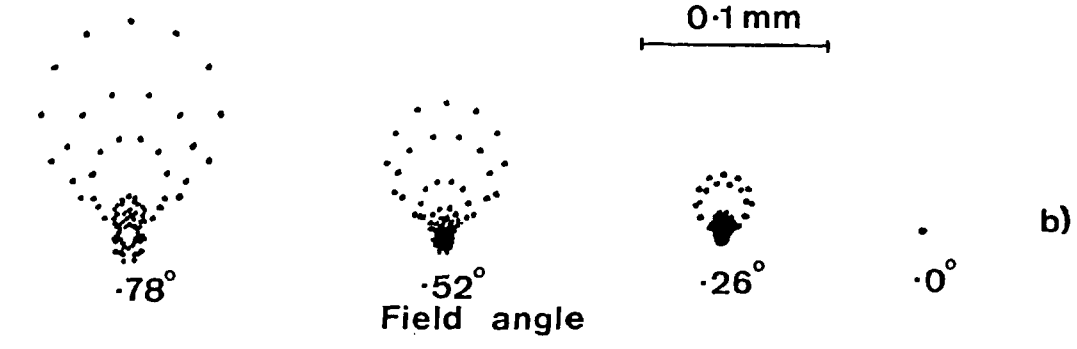
4.7. Examples corrected for constant optical path.

Using the techniques given over the last few sections the two lenses shown in Figs. (4.3) and (4.4) were designed. In each case the first surface is the aspheric and corrects the aberrations at 1·064μm. At the marginal zone the aspheric deviates from the paraxial sphere by approximately 8·0 and 0·65mm for the singlet and doublet respectively. System data is contained in an accompanying tables (4.1) and (4.2).

The singlet is bent so that OSC=0 at the marginal zone. The accompanying spot diagrams still show a large coma flare. This situation is improved in the case of the doublet.

Also shown in the diagram is the target illumination provided when the beam is assumed to be uniform. The target radius was taken to be 50μm, a typical value. Not surprisingly the curves resemble those of the strictly aplanatic lens in





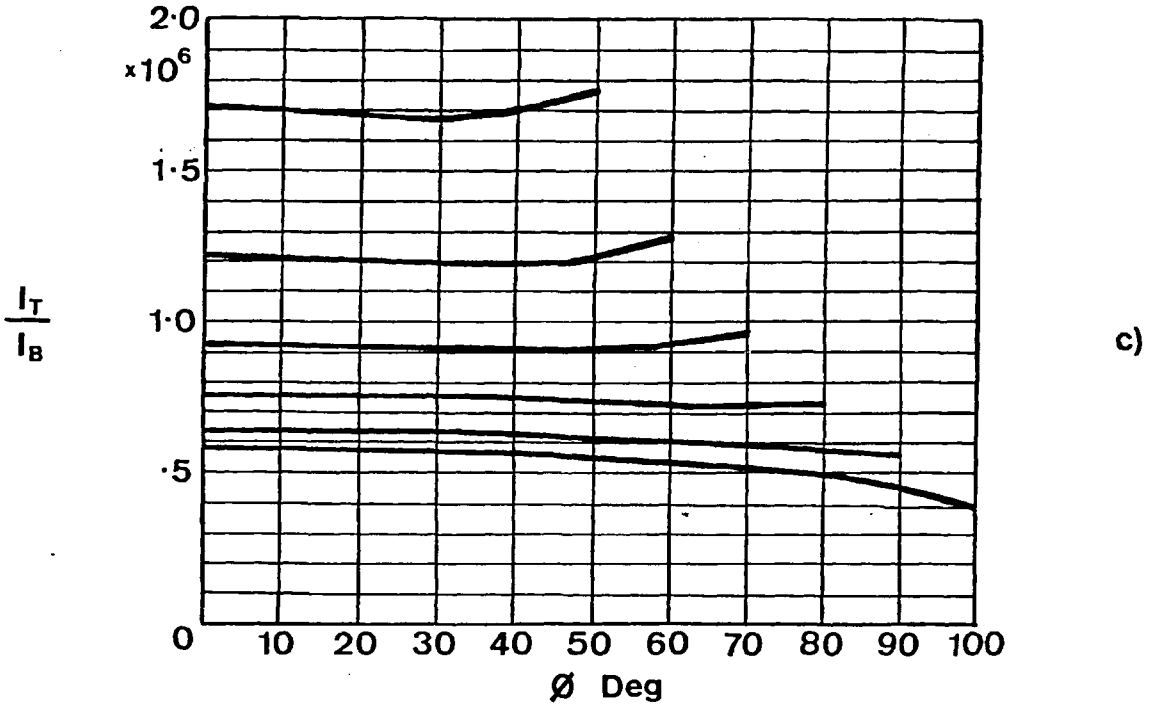


Fig. (4.3) a.) Singlet with an aspheric front surface that null corrects the lens at 1·064μm. b=90·2938mm, F=109·9427mm.

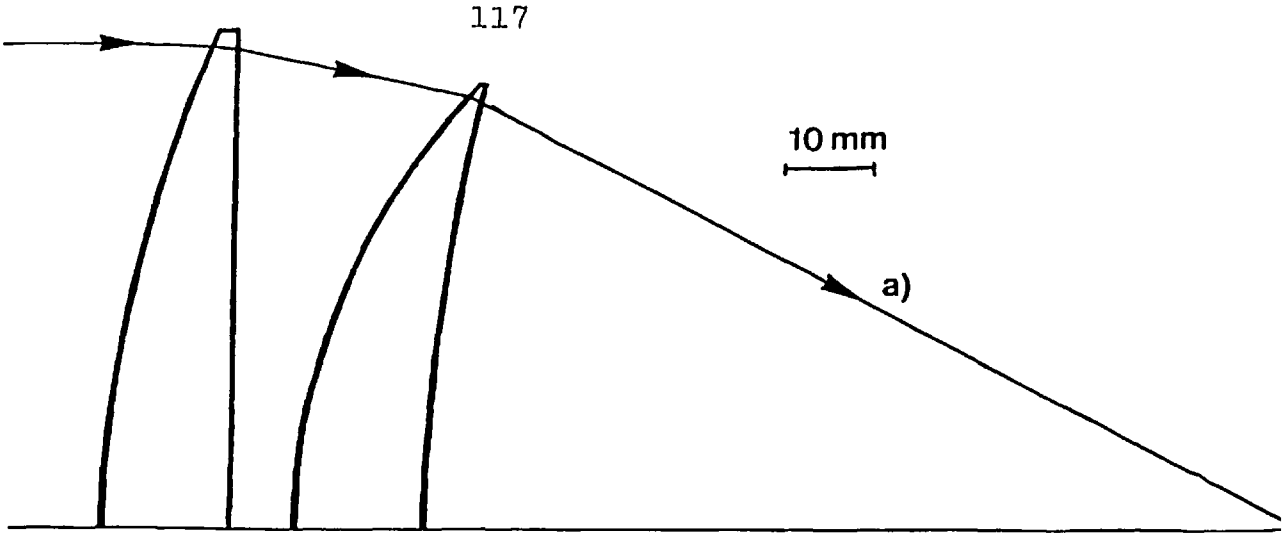
b.) Spot diagrams.

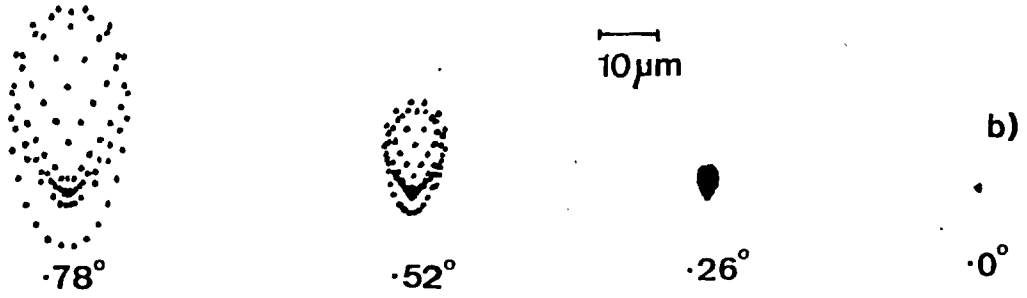
c.) Target illumination for various coverages. R_T=50μm.

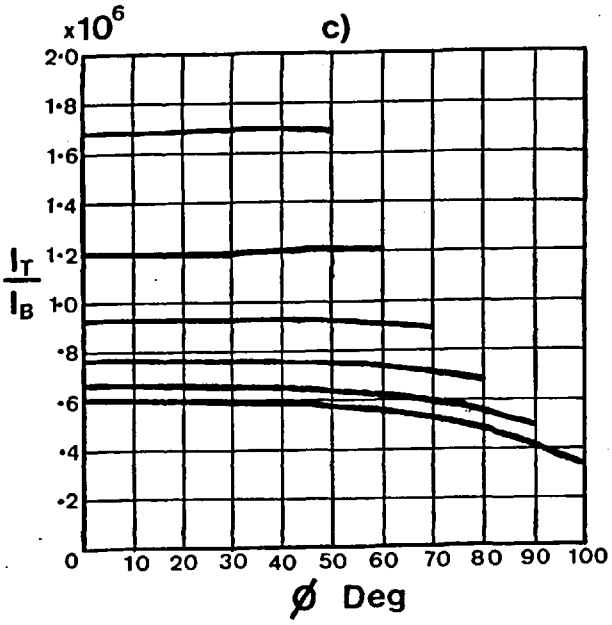
c_i (paraxial)	d_i
•016281467	
	32•6612
-•002053706	

j	a_{j1}
4	$-•3666015850 \times 10^{-6}$
6	$-•7332799117 \times 10^{-10}$
8	$-•1290773125 \times 10^{-13}$
10	$-•2303069395 \times 10^{-17}$
12	$-•4399495187 \times 10^{-21}$
14	$-•9031076404 \times 10^{-25}$
16	$-•1948109603 \times 10^{-28}$
18	$-•4323434205 \times 10^{-32}$
20	$-•9764994664 \times 10^{-36}$
22	$-•2236120629 \times 10^{-39}$
24	$-•5184215017 \times 10^{-43}$
26	$-•1215267389 \times 10^{-46}$

Table (4.1) System data for the singlet in fig. (4.3).







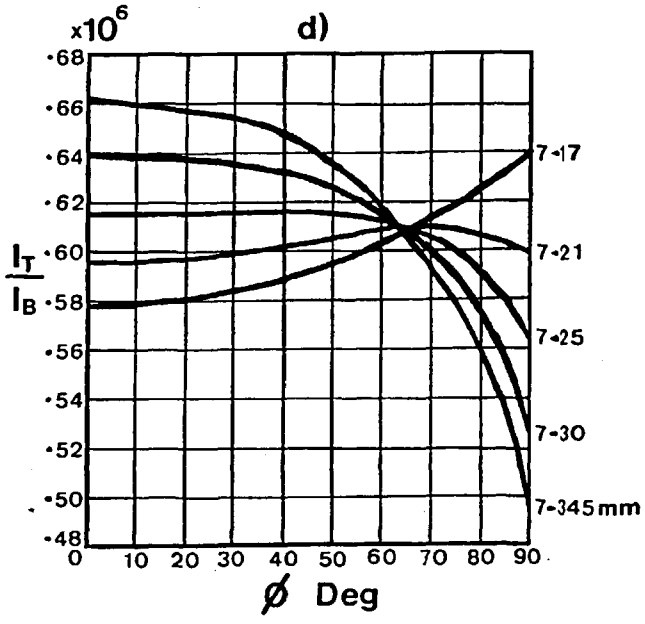


Fig. (4.4) a.) Doublet with an aspheric front surface that null corrects it at 1·064μm.
b=99·9081mm, F=119·6822mm.

b.) Spot diagrams.

c.) Target illumination for various coverages.
R_T=50μm.

d.) The effect of changing the element spacing on the illumination. The case of R_T=50μm and ϕ_m=-90°.

c_i (paraxial)	d_i
•0077101	
	14•5
•0	
	7•345
•01349528	
	14•5
•00455873	

j	a_{j1}
4	$-•6044701720 \times 10^{-7}$
6	$-•3118263746 \times 10^{-11}$
8	$-•1527101765 \times 10^{-15}$
10	$-•8807432364 \times 10^{-20}$
12	$-•6152588999 \times 10^{-24}$
14	$-•4969819604 \times 10^{-28}$
16	$-•4443624879 \times 10^{-32}$
18	$-•4265541439 \times 10^{-36}$
20	$-•4303783316 \times 10^{-40}$
22	$-•4500745266 \times 10^{-44}$
24	$-•4834692349 \times 10^{-48}$
26	$-•5303422890 \times 10^{-52}$

Table (4.2) System data for the doublet in fig. (4.4).

Fig. (2.3). As mentioned at the time, illumination profiles given by an aplanat tend to drop away at the outer zones and that this effect is increased by the incident beam whose intensity may also decrease at the outer zones. This will be particularly true if no attempt to smooth out the Gaussian profile is made. It is with this point that the doublet has a distinct advantage over the singlet. If we were to calculate the shape of the doublet's principal surface and put the result in equation (4.1), the result is the amount of aberration required to make the illumination uniform. The aberration may be approximated by decreasing the component spacing slightly, as illustrated by the second graph in Fig. (4.4). The case with coverage ϕ_m=-90° has been chosen, though no doubt it can be applied to any ϕ_m and R_T.

At the spacing 7•345mm the intensity drops 25% from center to edge. When decreased to 7•21mm the variation has been reduced to about 3%. Taking it further, we can actually get the illumination to increase with target angle for a uniform beam thus counteracting any decrease in beam profile.

The reflection losses of an uncoated lens could also be compensated for in this manner. Changes in spacing of only a millimeter or less are required. Enough to produce a significant change in the target intensity while leaving the losses virtually unaltered.

It therefore appears possible to optimize the illumination for any coverage, target size and beam intensity simply by adjusting the element spacing.

4.8. Uniform illumination using one aspheric surface or corrector plate.

To achieve an absolutely uniform illumination the aberration required is,

$$\frac{T(\rho)}{R_T} = -2\rho\sin(\frac{\phi_m}{2})\sqrt{1 - \rho^2\sin^2(\frac{\phi_m}{2})}$$

$$-2(\rho^2\sin(\frac{\phi_m}{2}) + \frac{F}{Y_m})\sin(\frac{\phi_m}{2})\tan U_K \qquad (4.91)$$

To provide it, a single aspheric surface is necessary. Thomas has also suggested that corrector plates could be used.

If we expand the above expression in powers of ρ by writing,

$$T(\rho) = \sum_{J=1,3,5,.} T^{(J)}\rho^J \qquad (4.92)$$

$$\tan U_K = \sum_{J=1,3,5,.} \tau_K^{(J)}\rho^J \qquad (4.93)$$

and $$\sqrt{1 - \rho^2\sin^2(\frac{\phi_m}{2})} = 1 - \sum_{J=2,4,6,.} \sigma_J\rho^J\sin^J(\frac{\phi_m}{2}) \qquad (4.94)$$

where the σ_J are the same constants used earlier and are listed in table (4.3).

$$\frac{1}{R_T}\sum_{J=1,3,5,.} T^{(J)}\rho^J = -2\rho\sin(\frac{\phi_m}{2})\{1 - \sum_{J=2,4,6,.} \sigma_J\rho^J\sin^J(\frac{\phi_m}{2})\}$$

$$-2\,(\rho^2\sin(\frac{\phi_m}{2}) + \frac{F}{Y_m})\sin(\frac{\phi_m}{2})\sum_{J=1,3,5,.} \tau_K^{(J)}\rho^J \qquad (4.95)$$

J	σ_J
2	•5
4	•125
6	•0625
8	•039063
10	•027344
12	•020508
14	•016113
16	•013092
18	•010910
20	•009274
22	•008009
24	•007008

Table (4.3) List of binomial coefficients.

The term in ρ is,

$$\frac{T^{(1)}}{R_T} = -2\sin(\frac{\phi_m}{2}) - \frac{2F}{Y_m}\sin(\frac{\phi_m}{2})\tau_K^{(1)} \tag{4.96}$$

Since $T^{(1)}=0$ this just gives

$$\tau_K^{(1)} = -\frac{Y_m}{F} \tag{4.97}$$

the paraxial ray angle.

For terms in ρ^q where $q=3,5,7,...$

$$\frac{T^{(q)}}{R_T} = 2\sigma_{q-1}\sin^q(\frac{\phi_m}{2}) - 2\sin^2(\frac{\phi_m}{2})\tau_K^{(q-2)} - \frac{2F}{Y_m}\sin(\frac{\phi_m}{2})\tau_K^{(q)} \tag{4.98}$$

$$T^{(q)} = 2R_T\{\sigma_{q-1}\sin^{q-1}(\frac{\phi_m}{2}) - \tau_K^{(q-2)}\sin(\frac{\phi_m}{2}) - \frac{F}{Y_m}\tau_K^{(q)}\}\sin(\frac{\phi_m}{2}) \tag{4.99}$$

The $\tau_K^{(q)}$ are in turn easily found, since

$$\tan U_K = \frac{M_K}{N_K} \tag{4.100}$$

or

$$N_K \tan U_K = M_K \tag{4.101}$$

$$\{\sum_{J=0,2,4,.} N_K^{(J)}\rho^J\}\{\sum_{J=1,3,.} \tau_K^{(J)}\rho^J\} = \sum_{J=1,3,.} M_K^{(J)}\rho^J \tag{4.102}$$

So the term in ρ^q is,

$$\sum_{J=1,3,5,.}^{q} \tau_K^{(J)} N_K^{(q-J)} = M_K^{(q)} \tag{4.103}$$

$$\tau_K^{(q)} = M_K^{(q)} - \sum_{J=1,3,.}^{q-2} \tau_K^{(J)} N_K^{(q-J)} \quad (q=3,5,7,.) \qquad (4.104)$$

and

$$\tau_K^{(1)} = \frac{-Y_m}{F} \qquad (4.105)$$

The procedure involves using (4.104) above to find the $\tau_K^{(q)}$ for a particular lens and then (4.99) gives the required aberration coefficients, $T^{(q)}$. The latter become the $T_{req}^{(q)}$ used in the aspherizing procedure already described.

By introducing the prescribed amount of aberration, we inevitably change tan U_K very slightly. So the process is iterative. Convergence is likely to be very rapid, however. This is borne out in practice. The examples given only needed a single iteration.

Starting from the singlet in Fig.(4.3), a plane-parallel plate 5mm thick and positioned 2mm in front of the lens can serve as the basis for a corrector plate. If $\phi_m=-90^0$ and target radii of 40μm, 50μm and 60μm are chosen, the aspheric polynomials for the second surface of the plate that result are shown in table (4.4).

The amount of aberration (proportional to the target size) is small. For the 60μm target it is a 0·368λ wavefront error at the marginal zone. The change in OSC was negligible.

Instead of using a corrector plate the first surface of the lens itself can be re-aspherized, say for the average target. For R_T=50μm table (4.5) shows the result of doing this. The price to pay for this is that the aspheric cannot be null tested during manufacture without auxiliary systems.

J	Target radius (μm)		
	40	50	60
4	$-\cdot6505899231 \times 10^{-10}$	$-\cdot8132374039 \times 10^{-10}$	$-\cdot9758848846 \times 10^{-10}$
6	$\cdot6556471326 \times 10^{-15}$	$\cdot8195589152 \times 10^{-15}$	$\cdot9834706981 \times 10^{-15}$
8	$\cdot4359742042 \times 10^{-18}$	$\cdot5449871347 \times 10^{-18}$	$\cdot6540078171 \times 10^{-18}$
10	$\cdot8536233014 \times 10^{-23}$	$\cdot1067167039 \times 10^{-22}$	$\cdot1280765954 \times 10^{-22}$
12	$-\cdot7852110855 \times 10^{-26}$	$-\cdot9815700474 \times 10^{-26}$	$-\cdot1177951488 \times 10^{-25}$
14	$-\cdot1443334797 \times 10^{-29}$	$-\cdot1804248644 \times 10^{-29}$	$-\cdot2165194559 \times 10^{-29}$
16	$-\cdot1095807889 \times 10^{-33}$	$-\cdot1369666098 \times 10^{-33}$	$-\cdot1643486782 \times 10^{-33}$
18	$\cdot3418255021 \times 10^{-38}$	$\cdot4276386032 \times 10^{-38}$	$\cdot5135944584 \times 10^{-38}$
20	$\cdot1775085788 \times 10^{-41}$	$\cdot2219098281 \times 10^{-41}$	$\cdot2663207217 \times 10^{-41}$
22	$\cdot1438902163 \times 10^{-45}$	$\cdot1798061727 \times 10^{-45}$	$\cdot2156994694 \times 10^{-45}$
24	$-\cdot1548215174 \times 10^{-49}$	$-\cdot1936645922 \times 10^{-49}$	$-\cdot2325627706 \times 10^{-49}$
26	$-\cdot5630085021 \times 10^{-53}$	$-\cdot7038130639 \times 10^{-53}$	$-\cdot8446384950 \times 10^{-53}$

Table (4.4) Aspheric polynomials for corrector plates

J	a_{J1}
4	$-\cdot 3665201591 \times 10^{-6}$
6	$-\cdot 7332518637 \times 10^{-10}$
8	$-\cdot 1290813998 \times 10^{-13}$
10	$-\cdot 2303117374 \times 10^{-17}$
12	$-\cdot 4399446887 \times 10^{-21}$
14	$-\cdot 9030876176 \times 10^{-25}$
16	$-\cdot 1948082671 \times 10^{-28}$
18	$-\cdot 4323421915 \times 10^{-32}$
20	$-\cdot 9765015439 \times 10^{-36}$
22	$-\cdot 2236125061 \times 10^{-39}$
24	$-\cdot 5184217020 \times 10^{-43}$
26	$-\cdot 1215266759 \times 10^{-46}$

Table (4.5) Aspheric coefficients required for the singlet in Fig.(4.3) in order to uniformly illuminate the target.
R_T=50μm, ϕ_m=-90°.

This does mean however, that there is no point in carrying out the testing at the awkward laser wavelength. A convenient wavelength in the visible is just as suitable.

CHAPTER 5

DESIGNS CONTAINING A SINGLE ASPHERIC (II): WAVELENGTH TUNABLE LENSES

5.1. Introduction

The principle behind the wavelength tunable lens is just a small step away from the doublet described in section 4.7. of the previous chapter. The manufacture of the aspheric surface is greatly simplified if it can be null tested at a visible spectral line. In doing this, of course, the lens will not be corrected at the laser wavelength, taken here to be 1·064μm. To recorrect in the infrared, the component spacing is adjusted and as can be seen from the example mentioned above, the illumination can be optimized by a slight further change to the spacing. It appears that the spacing required to optimize the illumination differs by only a fraction of a millimeter from the value required to correct the aberrations. The latter is usually a good starting point in searching for the ideal spacing.

Any change in power, back focal length or aberrations with wavelength is unimportant and correction for chromatic aberration or spherochromatism are not under consideration. What we are interested in doing is to recorrect spherical aberration at a new wavelength by using a degree of freedom; the element spacing.

The example of such a lens given by Nicholas, Pataky and Welford (17) was for use at the Rutherford Laboratory with a two beam Neodymium glass laser. The aspheric surface

is figured until the doublet gives a null pattern in a Twyman-Green interferometer using a mercury lamp line at 0·546μm. Decreasing the element spacing then recorrects the doublet at 1·064μm. In fact, for any wavelength over the visible and near infrared, correction can be achieved simply by changing the separation by a calculated amount.

Their design consisted of two plano-convex elements. These provide a flatter principal surface than a sphere, so the lens was not aplanatic. For a uniform beam, the effect of this is an increase of intensity with incident ray height over the converging wavefront. This counteracts the drop in the Gaussian beam actually used.

Theory concerning the behaviour of the target predicts that radiation absorption and compression processes become more efficient at shorter wavelengths. As a result experiments with frequency doubled (and tripled in some cases) outputs are being carried out at many laboratories, and appear to confirm the theory (see H.G. Ahlstron (20) for instance: a complete list of other references is included). The improved interactions may be worth the energy loss (about 30-50%) due to the harmonic generation.

Using a Nd: glass laser, outputs at 0·532μm and 0·355μm are generated as the second and third harmonics respectively. Therefore the use of wavelength tunable lenses at these harmonics would be advantageous. A completely new set of focussing optics would be unnecessary. The element spacing would be set to the calculated value and the whole lens shifted slightly to compensate for the back focal length

change. Bearing in mind the disadvantages of a purely reflecting system mentioned in Chapter 1, tunable refracting optics would certainly be of some use.

5.2. Recorrecting spherical aberration

Although the merit of a particular design is mainly on the target illumination, in order to be of any use at all the system must be reasonably well corrected for spherical aberration. In an extreme case, a lens would be useless if its image was so badly aberrated that it did not even focus the beam down to a size as small as the target. In practice the correction should be considerably better than this. Aberrated images contain caustic surfaces, one or two of which will be located about the focus. If a caustic intersects the target the illumination will certainly not be uniform. The aberrations must therefore be small enough to prevent this happening to a correctly positioned target. The procedure for "tuning" the system to the operational wavelength is therefore to correct it as well as possible and then relax the correction if the illumination can be improved.

The single aspheric surface is to provide complete correction at a visible wavelength. Recorrecting at the laser wavelength must therefore involve a balancing of the aberration orders. If the aspheric was generated by the method described, then these are already at hand to assess the state of correction.

The transverse ray aberration for any defocussed image plane is given by,

$$T = (T_\zeta^{(1)}\rho + T_\zeta^{(3)}\rho^3 + T_\zeta^{(5)}\rho^5 + \ldots)\zeta + (T^{(3)}\rho^3 + T^{(5)}\rho^5 + \ldots) \tag{5.1}$$

$$T = T_\zeta^{(1)}\rho\zeta + (T_\zeta^{(3)}\zeta + T^{(3)})\rho^3 + (T_\zeta^{(3)}\zeta + T^{(5)})\rho^5 + \ldots \tag{5.2}$$

where the $T_\zeta^{(J)}$ are the defocussing terms. We write T_ζ as $(T_\zeta^{(1)}\rho + T_\zeta^{(3)}\rho^3 + T_\zeta^{(5)}\rho^5 + \ldots)\zeta$. Also,

$$T_\zeta = \frac{M_K}{N_K}\zeta \tag{5.3}$$

or

$$N_K T_\zeta = M_K \zeta \tag{5.4}$$

This will give,

$$T_\zeta^{(q)} = M_K^{(q)} - \sum_{J=1,3,\ldots}^{q-2} T_\zeta^{(J)} N_K^{(q-J)} \tag{5.5}$$

and $T_\zeta^{(1)} = M_K^{(1)}$

The presence of any higher order aberrations $T_\zeta^{(J)}\zeta + T^{(J)}$ $(J=5,7,\ldots)$ can be balanced out by the presence of a related amount of third order aberration, $T_\zeta^{(3)}\zeta + T^{(3)}$. In all the common image assessing methods (e.g. the variance of the wavefront aberration, longitudinal aberration for marginal zone) these aberrations are just proportional to one another. For the best correction,

$$T_\zeta^{(3)}\zeta + T^{(3)} = K_{3J}(T_\zeta^{(J)}\zeta + T^{(J)}) \tag{5.7}$$

where K_{3J} is a constant.

For all higher orders, there is a summation on the right-hand side,

$$T_\zeta^{(3)}\zeta+T^{(3)} = K_{35}(T_\zeta^{(5)}\zeta+T^{(5)}) + K_{37}(T_\zeta^{(7)}\zeta+T^{(7)})+.. \quad (5.8)$$

$$= \sum_{J=5,7,..} K_{3J}(T_\zeta^{(J)}\zeta+T^{(J)}) \quad (5.9)$$

For instance, suppose the system is judged to be corrected when the marginal ray passes through the paraxial focus ($\zeta=0$). Then

$$T = T^{(3)} + T^{(5)} + T^{(7)} +...= 0 \quad (5.10)$$

$$\text{or} \quad T^{(3)} = -T^{(5)} - T^{(7)} - ... \quad (5.11)$$

Thus $K_{35}, K_{37},.. = -1$. The condition used to decide whether the system is corrected, determines the K_{3J}. Our degree of freedom, the element spacing, is used to satisfy the relation (5.9).

Since defocus is also a degree of freedom there is another relation similar to (5.9).

$$T_\zeta^{(1)}\zeta = K_{15}(T_\zeta^{(5)}\zeta+T^{(5)}) + K_{17}(T_\zeta^{(7)}\zeta+T^{(7)})+... \quad (5.12)$$

The optimum defocus is therefore,

$$\zeta = \frac{(K_{15}T^{(5)}+K_{17}T^{(7)}+...)}{(T_\zeta^{(1)}-K_{15}T_\zeta^{(5)}-K_{17}T_\zeta^{(7)}-...)} \quad (5.13)$$

This is the value of ζ to be used in equation (5.9). Table (5.1) gives the K_{3J} and K_{1J} for several possible criteria. Obviously the K_{3J} have to be negative.

5.3. Examples and discussion

The properties of a wavelength tunable lens are not

	$LA_m = T_m = 0$		Minimum Var(W)		Minimum image dia.	
J	K_{1J}	K_{3J}	K_{1J}	K_{3J}	K_{1J}	K_{3J}
5	0•0	-1•0	0•2	-1•0	0•31250	-1•25
7	0•0	-1•0	0•22857	-0•85714	0•48517	-1•37360
9	0•0	-1•0	0•21429	-0•71429	0•60175	-1•45254
11	0•0	-1•0	0•19048	-0•59523	0•68835	-1•50940
13	0•0	-1•0	0•16667	-0•5	0•75630	-1•55316
15	0•0	-1•0	0•14515	-0•42424	0•81155	-1•58826
17	0•0	-1•0	0•12727	-0•36364	0•85764	-1•61725
19	0•0	-1•0	0•11189	-0•31469	0•89684	-1•64171
21	0•0	-1•0	0•09890	-0•27473	0•93071	-1•66269
23	0•0	-1•0	0•08791	-0•24176	0•96033	-1•68094
25	0•0	-1•0	0•07857	-0•21429	0•98651	-1•69700

Table (5.1) Constants K_{1J} and K_{3J} for three criteria for judging the optimum correction.

amenable to analytic study. The most detailed to be attempted analytically is likely to be Seidel theory and this was found to be totally inadequate. For example, in some instances Seidel theory predicted that the spacing between the components should be altered in the opposite manner to that deduced from higher order considerations. As a result the following discussion is qualitative only.

The choice of visible wavelength for aspheric testing may be any spectral line, but bearing in mind that the laser wavelength is to be 1·064μm, then the closer they are together the better the correction in the infrared. Thus choosing the spectral line to be at the red end is likely to help. Here the cadmium line at 0·644μm was chosen.

By starting with the doublet illustrated in Fig. (4.4), we are able to derive from it the "equivalent" lens that is null corrected at 0·644μm. The starting system is corrected in the infrared when the spacing is set at 7·345mm. By increasing the spacing to 7·4135mm it is found that the variance of the wavefront error is minimized at 0·644μm. If the spacing is set to this value and the aspheric surface re-designed so that null correction is obtained at 0·644μm we arrive at this equivalent lens.

Both have the same curvatures and element thickness; the only difference is that the aspheric surface has been changed slightly. Fig.(5.1) shows the second lens together with its properties.

The meaning of "equivalent" lens extends a little further than the identical curvatures and thicknesses. In order to

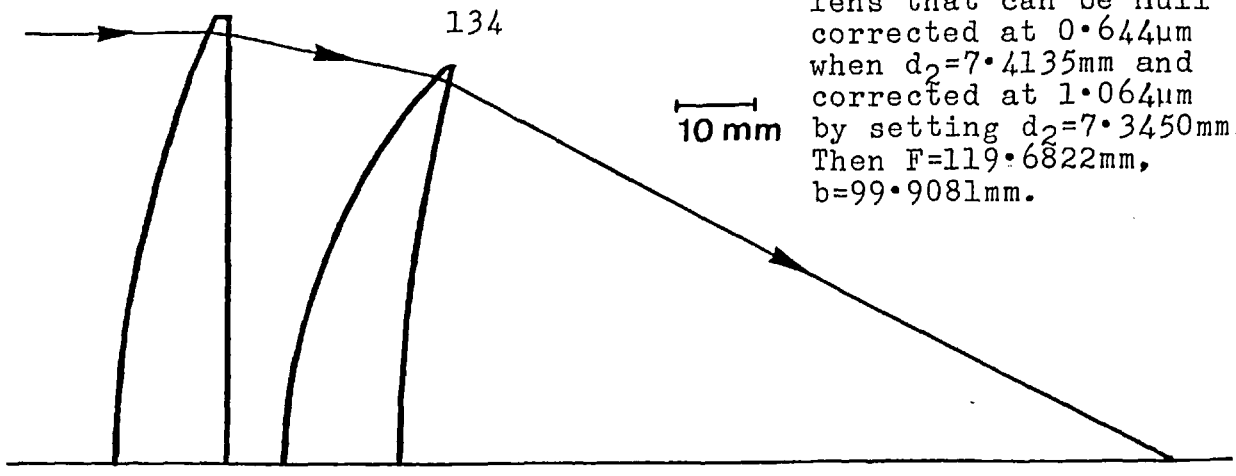
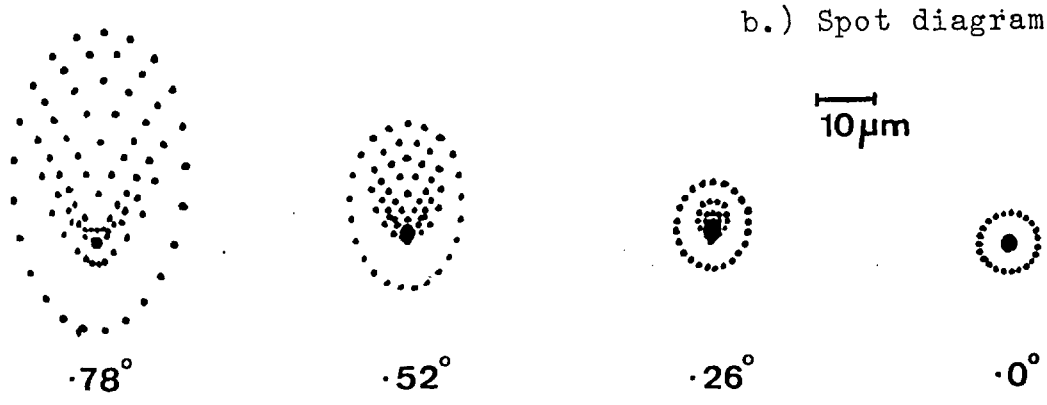


Fig. (5.1) a.) Tunable lens that can be null corrected at 0·644μm when d_2=7·4135mm and corrected at 1·064μm by setting d_2=7·3450mm. Then F=119·6822mm, b=99·9081mm.

b.) Spot diagrams.



c.) The element spacing required to correct the doublet to any wavelength.

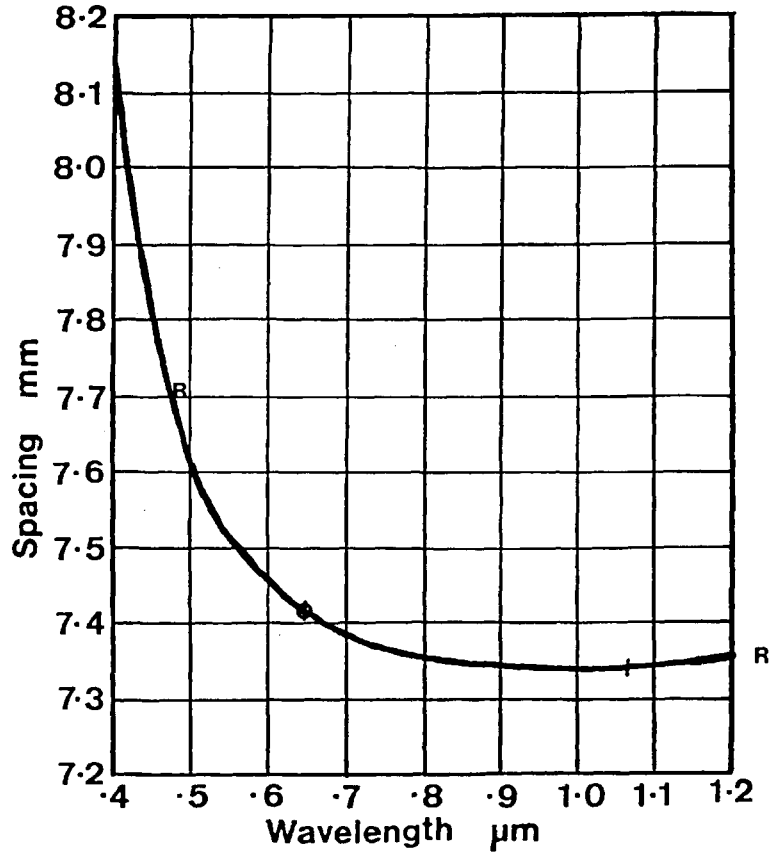
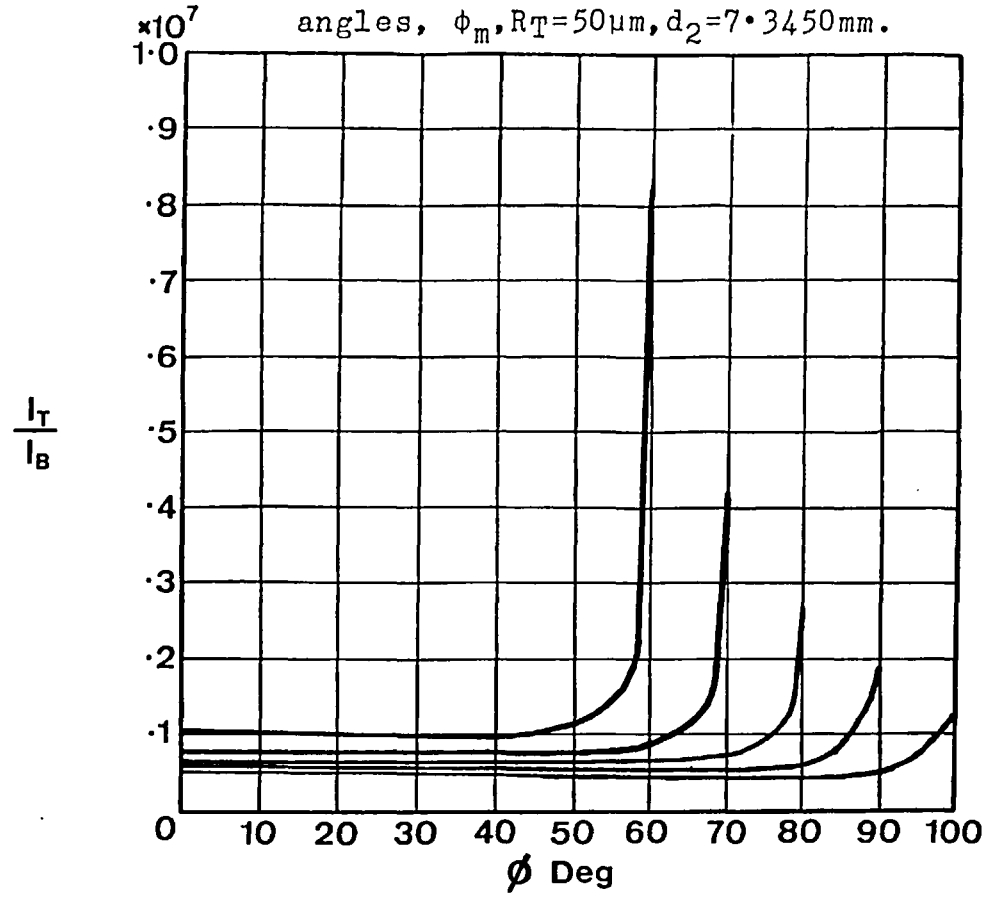
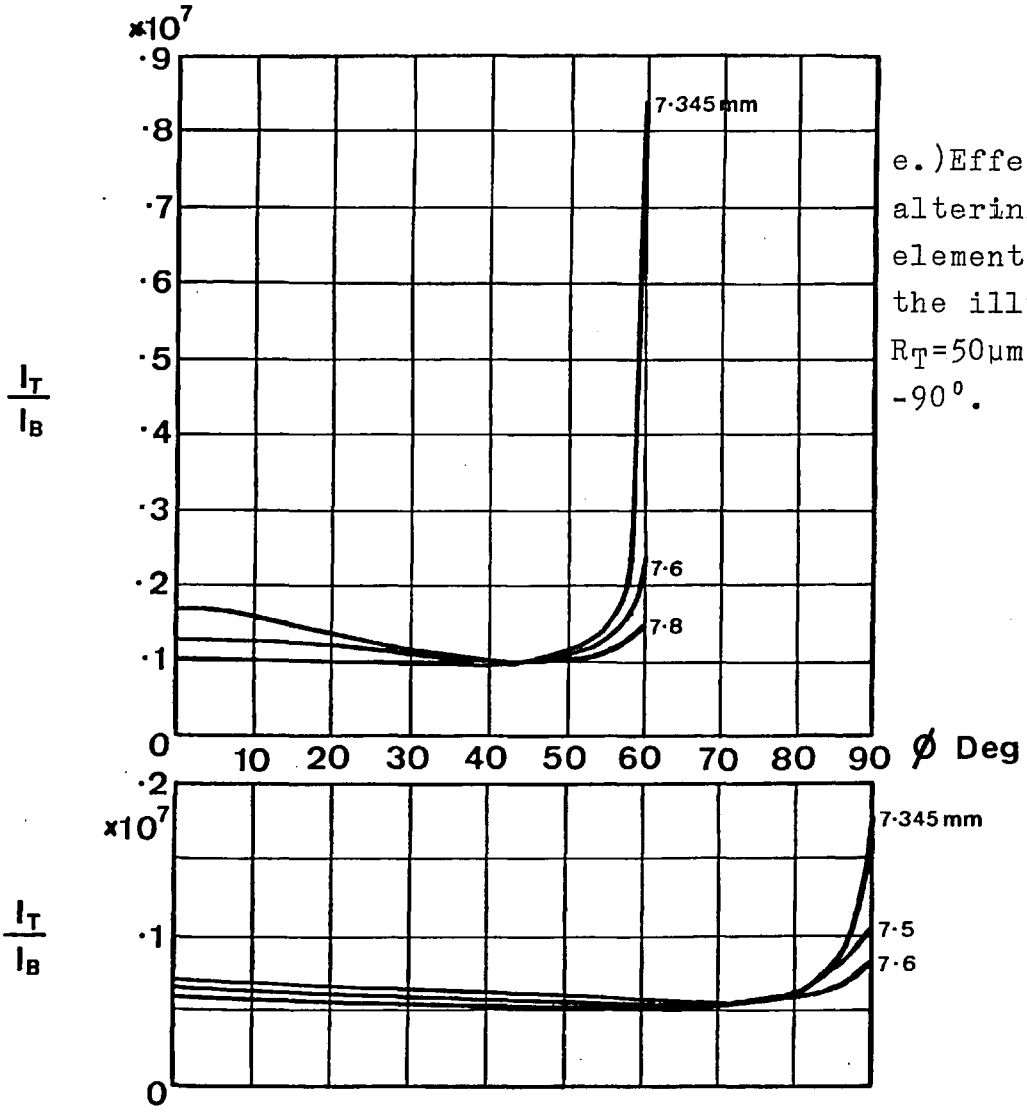


Fig. (5.1)

d.)Target illumination for various coverage angles, ϕ_m, R_T=50µm, d_2=7·3450mm.





e.)Effect of altering the element spacing on the illumination. R_T=50µm, ϕ_m=-60° & -90°.

Fig. (5.1)

c_i (paraxial)	d_i
•00771010	
	14•5
•0	
	Variable
•01349528	
	14•5
•00455873	

J	a_{J1}
4	$-•6060217640 \times 10^{-7}$
6	$-•3083729049 \times 10^{-11}$
8	$-•1479624227 \times 10^{-15}$
10	$-•8385714997 \times 10^{-20}$
12	$-•5767743125 \times 10^{-24}$
14	$-•4584501494 \times 10^{-28}$
16	$-•4035362388 \times 10^{-32}$
18	$-•3818046123 \times 10^{-36}$
20	$-•3800922341 \times 10^{-40}$
22	$-•3924536318 \times 10^{-44}$
24	$-•4164049760 \times 10^{-48}$
26	$-•4512829225 \times 10^{-52}$

Table (5.2) System data for lens in fig. (5.1)

correct it in the infrared the spacing has to be altered. On doing this it is found that the required spacing is 7·345mm, the same as the original lens. In fact if we determine the spacing required to correct each lens at any wavelength we find that they are the same (to an accuracy of 0·1μm anyway). The slight change in the aspheric also has a negligible effect on the amount of coma present at each wavelength. Thus Fig.(5.1c) also applies to the original lens in Fig. (4.4). Referring to this graph, the circle, O, on the curve denotes the position at which the lens is null corrected. Here at 0·644μm and a spacing of 7·4135mm. Thus by changing the aspheric, the O can be slid along the curve without altering it as a whole. The section bounded by the R's show the region for which the wavefront correction is to the Rayleigh $\lambda/4$ limit or better. This extends from about 0·5 to beyond 1·2μm.

It will be noticed that this curve has a minimum at about the 1·0μm mark. This means that the lens has the unusual property of being tuned to two wavelengths simultaneously.

The tunable properties may be altered by bending one of the elements. Fig.(5.2) shows the effect of doing this to the second element (the first element would produce a similar result) each time changing the aspheric so complete correction at the same wavelength is achieved. This ensures that the curve always passes through one particular point. It can be seen that the optimum spacing can be made to increase with wavelength, decrease or contain a minimum. In the latter case,

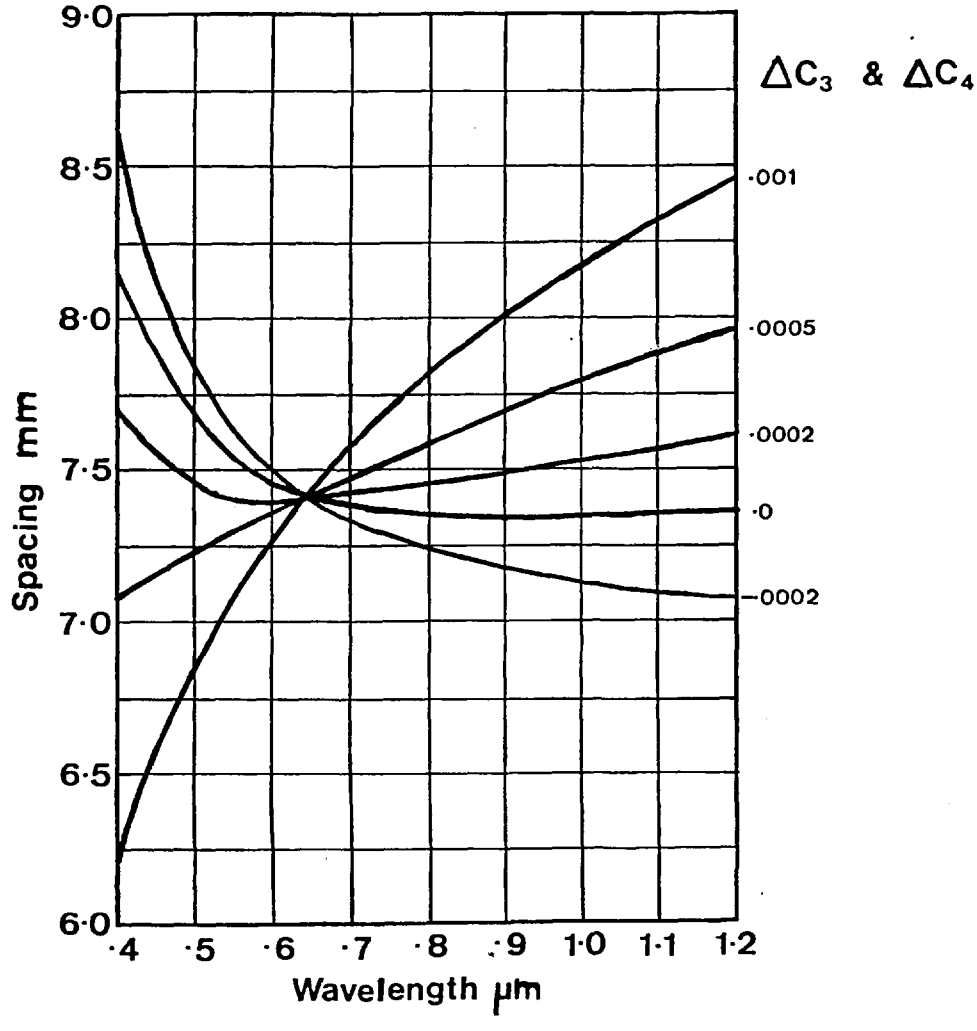


Fig. (5.2) The effect of bending the second element on the optimum spacing.

any second wavelength can be simultaneously corrected with the null-corrected wavelength.

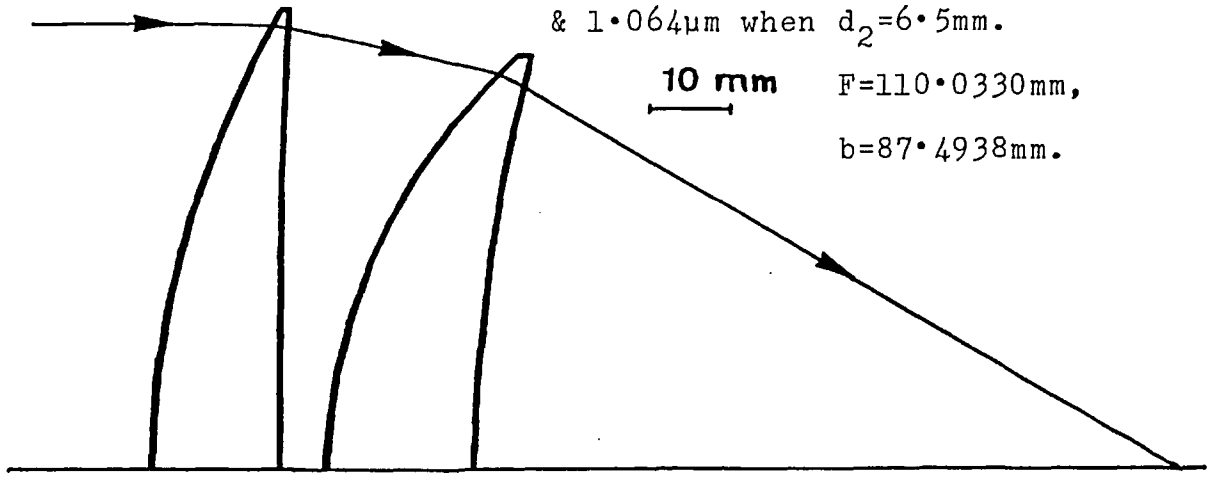
With the spacing set to 7·345mm, Fig. (5.1d) shows the target illumination for various coverage angles and R_T=50μm. It appears that I_T/I_B is very uniform except for the edge zone. The sudden rise is due to the close proximity of a caustic in the aberrated image. If this is intolerable then an overall improvement can be made by increasing the element spacing slightly. This is illustrated for the same target size and for coverages of 60° and 90° in Fig.(5.1e). In doing this the coma hardly changes; at least for changes of less than a millimeter. Thus coma correction can be preserved when optimizing the illumination.

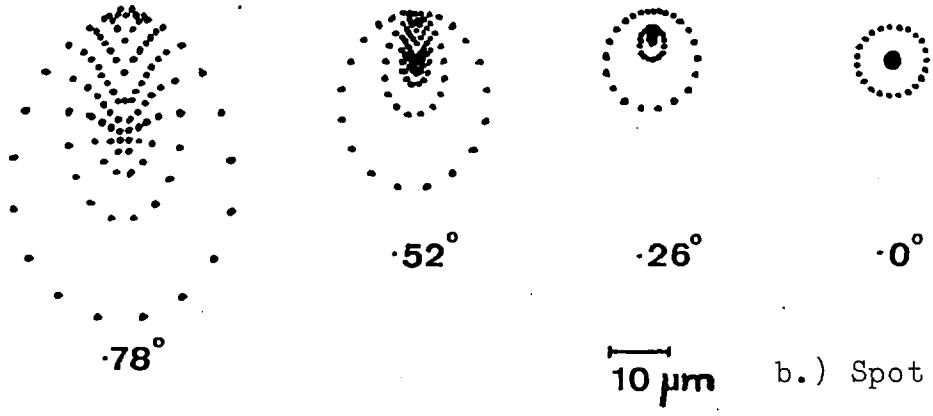
By making use of these observations the example in Fig.(5.3) was designed. Using an element shape as a degree of freedom and a spacing set at 6·5mm the lens can be null corrected at 0·644μm and simultaneously corrected at 1·064μm. By adjusting both element shapes OSC_m can be made to drop to zero in the infrared. Thus once the lens is null tested in the visible it is found to be aplanatic in the infrared.

The target illumination is similar to the previous example. Again an overall improvement can be made by adjusting the component spacing slightly. In both cases the uniformity is not so good as other designs. Here we have sacrificed uniformity for convenience in aspheric fabrication.

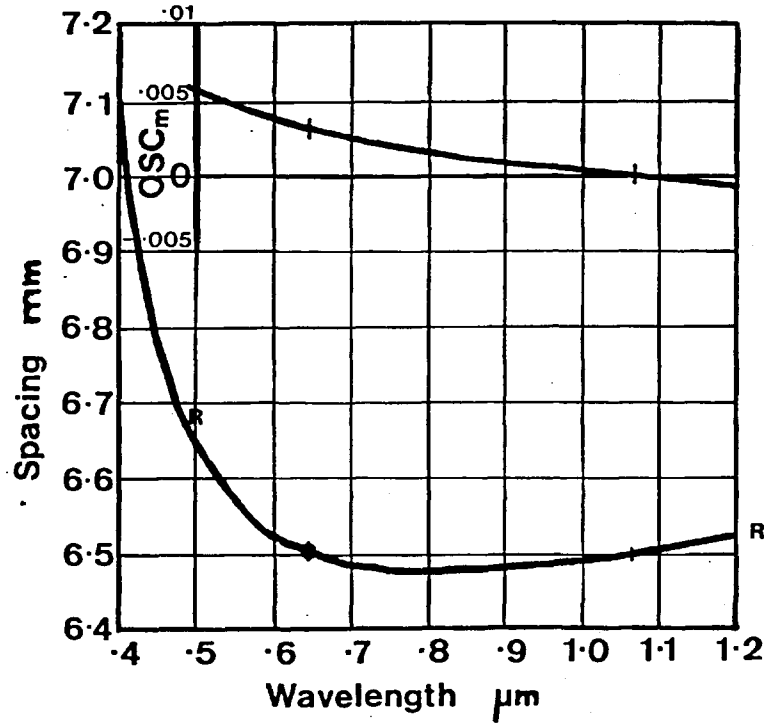
A third example is given in Fig.(5.4). Here the aspheric can again be null tested at 0·644μm with a 6·9414mm spacing but when increased to 7·0mm correction is obtained at both

Fig.(5.3) a.) Doublet that is simultaneously corrected at 0·644μm & 1·064μm when d_2=6·5mm. F=110·0330mm, b=87·4938mm.



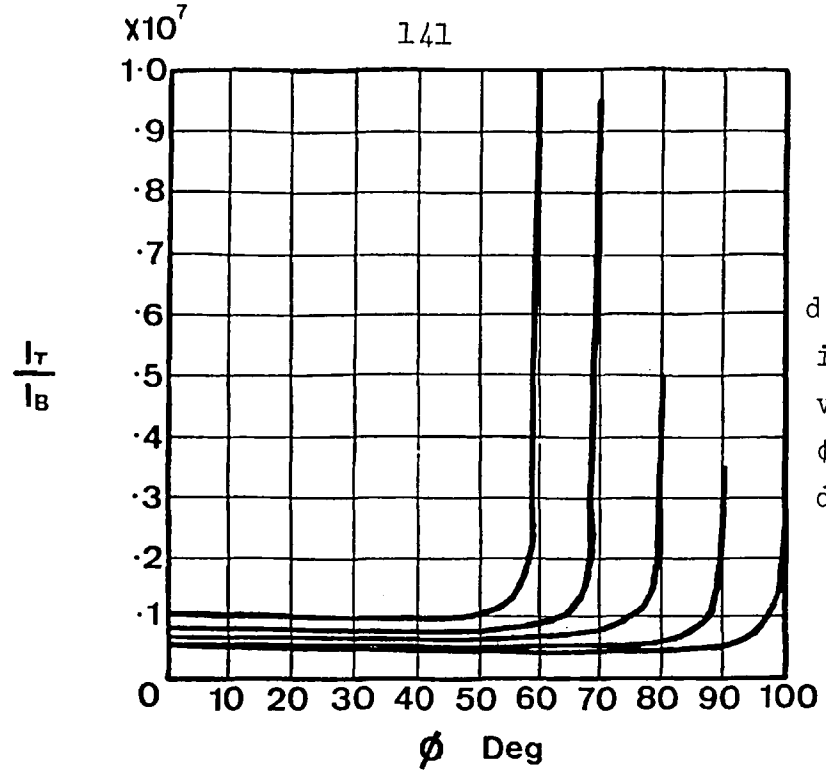


b.) Spot diagrams.

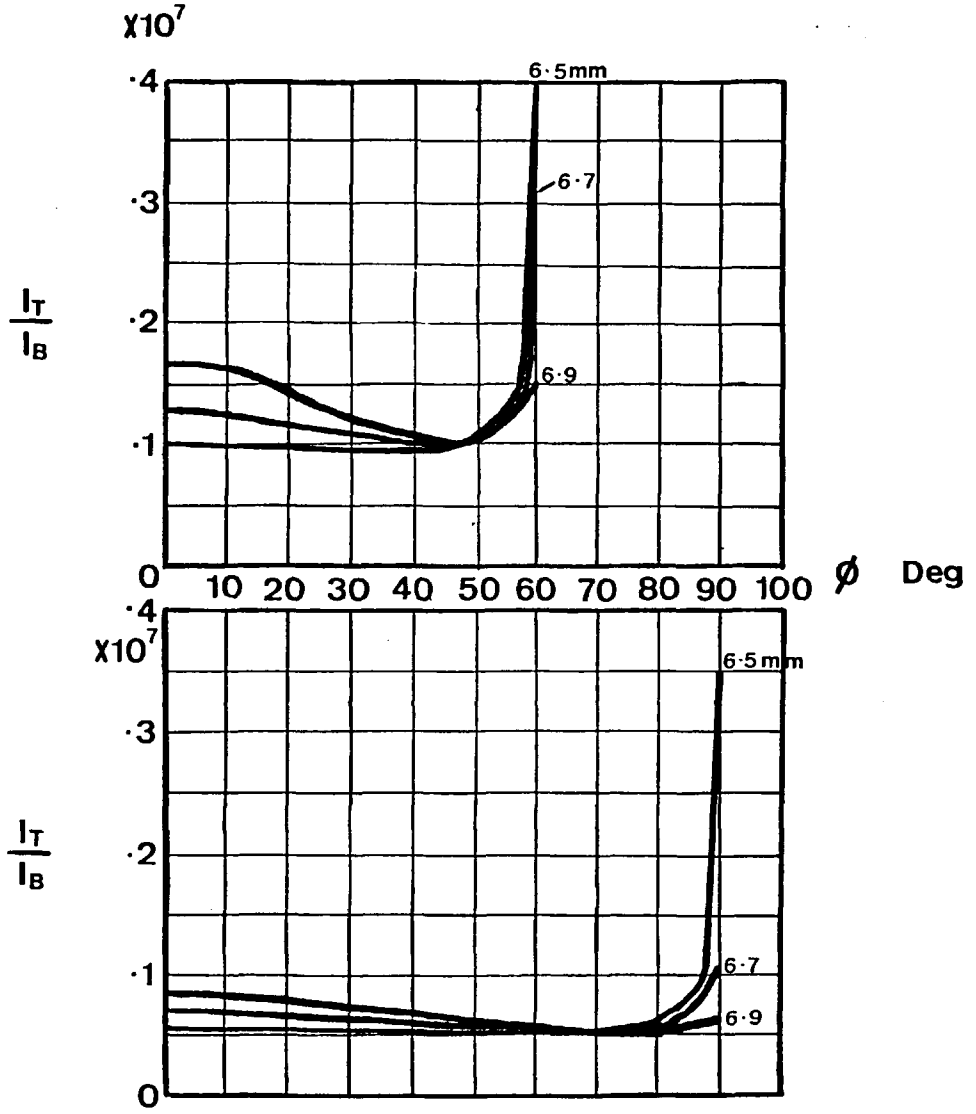


c.) Optimum element spacing for aberration correction and OSC_m when this spacing is set.

Fig. (5.3)



d.) Target illumination for various coverages, $\phi_m . R_T = 50\mu m$, $d_2 = 6 \cdot 5mm$.



e.) Effect of changing the spacing on the illumination. $R_T = 50\mu m, \phi_m = -60^0$ & -90^0.

Fig. (5.3)

c_i (paraxial)	d_i
•0092551543	
	15•0
•0006567903	
	Variable
•014528863	
	17•5
•005173892	

J	a_{J1}
4	$-•8212075722 \times 10^{-7}$
6	$-•5729408897 \times 10^{-11}$
8	$-•3665687028 \times 10^{-15}$
10	$-•2488529825 \times 10^{-19}$
12	$-•1858511418 \times 10^{-23}$
14	$-•1527868675 \times 10^{-27}$
16	$-•1371161376 \times 10^{-31}$
18	$-•1329052355 \times 10^{-35}$
20	$-•1374834533 \times 10^{-39}$
22	$-•1499596777 \times 10^{-43}$
24	$-•1705921013 \times 10^{-47}$
26	$-•2005519330 \times 10^{-51}$

Table (5.3) System data for the lens in fig. (5.3)

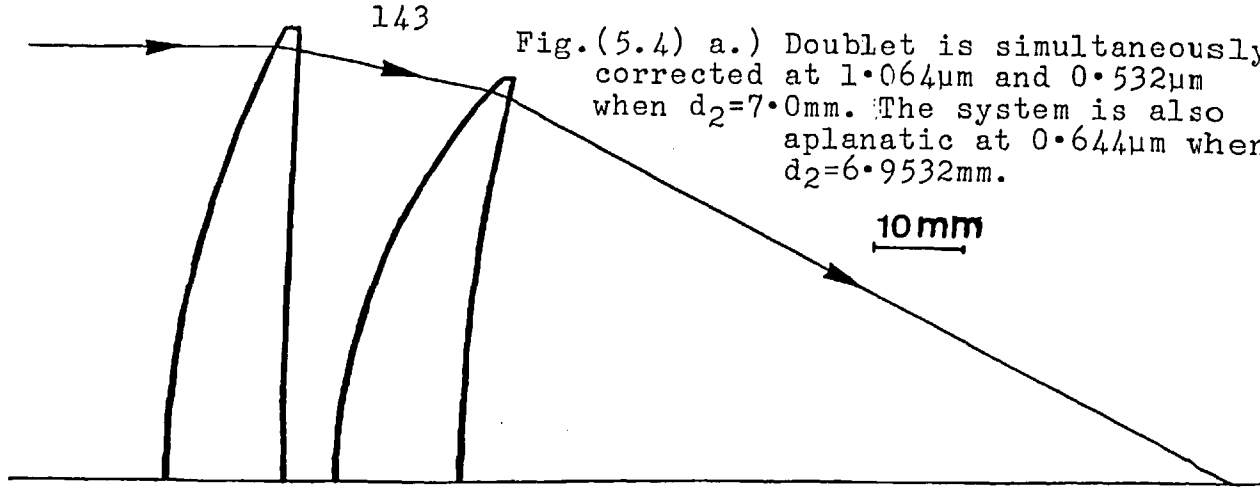
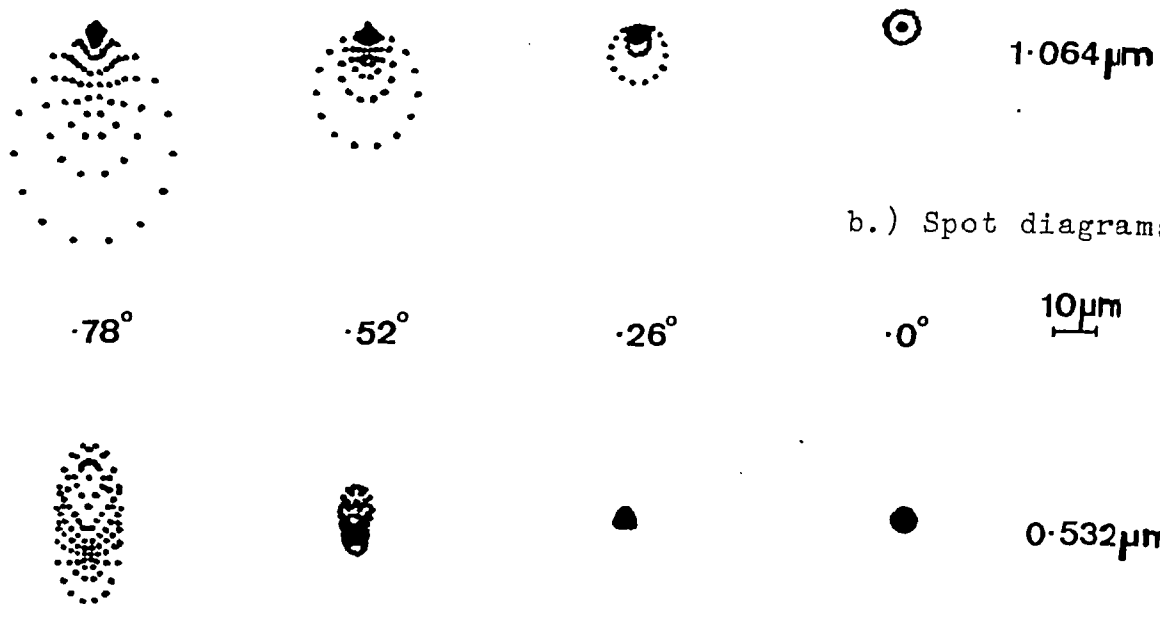
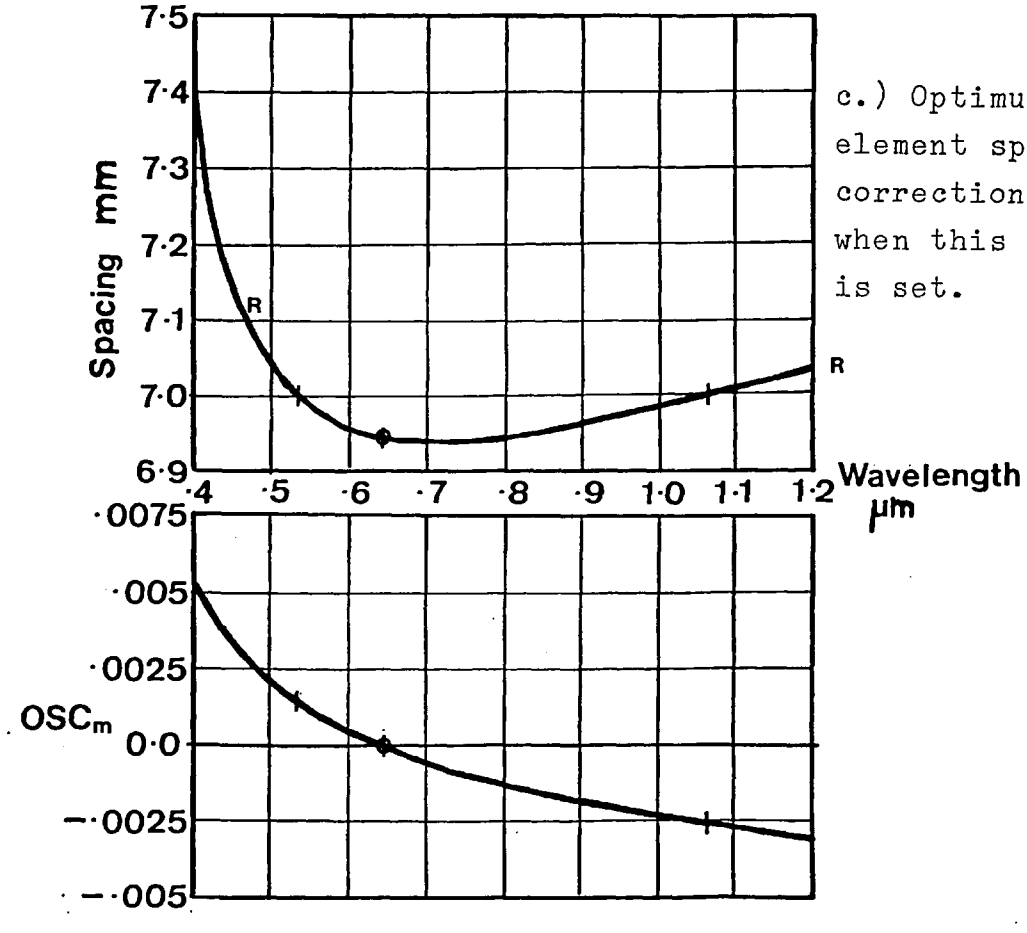


Fig.(5.4) a.) Doublet is simultaneously corrected at 1·064μm and 0·532μm when d_2=7·0mm. The system is also aplanatic at 0·644μm when d_2=6·9532mm.

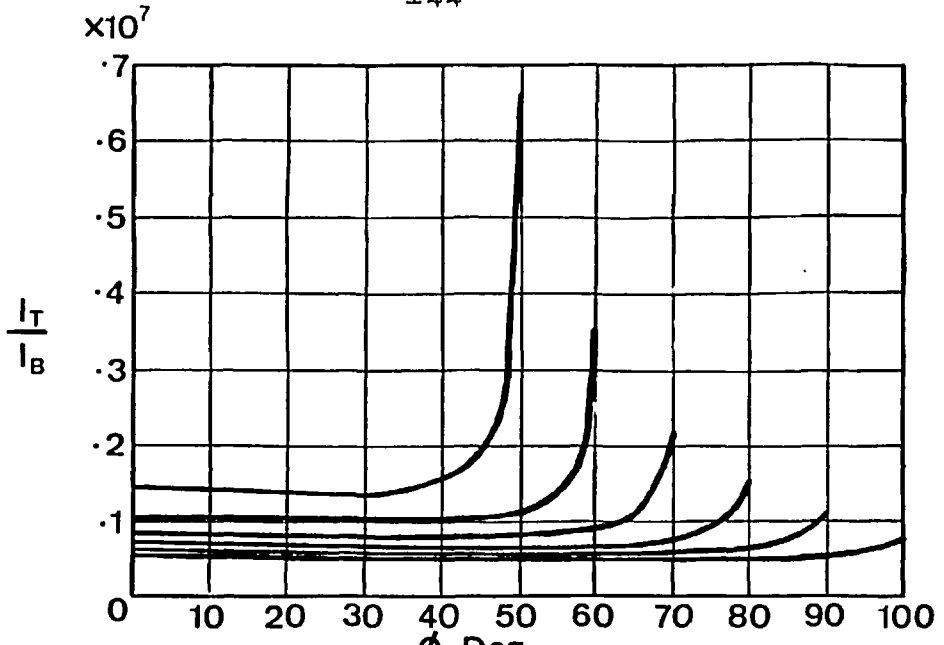


b.) Spot diagrams.

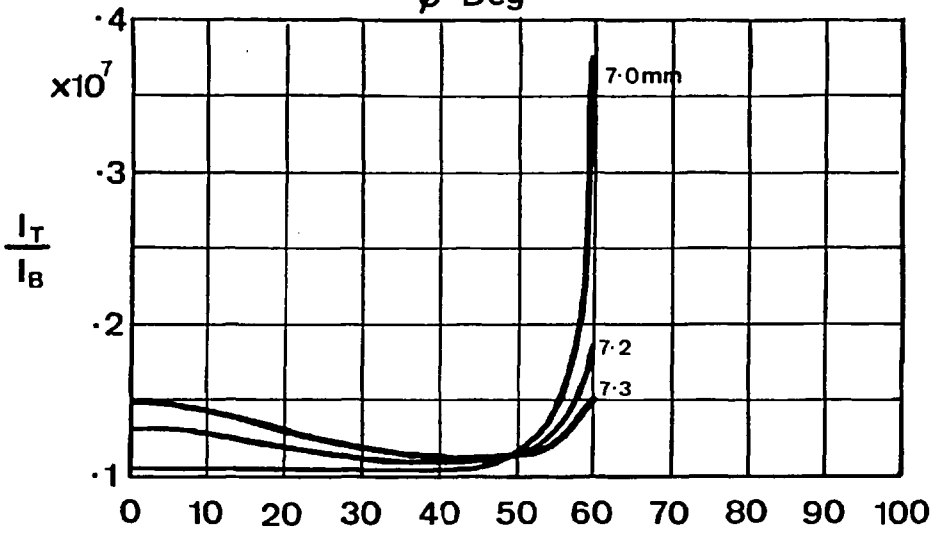


c.) Optimum element spacing for correction and OSC_m when this spacing is set.

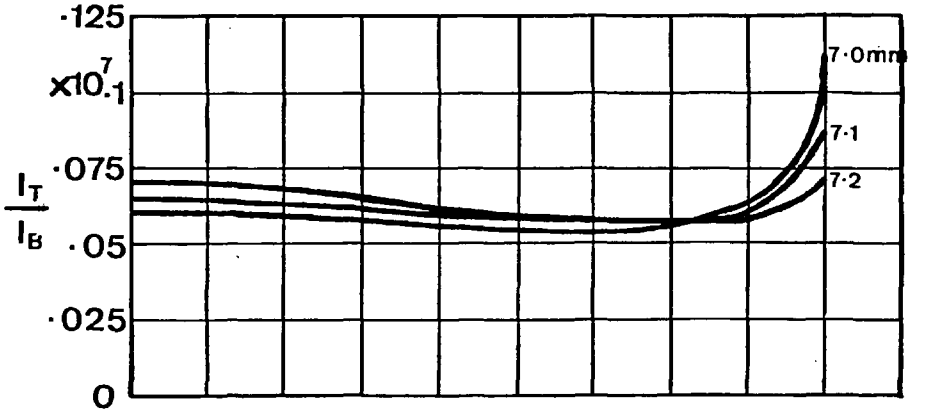
Fig. (5.4)



d.)



e.)

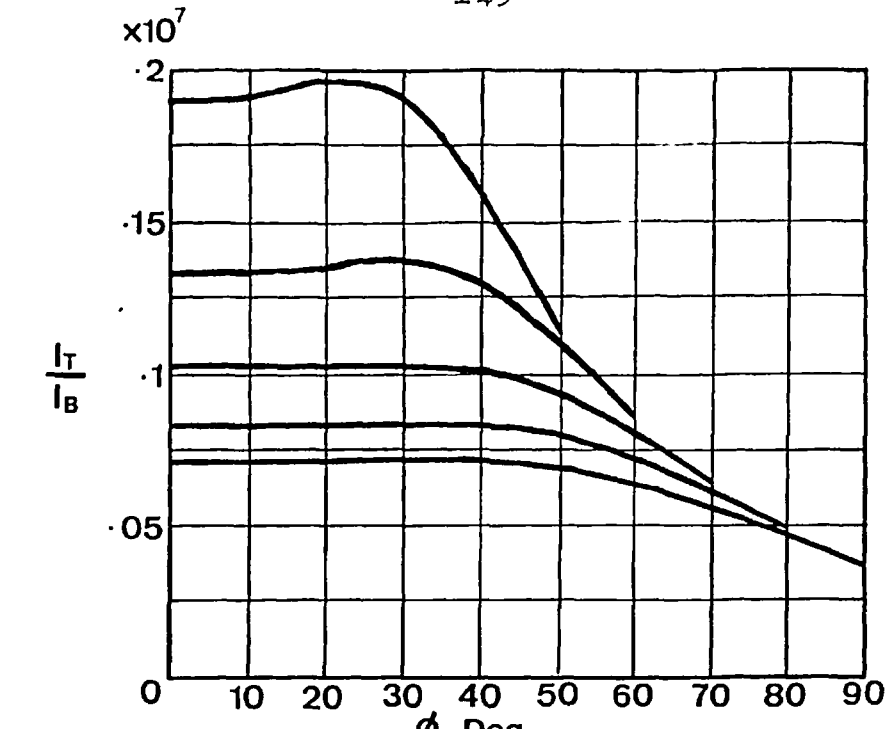


Illumination at 1·064μm: R_T=50μm.

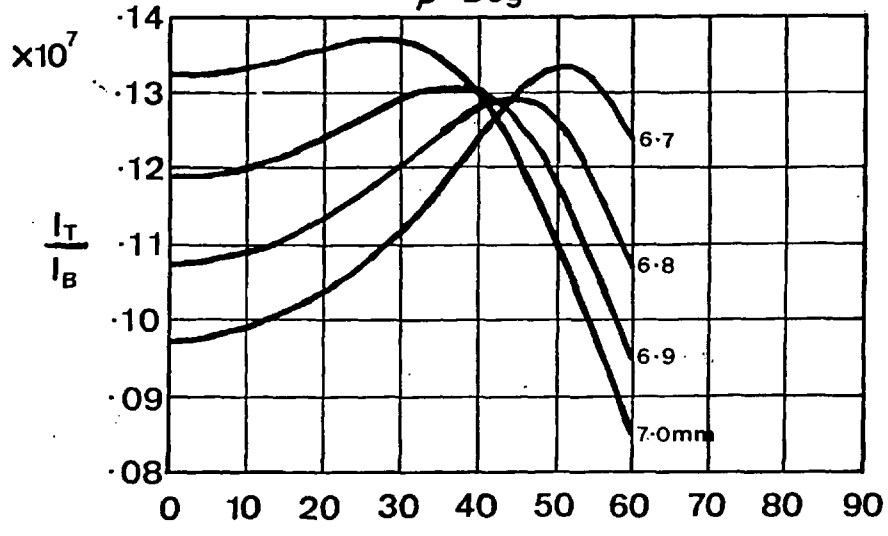
d.) Illumination for various coverages, ϕ_m.

e.) Effect of changing the spacing.
ϕ_m=-60° & -90°.

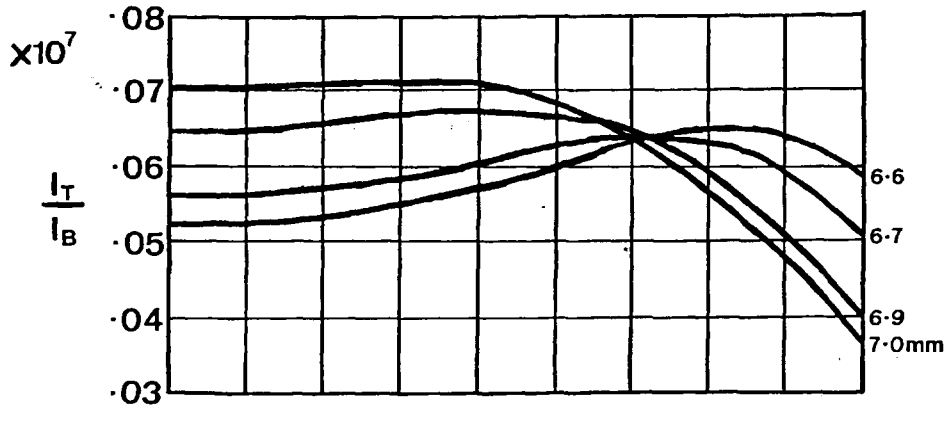
Fig. (5.4)



f.)



g.)



Illumination at 0·532μm: R_T=50μm.

f.) Illumination for various coverages, ϕ_m.

g.) Effect of changing the spacing.
ϕ_m=-60° & -90°.

Fig. (5.4)

c_i (paraxial)	d_i
•008632551	
	14•5
•000922451	
	Variable
•013376329	
	15•5
•004447709	

J	a_{J1}
4	$-•6699100727 \times 10^{-7}$
6	$-•4192364503 \times 10^{-11}$
8	$-•2422623623 \times 10^{-15}$
10	$-•1484293223 \times 10^{-19}$
12	$-•1001454291 \times 10^{-23}$
14	$-•7478792337 \times 10^{-28}$
16	$-•6134128914 \times 10^{-32}$
18	$-•5452123993 \times 10^{-36}$
20	$-•5172208289 \times 10^{-40}$
22	$-•5163268408 \times 10^{-44}$
24	$-•5360337168 \times 10^{-48}$
26	$-•5734781602 \times 10^{-52}$

Table (5.4) System data for lens in fig. (5.4)

1·064 and 0·532μm. Wavefront errors are less than λ/4 at both wavelengths. To control coma, the system was made aplanatic at 0·644μm. The spot diagrams show that coma correction at an intermediate wavelength is satisfactory; especially at 0·532μm.

The residual aberrations at the infrared and visible regions will be of opposite sign, producing the opposing illumination profiles in Figs. (5.4d) and (5.4f). They may be improved by corresponding opposite adjustments in the spacing; an increase for the infrared and a decrease for the visible.

5.4. Conclusions

With a single aspheric surface that can be conveniently tested at a visible wavelength, the tunable lens is likely to be the cheapest alternative for the focussing optics.

The lenses can be corrected to any wavelength over a range from blue-visible to infrared and remain diffraction limited over a fair part of it. Because of this, there is no reason why they cannot replace several sets of lenses, each originally designed for a single wavelength. Extending the range of use to third and fourth harmonics would no doubt be possible though U.V. transmitting material such as quartz or fused silica would have to replace the usual BK7 glass.

Although we can optimize the target illumination for any target size, coverage and incident beam intensity, the optimum solution does not give a uniformity (20-50% variations)

as good as previous examples. The fact that their properties are not easily investigated analytically means it would be very difficult to know whether this situation can be improved. If the disadvantages of a purely reflecting system can be tolerated, they could no doubt provide greater uniformity.

CHAPTER 6

DUAL-WAVELENGTH LENSES

6.1. Introduction

The increasing interest in using harmonics of the fundamental laser frequency means that an entirely new set of focussing optics would be necessary. At least, all the refracting components have to be replaced. The wavelength tunable lenses illustrate that it would be possible to replace a set of lenses by one doublet. The residual aberrations that exist at all the wavelengths except one, are responsible for the rather large non-uniformities in target intensity. Not surprisingly, if exact stigmatism is obtained the uniformity is generally better. Thus it appears that if we require good uniformity at the harmonics, the optics should be stigmatic at each or aberrated by the amount prescribed by equation (2.68),

$$\frac{T(\rho)}{R_T} = -2\rho\sin\left(\frac{\phi_m}{2}\right)\sqrt{1-\rho^2\sin^2\left(\frac{\phi_m}{2}\right)} \\ -2\left(\rho^2\sin\left(\frac{\phi_m}{2}\right)+\frac{F}{Y_m}\right)\sin\left(\frac{\phi_m}{2}\right)\tan U_K \qquad (6.1)$$

Using one aspheric surface it is possible to satisfy the constant optical path condition or to produce a selectively aberrated system at a single wavelength. Using two aspheric surfaces Chapter 3 has showed that it is possible to achieve exact axial stigmatism and obtain a specified principal surface, also at a single wavelength. Alternatively, two

aspheric surfaces can be used to satisfy the constant optical path condition at two wavelengths. In general, N aspherics could correct a system at N wavelengths. Here though we shall restrict ourselves to just two.

Therefore, consider a doublet whose first and third surfaces are aspheric, and are designed to satisfy the constant optical path condition at the wavelengths 1·064 and 0·532μm. Such an arrangement should be compared with two singlets, one for infrared use and the other for its harmonic. Those of the type illustrated in Fig.(4.3) for example. In each case there are two elements and two aspheric surfaces, so the cost of manufacture is likely to be about the same. In the singlet case, one set of lenses will have to be removed from the target chamber and replaced by the other set. For the doublet only a slight movement of the lens as a whole is required. The trouble of beam alignment has been substantially reduced for each change-over.

Another advantage is that the doublet can be optimized to any target size, coverage and beam profile in exactly the same manner as the tunable lenses. For a singlet, there is no degree of freedom to do this. Of course each singlet can be replaced by a doublet but then four elements have to be made.

Thus by effectively combining a pair of singlets to form a doublet that is stigmatic at two wavelengths, we achieve a system with the tunable advantages of a doublet while being cost-wise, the same as a pair of singlets.

The problem of designing such a pair of aspheric surfaces

by the point-characteristic method or most of the other common techniques would be extremely difficult. Although a ray traced at two wavelengths would fall on the first surface at the same point, after passing through the first element it will not hit the other surfaces at the same place. The problem is complicated by the fact that a set of three or four points and gradients are interrelated.

The method described in Chapter 4, however, can be easily extended to cope with this problem. In fact it allows the system to be selectively aberrated at one or both wavelengths. As in the case for designing a single aspheric, it is applicable to any pair of surfaces.

6.2. Correcting spherical aberration at two wavelengths by the use of two aspheric surfaces

Referring back to equation (4.65), the contribution to $n_K M_K^{(1)} T^{(q)}$ from the terms a_{q+1i} is,

$$\sum_{i=1,2,.}^{K} (q+1) a_{q+1i} (n_i - n_{i-1}) Y_i^{(1)q+1} \qquad (6.2)$$

Suppose the two surfaces are the ith and Jth. Then if $T_{req}^{(q)}$ is the required value of $T^{(q)}$ for the whole system at wavelength λ, we can write,

$$n_K M_K^{(1)} T^{(q)} + (q+1) a_{q+1i} (n_i - n_{i-1}) Y_i^{(1)q+1} + (q+1) a_{q+1J} (n_J - n_{J-1}) Y_J^{(1)q+1} = n_K M_K^{(1)} T_{req}^{(q)} \qquad (6.3)$$

or

$$a_{q+1i}(n_i-n_{i-1})Y_i^{(1)q+1}+a_{q+1J}(n_J-n_{J-1})Y_J^{(1)q+1}$$

$$= \frac{n_K M_K^{(1)}(T_{req}^{(q)} - T^{(q)})}{(q+1)} \qquad (6.4)$$

For the second wavelength, λ^*, there is a similar expression,

$$a_{q+1i}(n_i^*-n_{i-1}^*)Y_i^{(1)*q+1} + a_{q+1J}(n_J^*-n_{J-1}^*)Y_J^{(1)*q+1}$$

$$= \frac{n_K^* M_K^{(1)*}(T_{req}^{(q)*} - T^{(q)*})}{(q+1)} \qquad (6.5)$$

Equations (6.4) and (6.5) can therefore be simultaneously solved for a_{q+1i} and a_{q+1J}. This apparently difficult problem is therefore reduced to a simple extension of the single surface method used. Its generalization to any number of surfaces and wavelengths is obvious.

A flow chart for the process is shown in Fig. (6.1).

6.3. Design examples

Fig. (6.2) shows the first example of such a doublet. With the spacing set at 5·0mm the constant optical path condition is obeyed for radiation at 0·532μm and 1·064μm. By adjusting the spacing to the value shown in Fig. (6.2c) the diffraction limited region can be made to extend over a very wide region.

In principle the two component shapes could be used to correct for coma at both wavelengths or by bending say the whole system, at an intermediate wavelength. Seidel theory

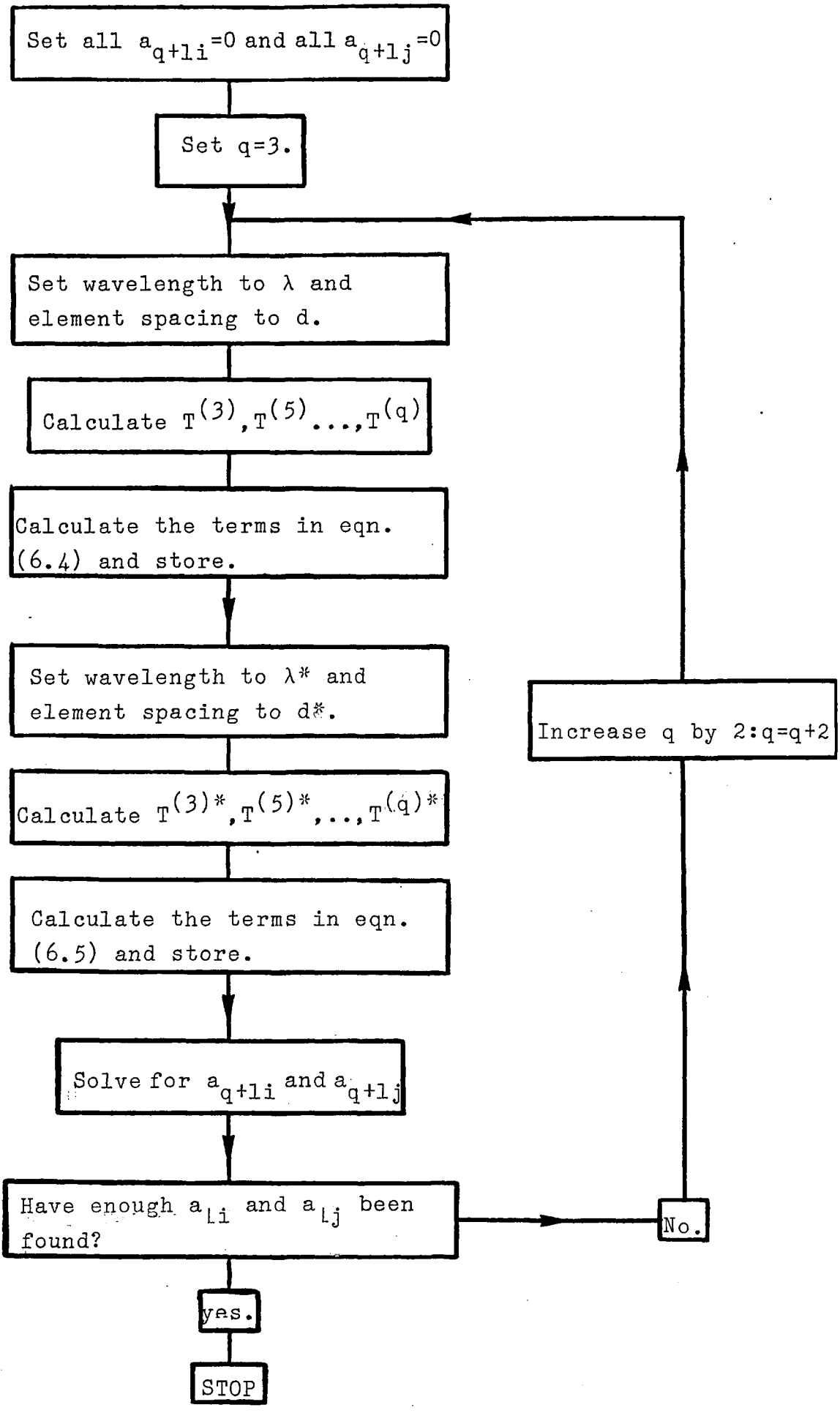


Fig.(6.1) Flowchart for generating two aspheric surfaces for correction at two wavelengths.

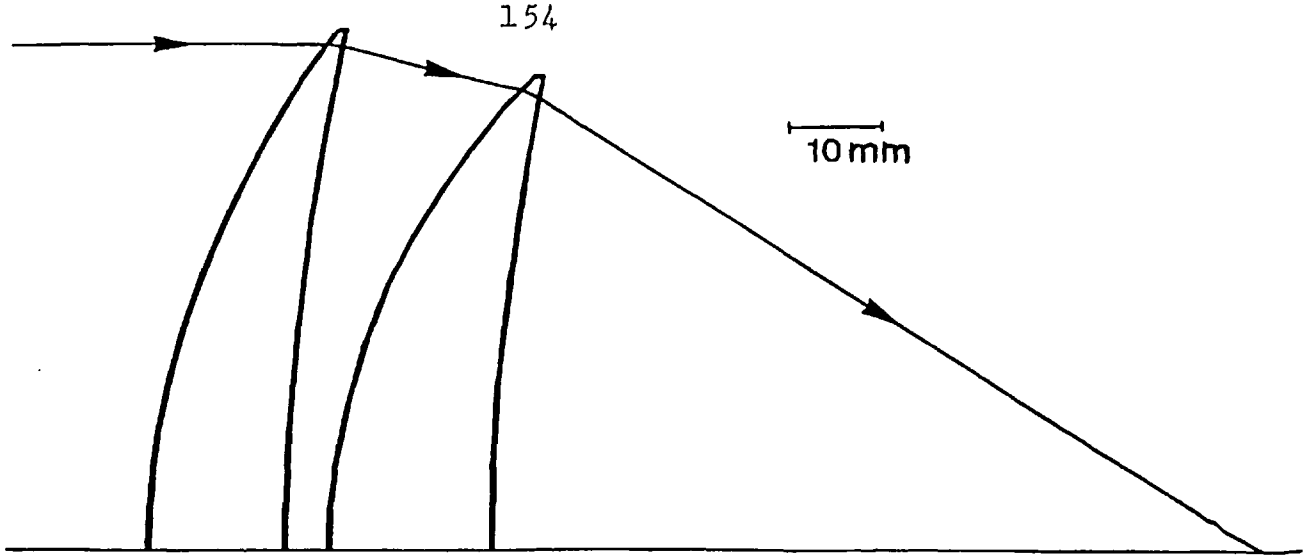
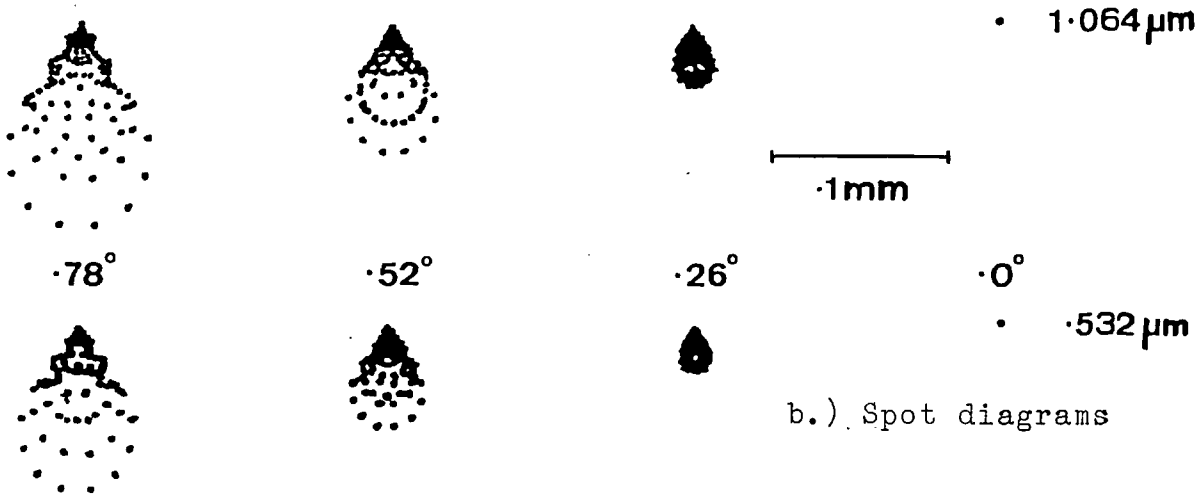
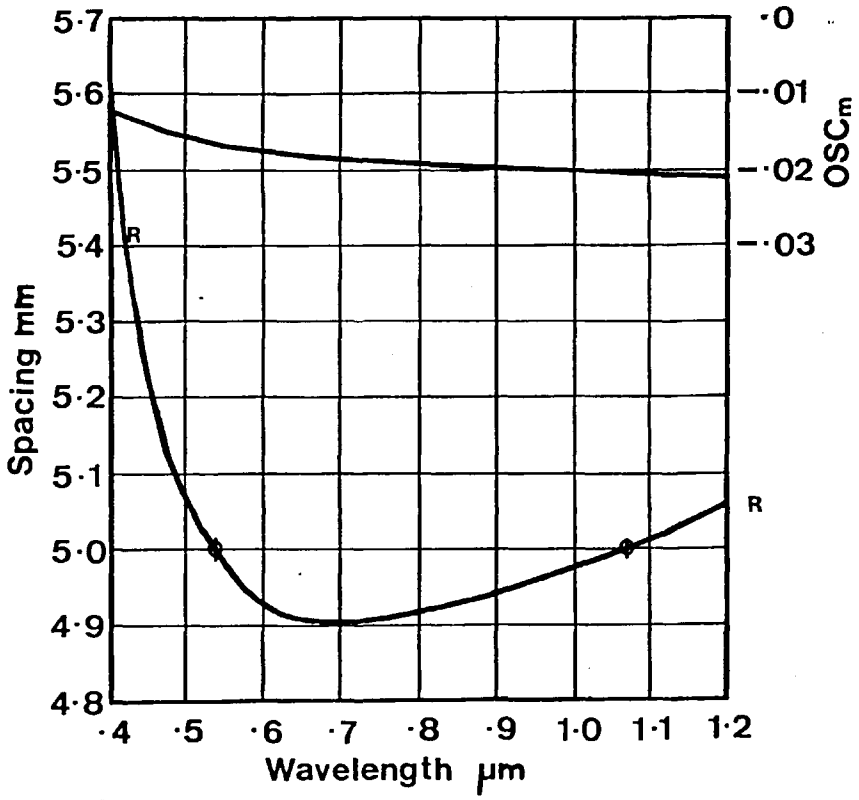


Fig. (6.2) a.) F/1 lens corrected at 1·064μm and 0·532μm.

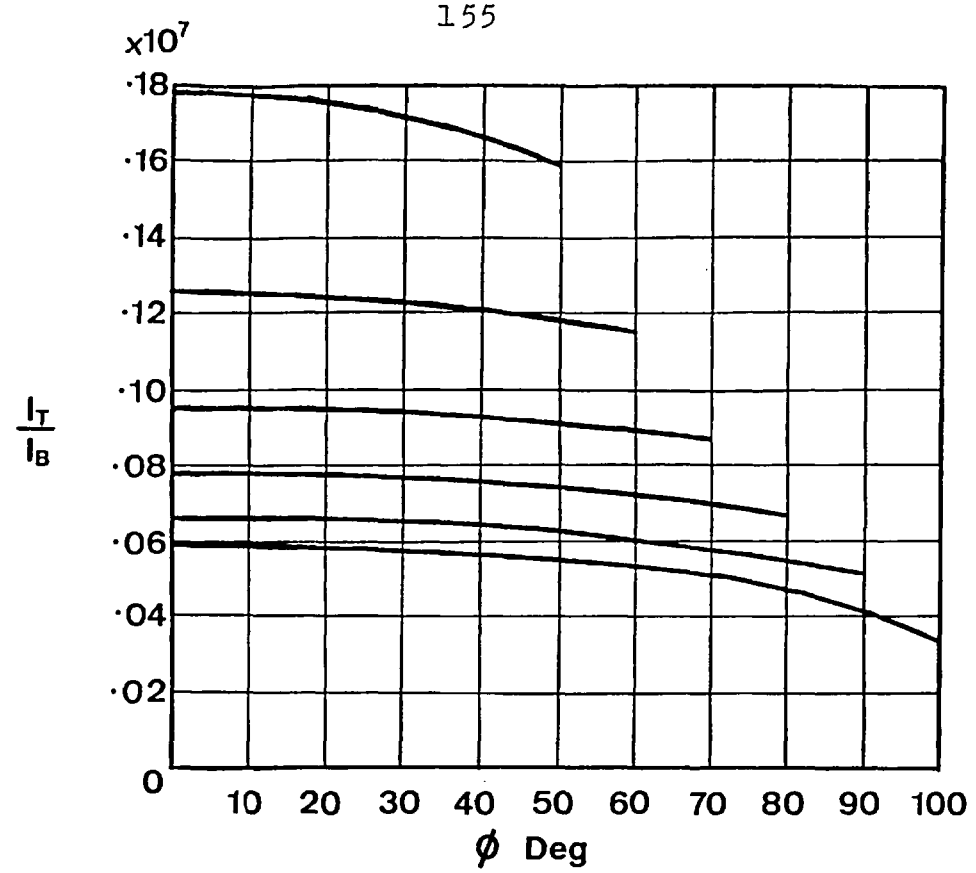


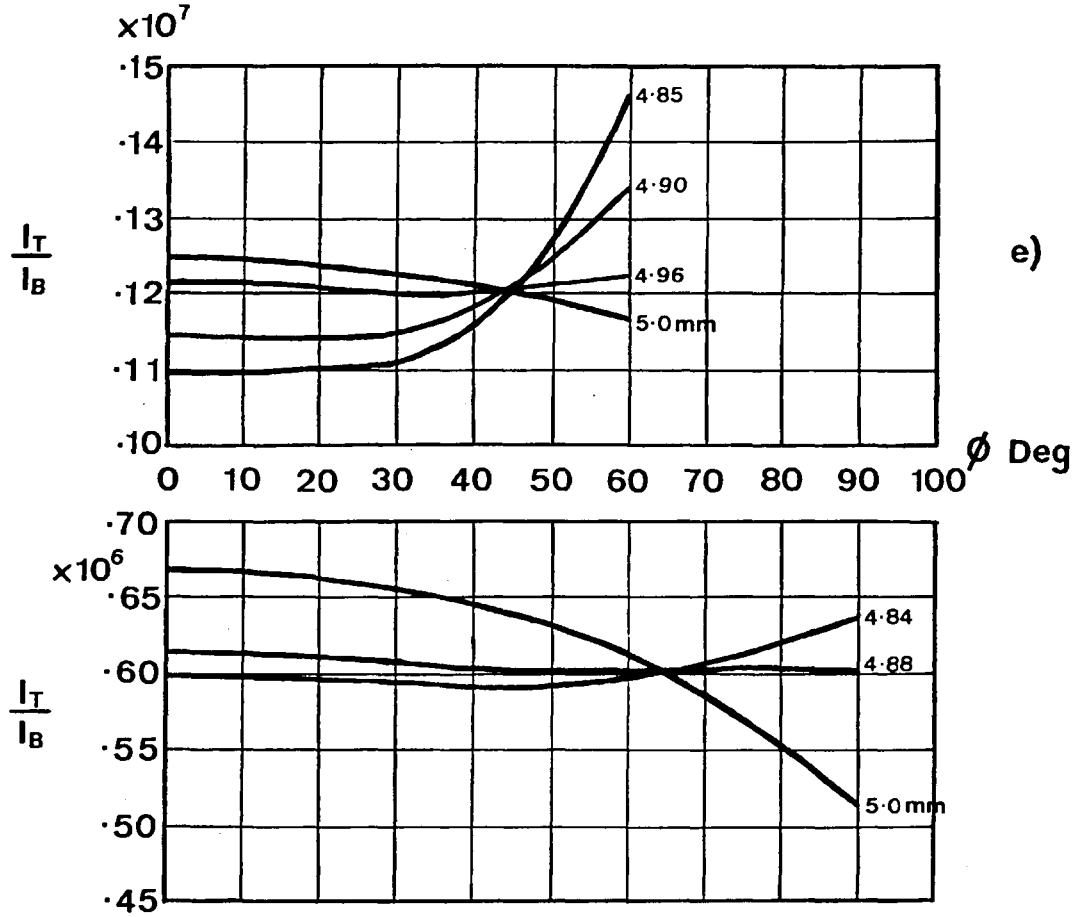
b.) Spot diagrams



c.) Element spacing for aberration correction and OSC_m when this is set.

Fig. (6.2)



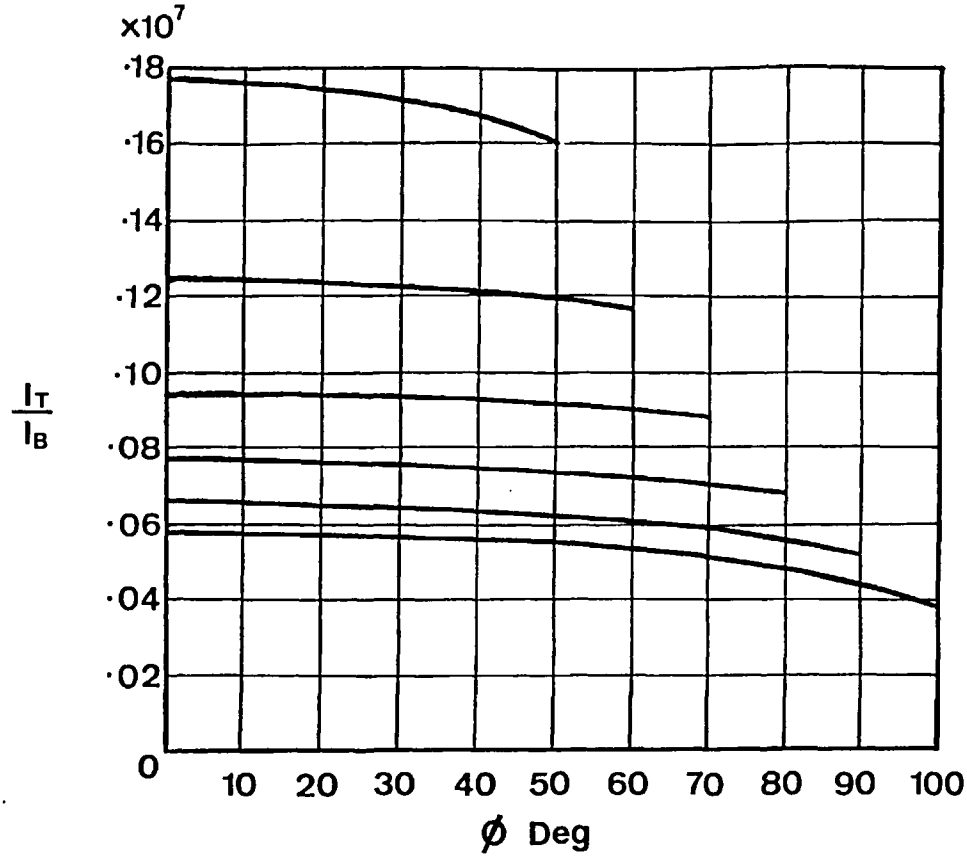


Illumination at 1·064μm.

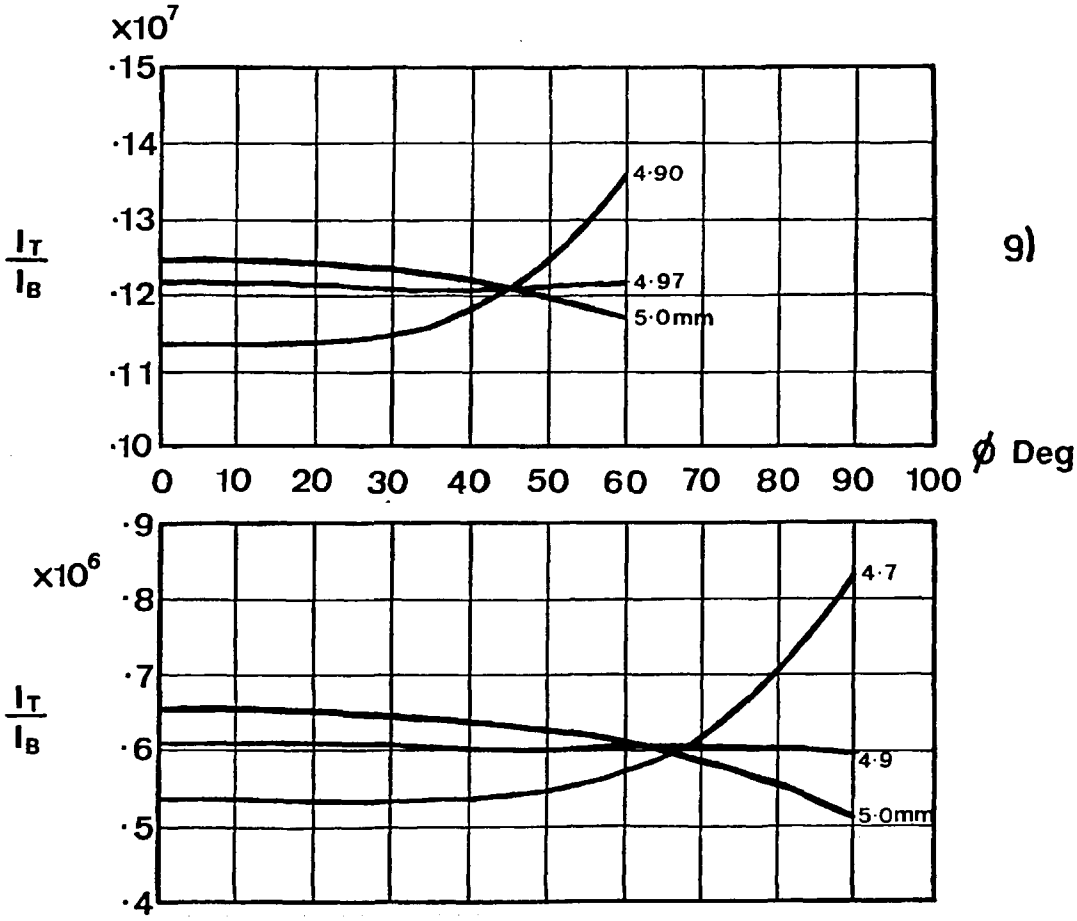
d.) Illumination for various coverages, ϕ_m. d_2=5·0mm.

e.) The effect of changing the spacing. ϕ_m=-60° & -90°.

Fig. (6.2)



f)



g)

Illumination at 0·532μm.

f.) Illumination for various coverages, ϕ_m. d_2=5·0mm.

g.) The effect of changing the spacing.

ϕ_m=-60° & -90°.

Fig. (6.2)

c_i (paraxial)	d_i
•012153288	
	15•0
•004006060	
	5•0 (at 0•532 & 1•064μm)
•014217070	
	17•5
•004006060	

F = 108•1484mm, b = 85•5899mm (1•064μm)

F = 105•5060mm, b = 83•1236mm (0•532μm)

J	a_{J1}	a_{J3}
4	$-•1070275145 \times 10^{-6}$	$-•4732954656 \times 10^{-7}$
6	$-•1857072804 \times 10^{-10}$	$•7596262953 \times 10^{-12}$
8	$-•1352049922 \times 10^{-14}$	$-•1151189909 \times 10^{-14}$
10	$-•1907084698 \times 10^{-18}$	$-•1270947070 \times 10^{-18}$
12	$-•1310731182 \times 10^{-22}$	$-•4016004057 \times 10^{-22}$
14	$-•2617292581 \times 10^{-26}$	$-•5260173762 \times 10^{-26}$
16	$-•1619759082 \times 10^{-30}$	$-•1348384627 \times 10^{-29}$
18	$-•4205569198 \times 10^{-34}$	$-•2009094077 \times 10^{-33}$
20	$-•2011868258 \times 10^{-38}$	$-•4587494775 \times 10^{-37}$
22	$-•7397365829 \times 10^{-42}$	$-•7375641824 \times 10^{-41}$
24	$-•2366785909 \times 10^{-46}$	$-•1582205947 \times 10^{-44}$
26	$-•1386375140 \times 10^{-49}$	$-•2687039867 \times 10^{-48}$

Table (6.1) System data for the doublet in fig. (6.2)

J	a_{J1}	a_{J3}
4	$-\cdot1087930034 \times 10^{-6}$	$-\cdot4473200217 \times 10^{-7}$
6	$-\cdot1852934671 \times 10^{-10}$	$\cdot8389814655 \times 10^{-12}$
8	$-\cdot1366293638 \times 10^{-14}$	$-\cdot1107150508 \times 10^{-14}$
10	$-\cdot1882966130 \times 10^{-18}$	$-\cdot1258271692 \times 10^{-18}$
12	$-\cdot1346498297 \times 10^{-22}$	$-\cdot3884560905 \times 10^{-22}$
14	$-\cdot2562696125 \times 10^{-26}$	$-\cdot5224010221 \times 10^{-26}$
16	$-\cdot1710220071 \times 10^{-30}$	$-\cdot1311305822 \times 10^{-29}$
18	$-\cdot4060210076 \times 10^{-34}$	$-\cdot1998389929 \times 10^{-33}$
20	$-\cdot2243800979 \times 10^{-38}$	$-\cdot4478320342 \times 10^{-37}$
22	$-\cdot7022849144 \times 10^{-42}$	$-\cdot7338984770 \times 10^{-41}$
24	$-\cdot2966673624 \times 10^{-46}$	$-\cdot1550290846 \times 10^{-44}$
26	$-\cdot1290116594 \times 10^{-49}$	$-\cdot2673185348 \times 10^{-48}$

Table (6.2) Alternative aspheric surface coefficients for the system in fig (6.2) in order to provide uniform target illumination when $R_T = 50\mu m$ and $\Phi_m = -90^0$

suggests that this would be possible.

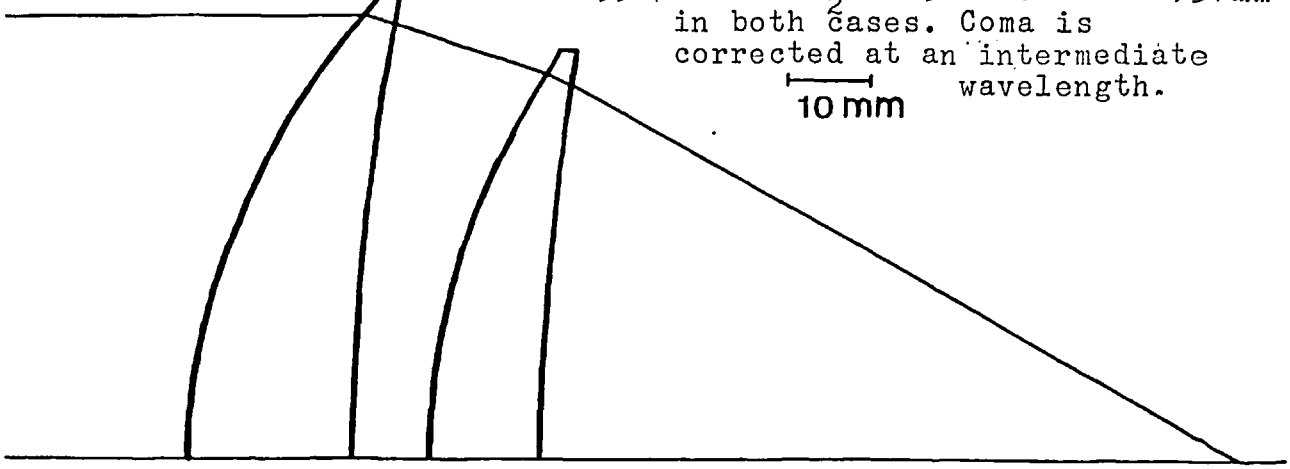
On attempting either of these it was found they could not be satisfied. The OSC is negative for the doublet and remains so with respect to all variations in the element shapes. With each change in shape, the aspherics are altered to re-satisfy the constant optical path condition. No doubt the reason for failure is that Seidel theory alone is no longer accurate at such high apertures. Therefore, as shown in Fig.(6.2c) OSC_m has not been fully corrected.

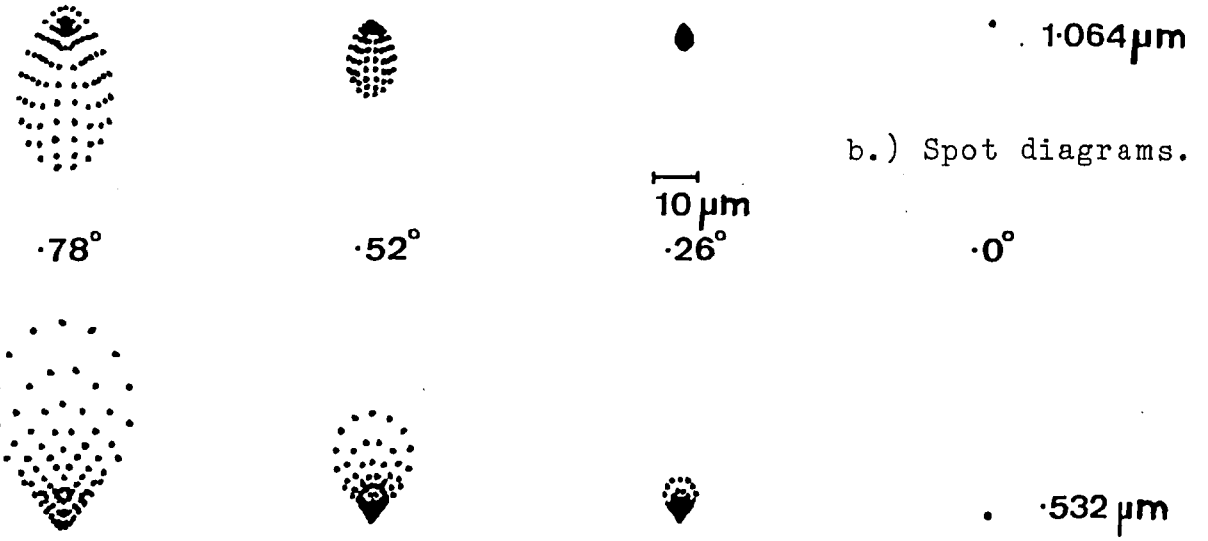
In Figs.(6.2d) to (6.2g) the target illumination properties are shown at both wavelengths. By adjusting the spacing slightly, uniformity to within 5% is possible.

For absolute uniformity the process in Chapter 4 of selective aberrating has to be applied. For R_T=50μm and ϕ_m=-90° table (6.2) gives the new aspheric profiles required. The amount of aberration introduced is no more than a 0·71λ and 0·41λ wavefront error for the visible and infrared wavelengths respectively.

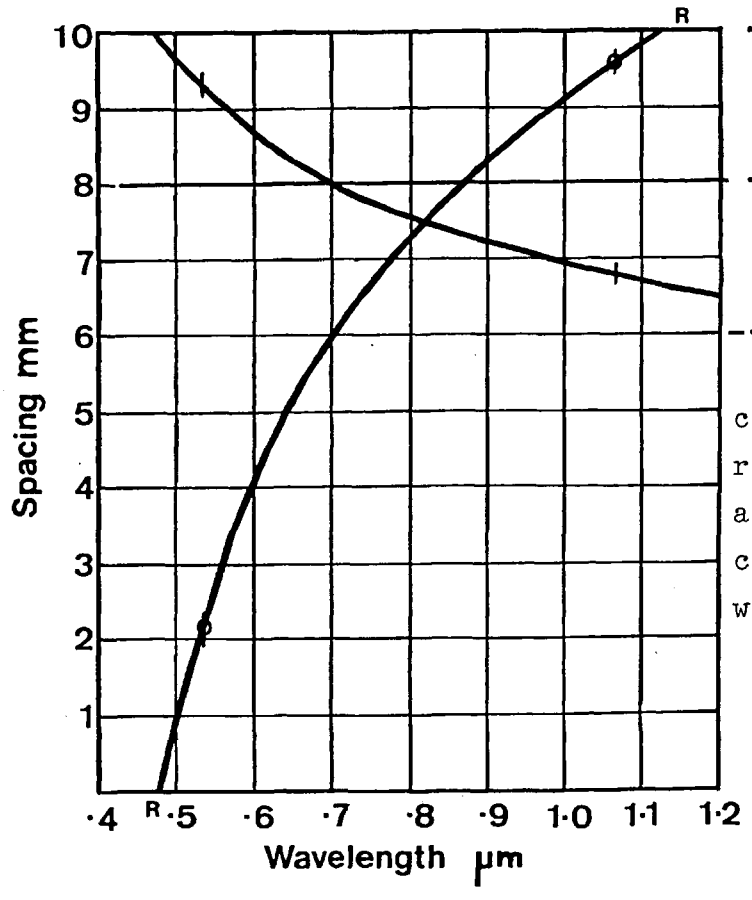
The coma may be corrected at an intermediate wavelength by relaxing the condition that the component spacing is the same for the infrared and its harmonic. The doublet just described above was fully corrected at both wavelengths when the separation is 5·0mm. The system illustrated in Fig.(6.3) is corrected at 1·064μm when the spacing is 9·6318mm. On changing the wavelength to 0·532μm, exact stigmatism is only achieved when reduced to 2·1501mm. This alteration has allowed the amount of coma to be considerably reduced, as shown in Fig.(6.3b) and (6.3c). OSC_m = 0 at 0·7μm; this gives

Fig. (6.3) a.)F/1 doublet that is corrected at 1·064μm when d_2=9·6318mm and at 0·532μm when d_2=2·1501mm. b=86·8937mm in both cases. Coma is corrected at an intermediate wavelength.



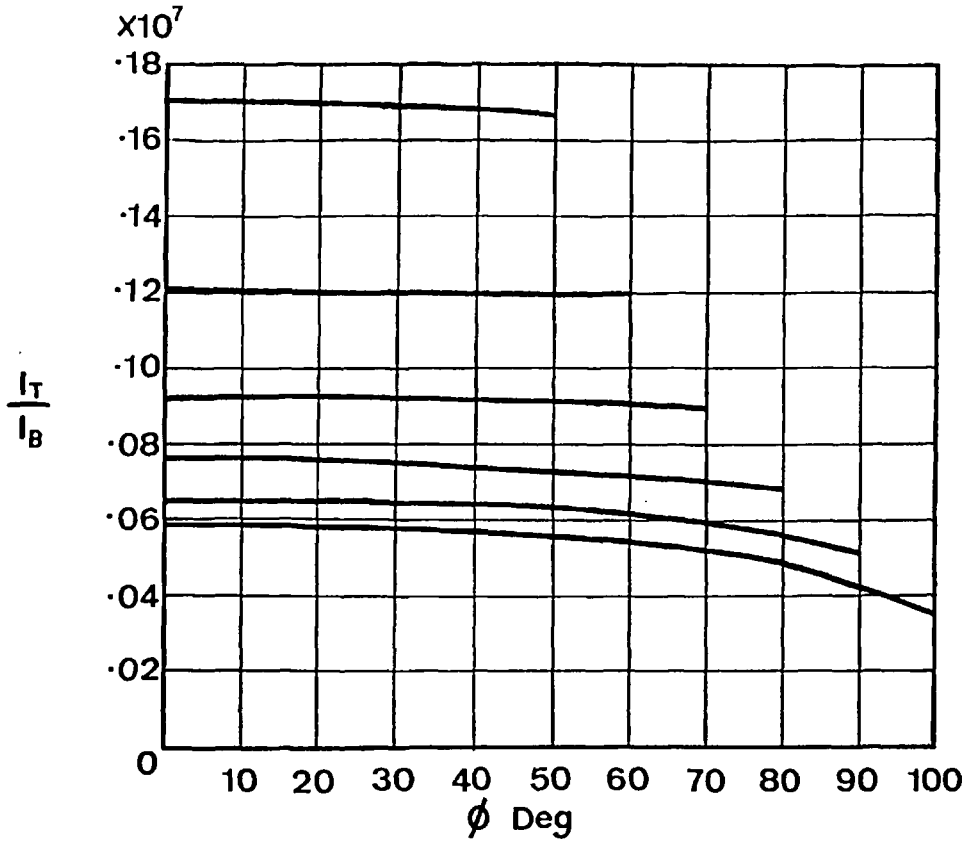


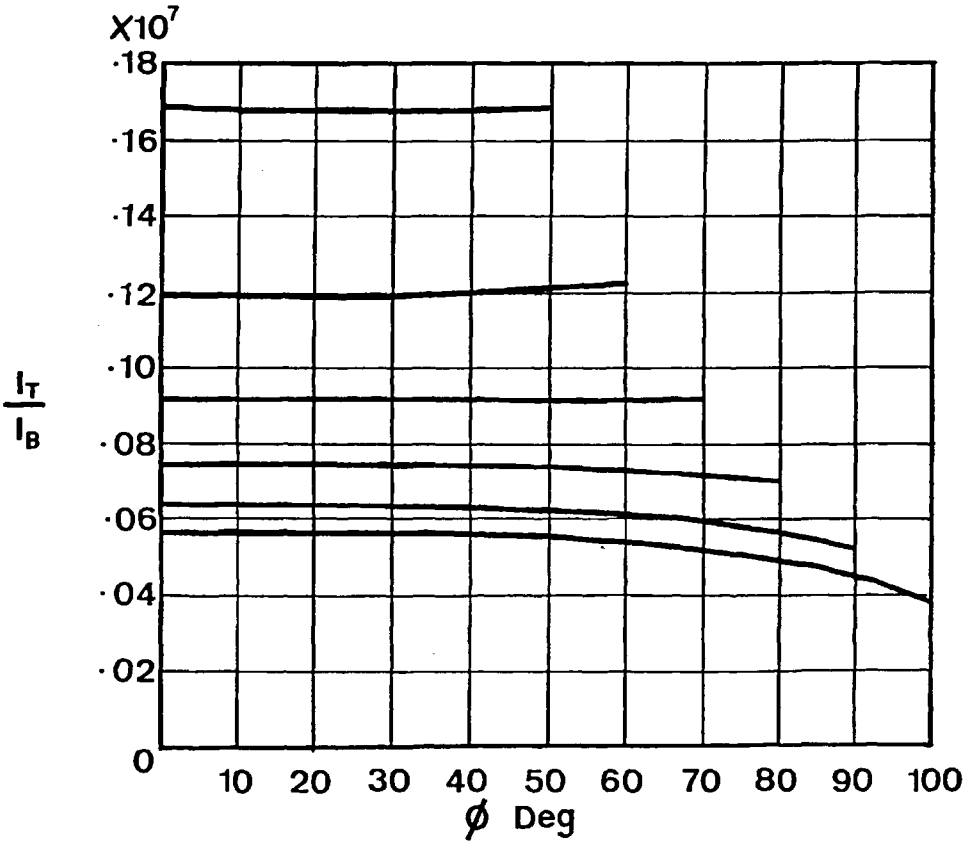
b.) Spot diagrams.



c.) Element spacing required for aberration correction and OSC_m when this is set.

Fig. (6.3)





d.) Target illumination at 1·064μm and d_2=9·6318mm.
e.) Target illumination at 0·532μm and d_2=2·1501mm.

Fig. (6.3)

c_i (paraxial)	d_i
•01300773	
	20•5
•00300773	
	9•6318 at 1•064μm
	2•1501 at 0•532μm
•01111285	
	13•5
•00300773	

F = 112•7995mm, b = 86•8937mm (1•064μm)

F = 108•0721mm, b = 86•8937mm (0•532μm)

J	a_{J1}	a_{J3}
4	$-•1347179333 \times 10^{-6}$	$-•4156143470 \times 10^{-7}$
6	$-•1738891232 \times 10^{-10}$	$-•5525404628 \times 10^{-11}$
8	$-•2156242589 \times 10^{-14}$	$-•5723107913 \times 10^{-15}$
10	$-•2738644375 \times 10^{-15}$	$-•5659351854 \times 10^{-19}$
12	$-•3589946825 \times 10^{-22}$	$-•5640582938 \times 10^{-23}$
14	$-•4847020411 \times 10^{-26}$	$-•5739743284 \times 10^{-27}$
16	$-•6713485876 \times 10^{-30}$	$-•5951753720 \times 10^{-31}$
18	$-•9500542160 \times 10^{-34}$	$-•6264318762 \times 10^{-35}$
20	$-•1368799455 \times 10^{-37}$	$-•6675272528 \times 10^{-39}$
22	$-•2001931885 \times 10^{-41}$	$-•7191067742 \times 10^{-43}$
24	$-•2965130009 \times 10^{-45}$	$-•7822919969 \times 10^{-47}$
26	$-•4439057935 \times 10^{-49}$	$-•8585247193 \times 10^{-51}$

Table (6.3) System data for the doublet in fig. (6.3)

J	a_{J1}	a_{J3}
4	$-\cdot1345675908 \times 10^{-6}$	$-\cdot4173875599 \times 10^{-7}$
6	$-\cdot1737399710 \times 10^{-10}$	$-\cdot5545290436 \times 10^{-11}$
8	$-\cdot2154843161 \times 10^{-14}$	$-\cdot5740875021 \times 10^{-15}$
10	$-\cdot2737299158 \times 10^{-18}$	$-\cdot5674040146 \times 10^{-19}$
12	$-\cdot3588608452 \times 10^{-22}$	$-\cdot5652941580 \times 10^{-23}$
14	$-\cdot4845636448 \times 10^{-26}$	$-\cdot5750757191 \times 10^{-27}$
16	$-\cdot6711994126 \times 10^{-30}$	$-\cdot5962216556 \times 10^{-31}$
18	$-\cdot9496864798 \times 10^{-34}$	$-\cdot6274888385 \times 10^{-35}$
20	$-\cdot1368603016 \times 10^{-37}$	$-\cdot6686656724 \times 10^{-39}$
22	$-\cdot2001693096 \times 10^{-41}$	$-\cdot7204243919 \times 10^{-43}$
24	$-\cdot2964830042 \times 10^{-45}$	$-\cdot7839261088 \times 10^{-47}$
26	$-\cdot4438670343 \times 10^{-49}$	$-\cdot8606832434 \times 10^{-51}$

Table (6.4) Alternative aspheric coefficients for the doublet in fig. (6.3) in order to provide uniform target illumination when R_T = 50μm and $\phi_m = -90^0$.

approximately equal and opposite amounts at the two wavelengths.

Another property the doublet has is that its back focal length is the same for both wavelengths (86·8937mm). If $\lambda<\lambda^*$ then $n>n^*$ and when the spacing is kept constant $b<b^*$. So in order to restore the original back focal length in changing from λ to λ^* the spacing would have to be increased. Thus $d^*>d$. The shift may be easily deduced by a paraxial raytrace.

If the front element is the one moved, the second element remaining fixed with respect to the target position, then there should be no re-alignment problems. The slight change in power will alter the coverage, ϕ_m, by a few degrees but this is unlikely to be significant. The pros and cons of whether it is better to move the whole lens or just one component have to be considered.

Figs.(6.3d) and (6.3e) show the target illumination provided at each wavelength. Again they may be improved in the usual manner, but the back focal length will no longer be the same in each case.

In Table (6.4) the corresponding aspherics for the selectively aberrated system are shown. Again uniform illumination for $R_T=50\mu m$ and $\phi_m=-90^0$ was used.

CHAPTER 7

NONLINEAR EFFECTS AND DAMAGE

7.1. Nonlinear effects

As mentioned back in Chapter 1, the transmission of a high powered laser beam through a medium may lead to damage from the nonlinear effects. The latter being caused by the terms in E^2, E^3,... of the relation,

$$P = \varepsilon_0 \chi_1 E + \varepsilon_0 \chi_2 E^2 + \varepsilon_0 \chi_3 E^3 + \ldots . \qquad (7.1)$$

This applies not only to the focussing optics but all the components within the laser. If not actually resulting in damage, the propagation of the beam through large amounts of glass can cause it to break up; this is a consequence of the variation of index with intensity given by,

$$n_{NL} = n + n_I I \qquad (7.2)$$

The damage can appear in many ways; holes, cracks, bubbles and even shattered elements. Any one of these defects which does occur will then become the source for further trouble. A bubble inside the glass, for example, caused either by damage already sustained or produced when the glass blank was cast, will strongly absorb radiation and inevitably lead to greater damage. The original glass blanks must therefore be of the highest quality; have high homogeneity, be free from bubbles, straie etc. and the polish on each surface must also be of very high quality. All pits and scratches must be completely removed. The finished lens must also be scrupulously cleaned.

Those nonlinear effects which are believed to be chiefly responsible for damage are briefly described as follows.

a.) Self focussing

The variation of refractive index with beam intensity is responsible for this effect, (Equation (7.2)). n_I is positive, so the index increases with intensity. Thus for a beam with a higher central intensity the refractive index is also higher for the central zones and therefore the propagation speed is less. The central portion of the wavefront will lag behind the outer regions and the effect is a converging beam. The eventual result is a very bright filament that rapidly becomes unstable because of its effect on the glass medium. The damage left behind is in the form of a chain of bubbles.

b.) Brillouin Scattering

In this phenomena the radiation exerts a pressure on the medium producing a sound wave. This acoustic wave can then fracture the glass. Scattered light of a slightly lower frequency also results from the process; the energy difference between incident and scattered light appears in the sound wave.

c.) Thermal Shock

As already said, any defect already present will strongly absorb radiation causing a serious over-heating; hence the thermal shock. It can lead to melting, vapourization and fracture.

The damage threshold varies slightly from material to material; but appears to virtually half for a surface coating

Sample	Pulse duration 0•125ns	0•60ns	1•10ns
Fused silica	10•0±1•0	23•0±3•0	27•0±4•0
BK7	9•2±1•5	19•5±3•5	25•0±3•0
99% reflector	5•1±1•0	9•5±2•0	11•5±2•5
Antireflecting film	5•5±1•0	11•0±2•0	9•5±2•0

Table (7.1) List of measured damage thresholds at 1•064μm obtained by Smith, Milam and Weber (2). Units are J/cm^2.

either reflecting or antireflective. Smith, Milan and Weber (2) give a table of measured surface damage thresholds for three pulse durations and for various materials. Apparently, the threshold tends to vary as the square root of the pulse duration. Thus increasing the pulse duration by a factor of 4 will lead to a doubling of the damage threshold. Broadly speaking, damage can be sustained with energy densities in the range 5 to 25J/cm^2 for 0·1 to 1·0ns pulses. Uncoated surfaces occupy the upper half of this range and coated surfaces the lower half. Table (7.1) gives some of these figures.

Surface coatings are therefore a weak link in high powered laser optics and for this reason are often dispensed with. For the highly curved surfaces, the production of an even coating which has identical reflectivities for all zones is very difficult.

7.2. Ghost images

Provided high quality glass free of defects is used for the focussing optics and the incident beam diameter is large enough to keep the energy density below the damage threshold there should be no damage from the beam as it propagates directly through the elements. Unfortunately, problems arise from "ghost images". These are images that result from multiple reflections from the surfaces and even from the target plasma.

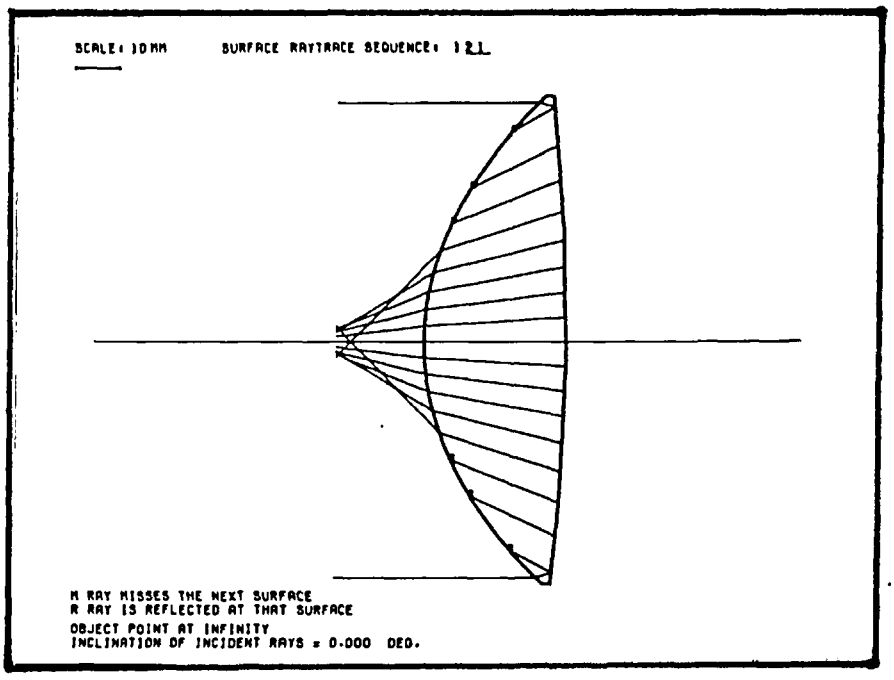
In general, they will be highly aberrated images and if any fall inside the elements, the high energy density at a

focus or in a caustic is likely to result in damage.

Even though only 4% (for paraxial zone) of the energy is reflected at each surface, its concentration in a ghost image can easily damage the elements. At the marginal zone where the angles of incidence are higher, the losses per reflection vary from about 2·5% to 5·8% for external and 0·5% to 10·6% for internal reflections. A 30^0 incidence angle and an index of 1·5 assumed. The first of these figures applies to the case when the plane of polarization lies in the incidence plane and the second, perpendicular to it. Since it is usual for the beam to be polarized, there will be this variation with azimuth.

The location of the ghost images and caustics may be found by raytracing. A program for doing this must have facilities for tracing rays back and forth through the system to follow a possible combination of internal reflections. In addition, the tracing of a particular ray should be discontinued if it is deflected out to one side or undergoes total internal reflection and therefore cannot follow the path originally assigned. Other combinations will trace these internally reflected rays.

For singlets nothing can be done to control their location. By considering enough reflections off the front and back surfaces, rays will always focus inside the lens. F/1 singlets will produce an internal image after only two reflections; from the rear and front surfaces respectively. This is illustrated in Fig. (7.1) which shows the graphical output from a "ghost-image program". The singlet in Fig.(4.3)



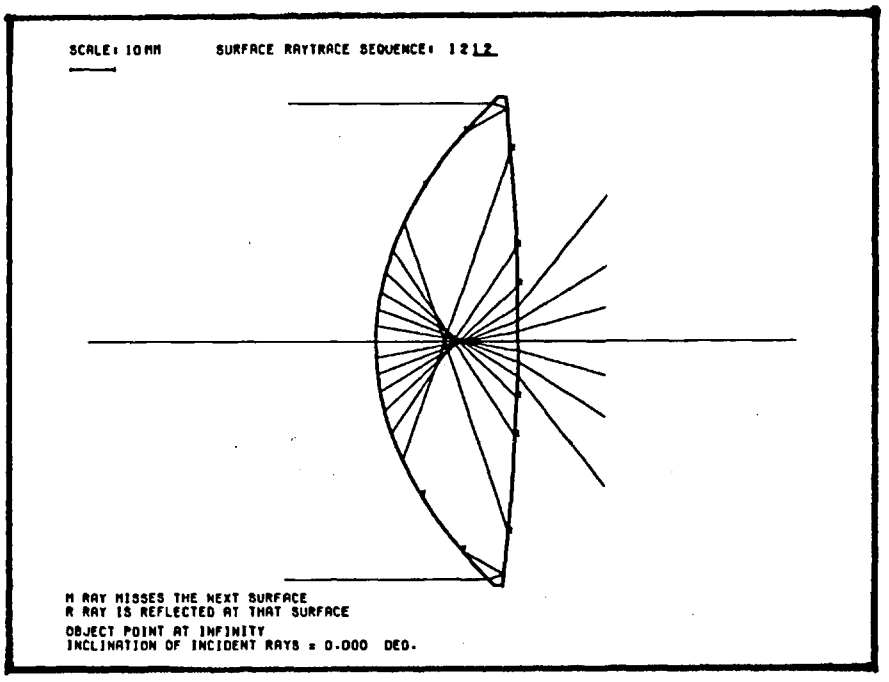
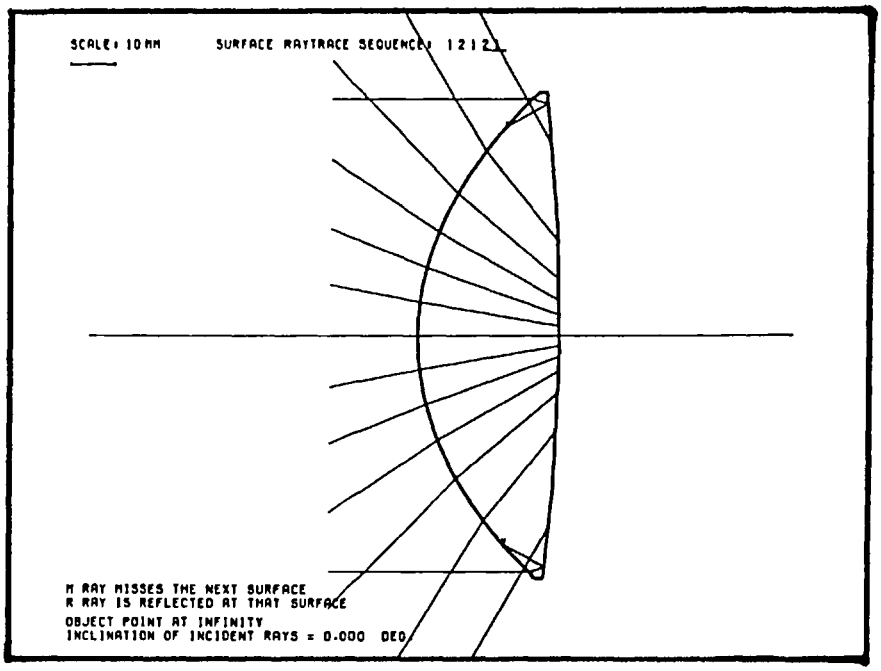
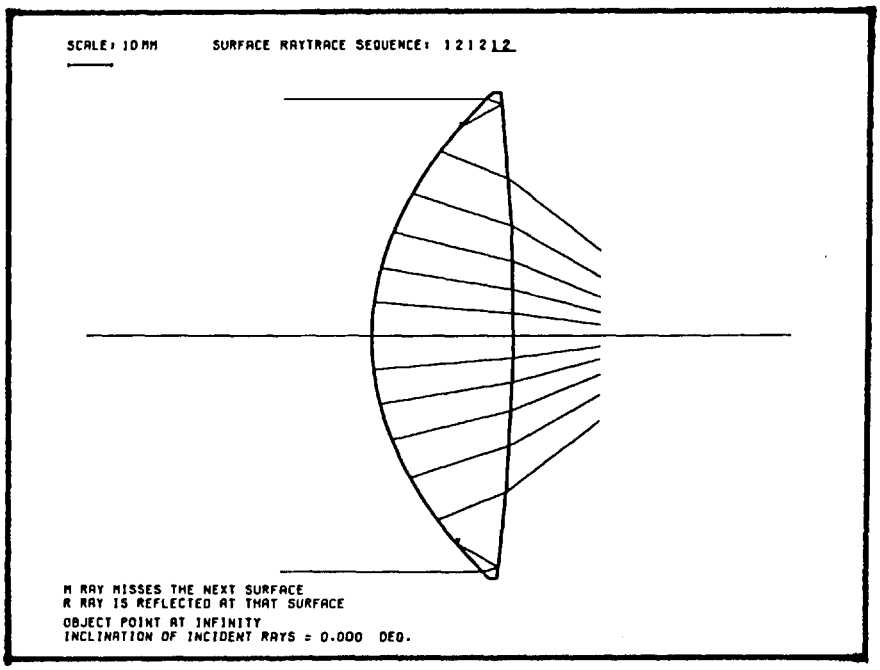


Fig. (7.1) The output from a ghost-image location program applied to the lens in fig.(4.3). The "surface raytrace sequence" at the top of each diagram shows the path through which the rays are traced and the underlined part the section for which they have been drawn.
Up to 4 successive reflections are drawn but only one case produces an internal ghost-image. This can be dealt with by cutting a central hole through the lens or an annular ring in from the front surface as illustrated in fig. (7.3).


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 1 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 1 2 1 2
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

has been used and up to 4 successive reflections are shown.

As well as looking for images formed within the lens itself, some attention should also be given to those which occur back down the beam path just in case they occur within other optical components.

For doublets some attempt could be made to choose the spacing so that the main ghost images would fall within the spacing. Complete success is unlikely however. The number of surfaces is now 4 and the possible combinations that can produce these trouble spots begin to increase dramatically. In the singlet's case, there are 2 combinations for 1 reflection and only 1 combination for more than 1 reflection. For the doublet, there are 4 combinations of 1 reflection, 6 for 2 reflections and 14 for 3 reflections. Fig.(7.2) shows all these drawn out for the lens in Fig.(4.4). As can be seen, there are several combinations which focus the radiation inside or just outside an element. Notice, that a couple appear in the same place, so their effects may add if the optical path difference does not exceed the pulse length. (30cm to 3cm for 1ns to 100ps pulses respectively).

The only suitable method for dealing with these ghost images is to cut a central hole through each element. This has the unfortunate consequence of producing a small shadow area on the target, but is the price to pay for preventing damage. For a singlet, an annular saw cut into the lens from one side would be an alternative (3). This would disrupt the focussed rays without removing the central core and therefore minimizing the area of shadow. Fig.(7.3) illustrates

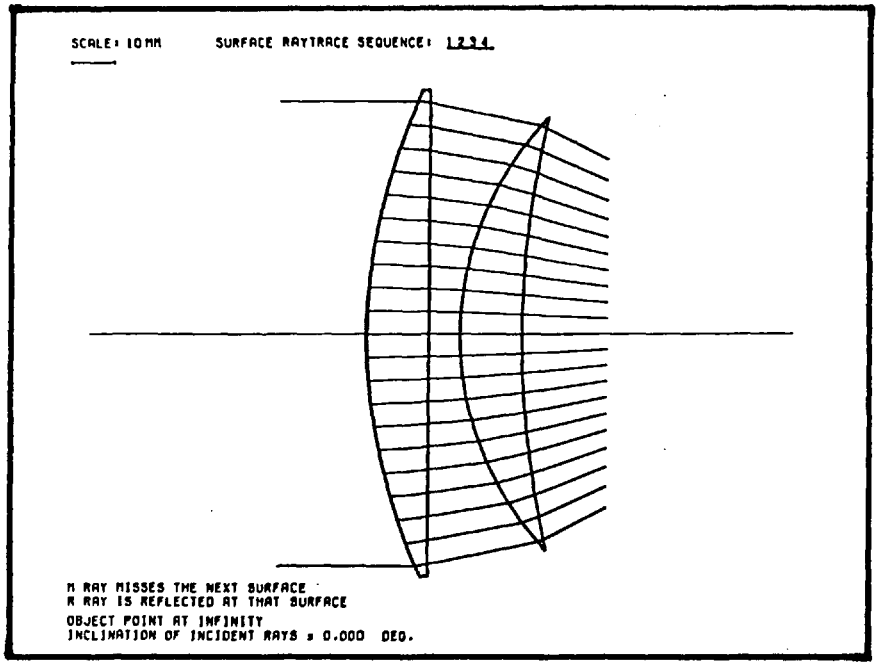
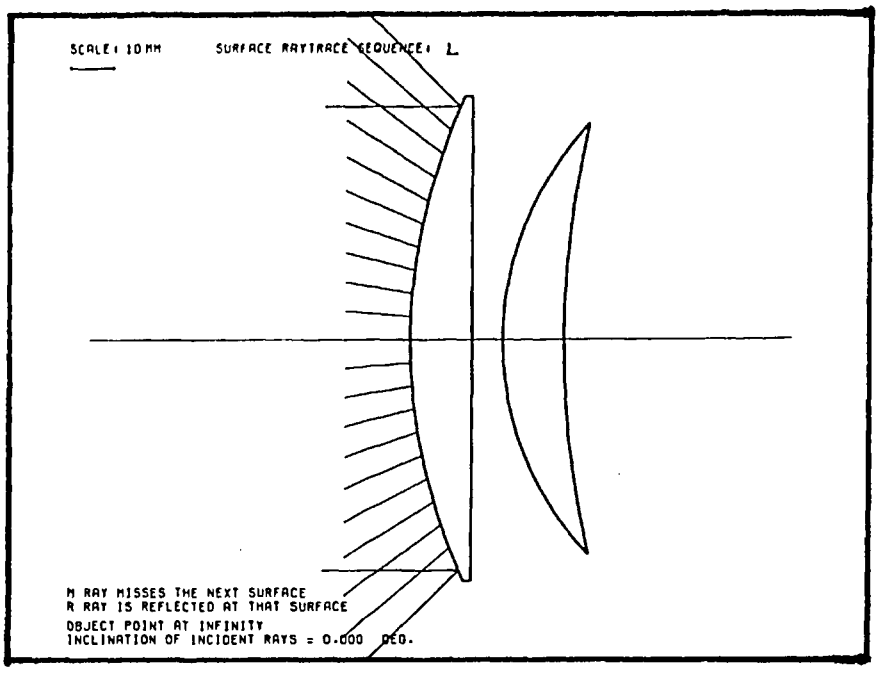
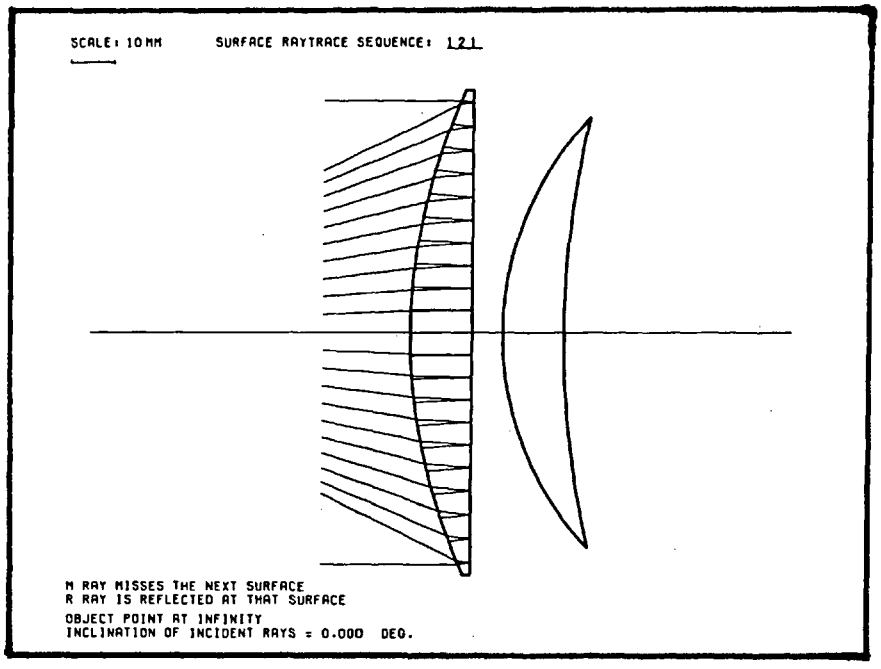
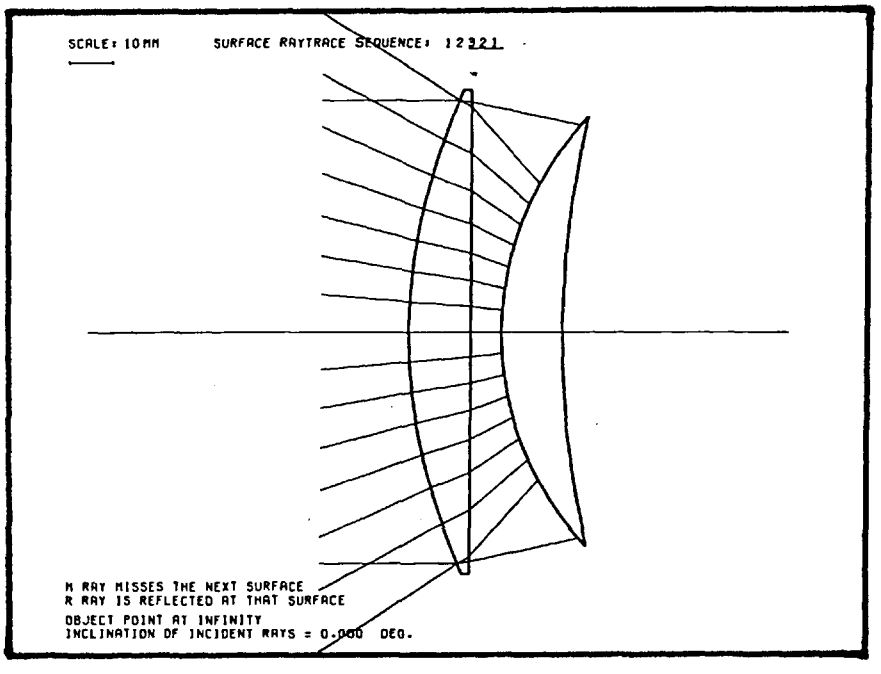
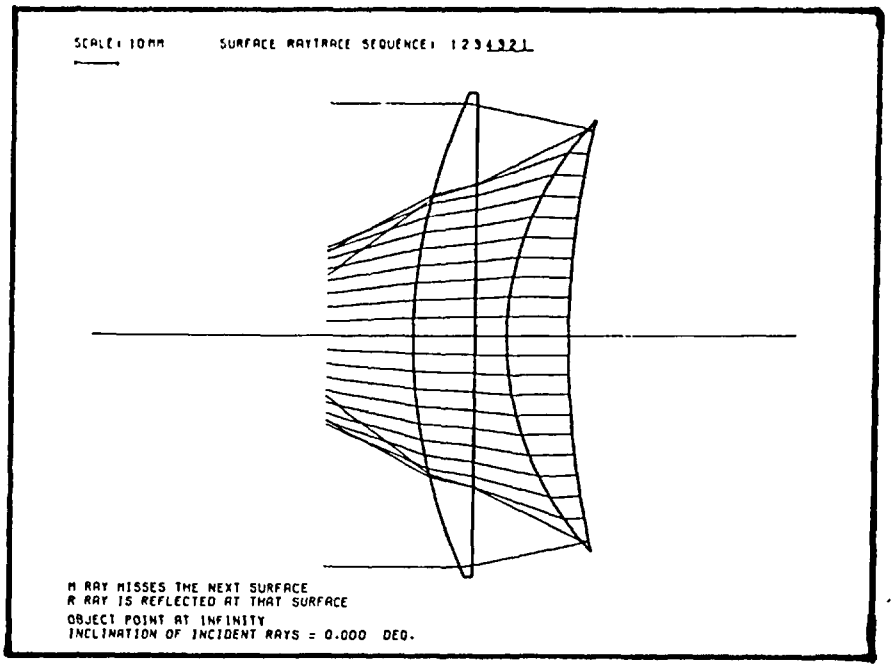


Fig. (7.2) The program output for the doublet in fig. (4.4). Up to 3 successive reflections are shown producing 24 combinations.

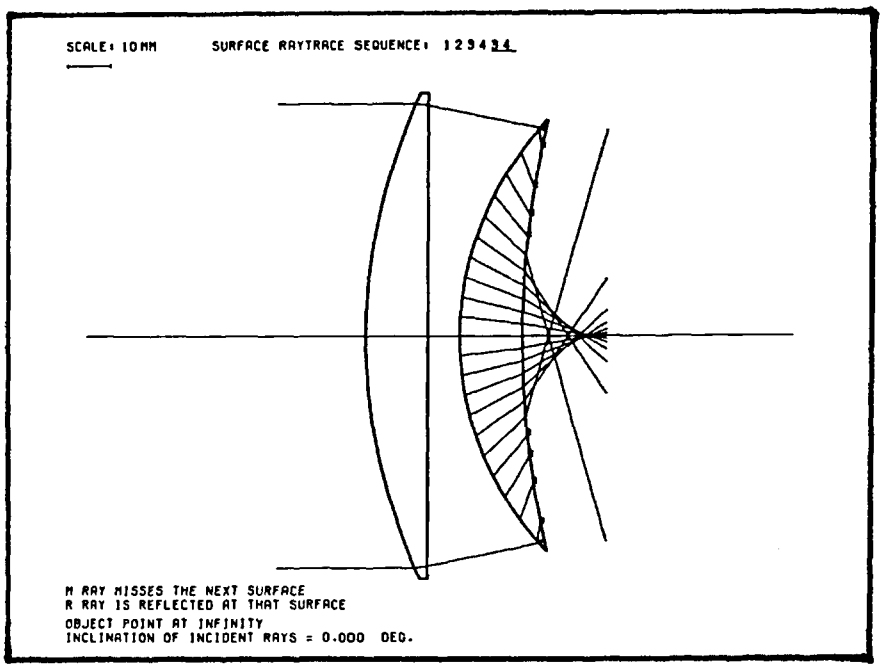

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

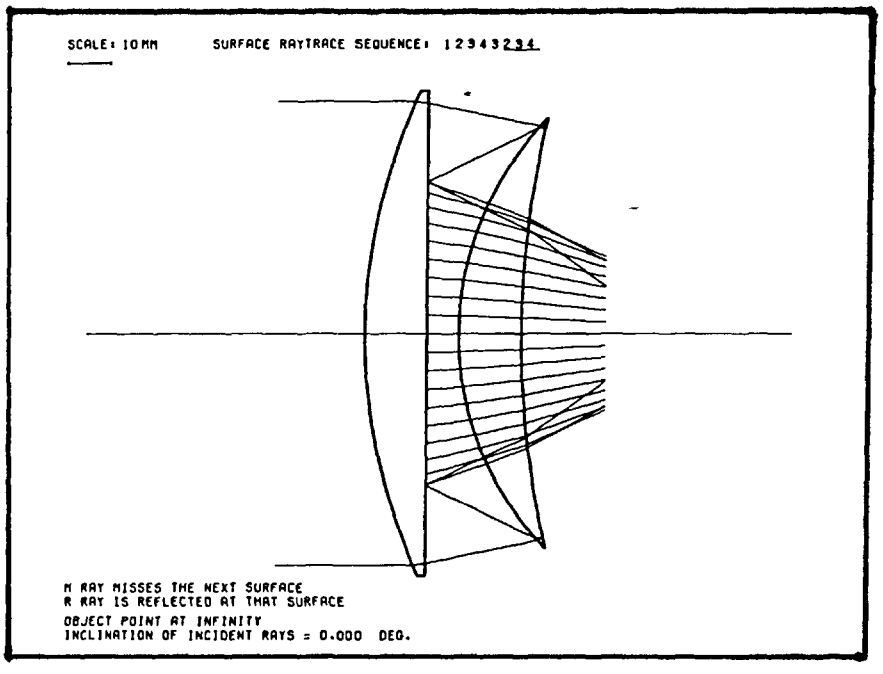

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 121
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

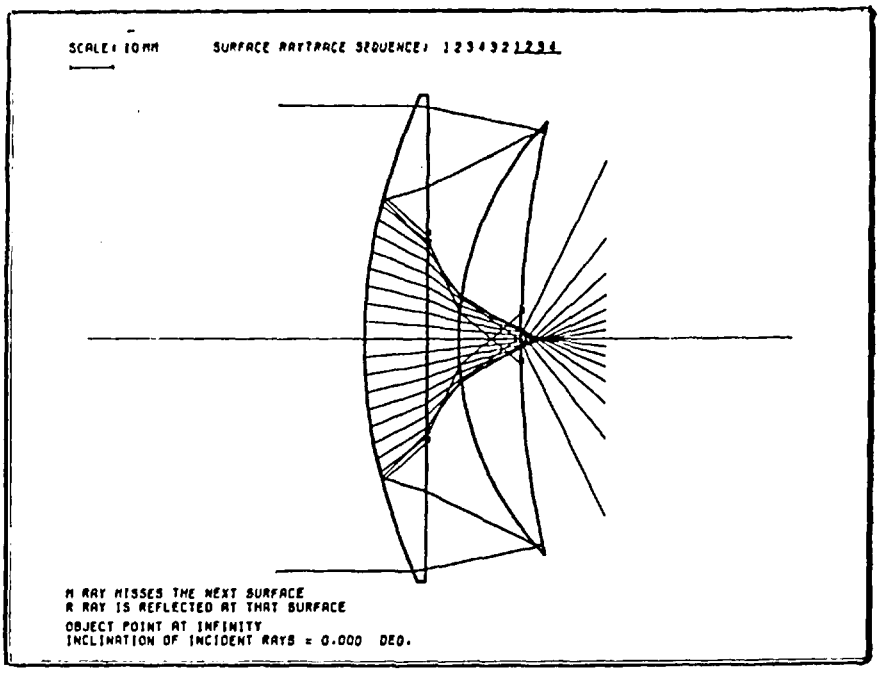

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 12321
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

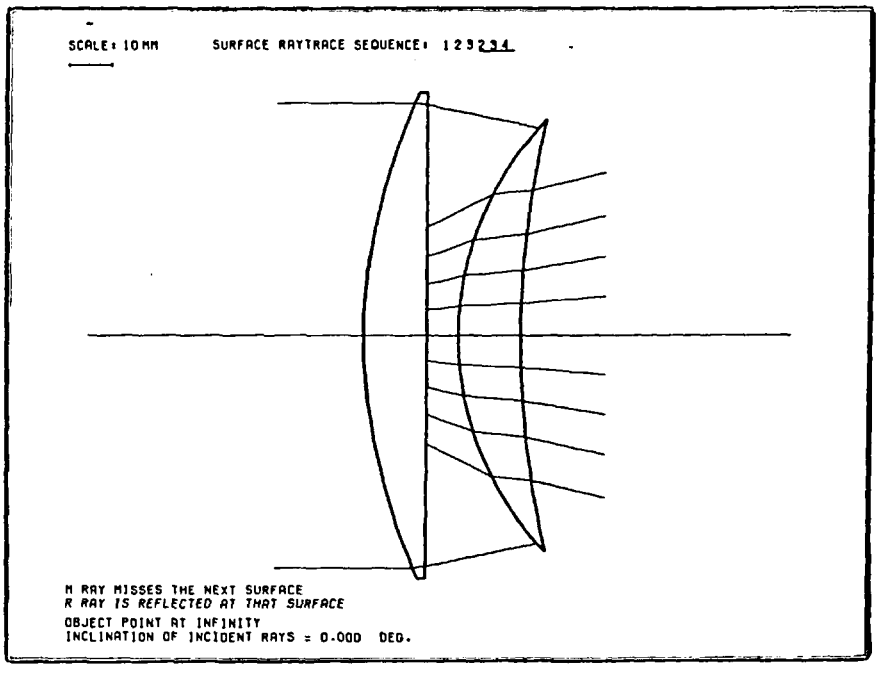


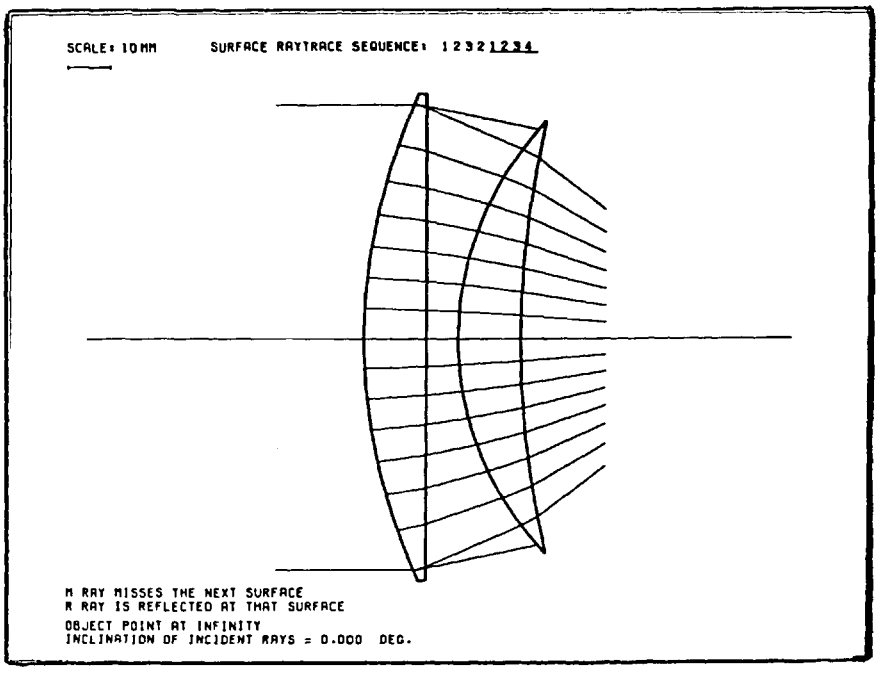
2 Reflections

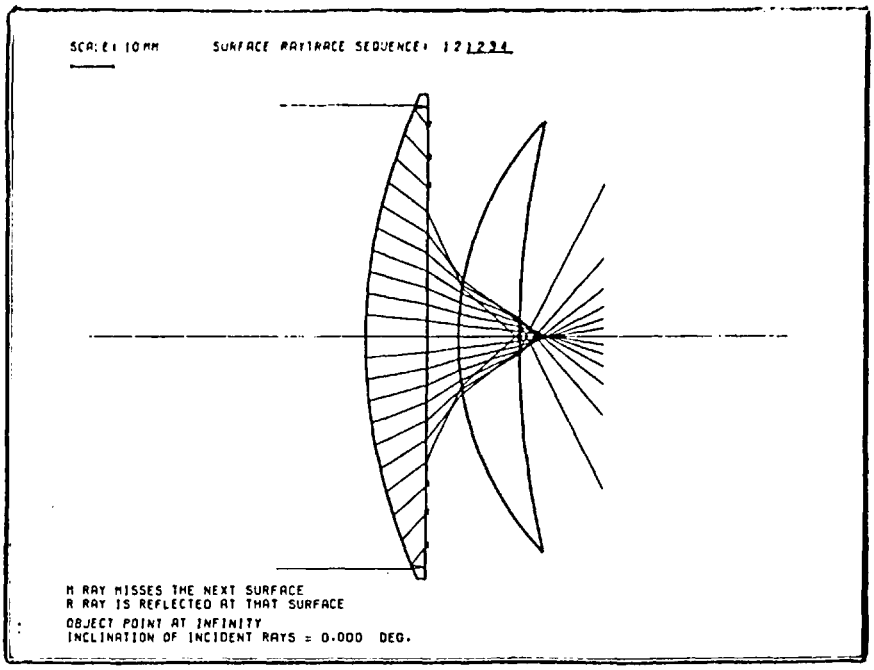




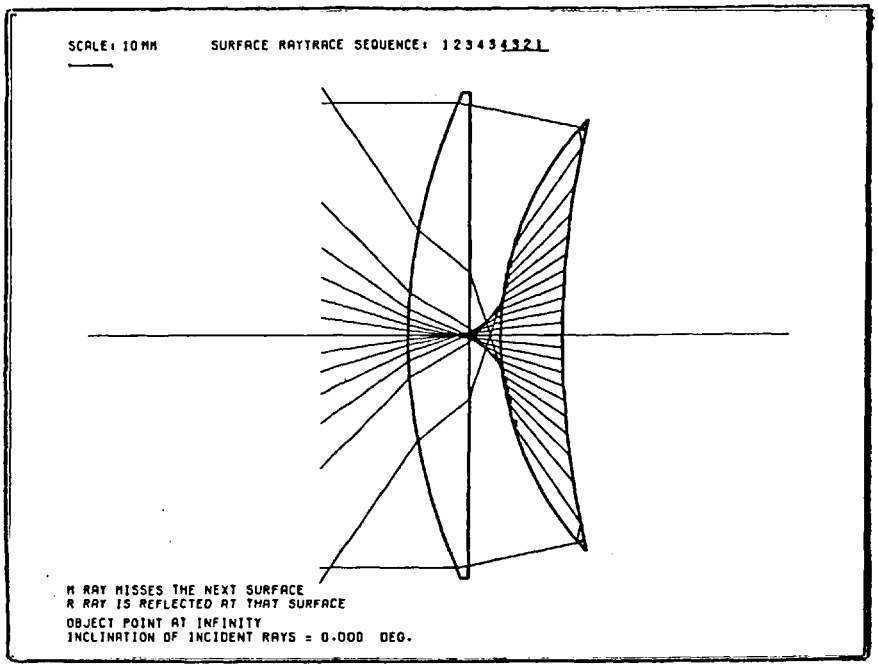

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 4 3 2 1 2 3 4
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

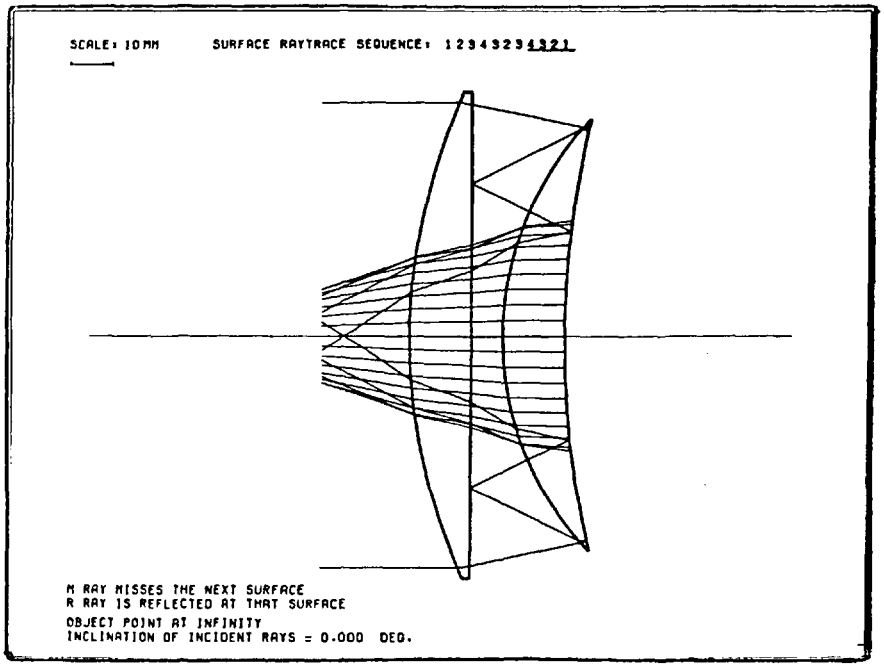

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 3 4
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

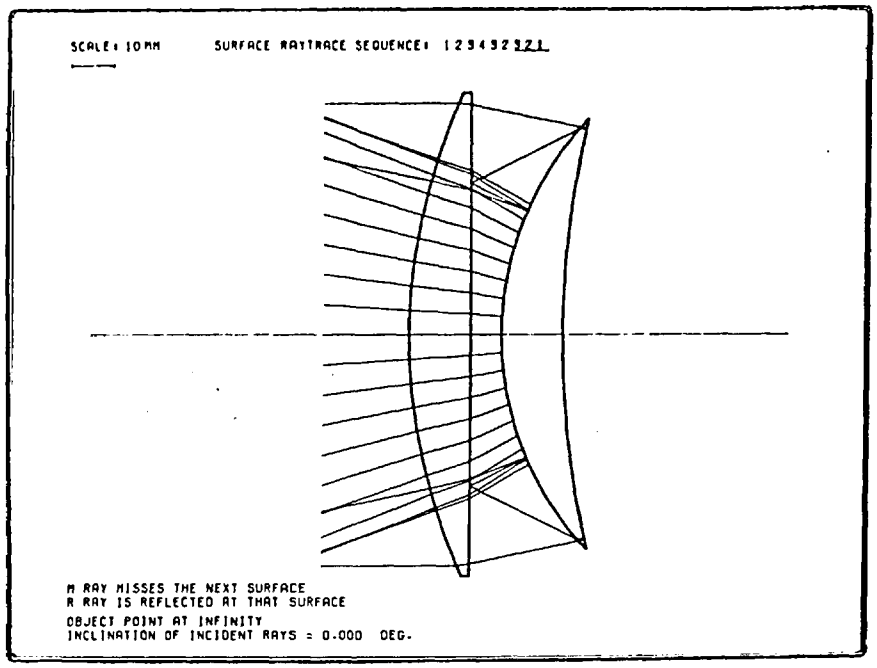

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 1 2 3 4
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

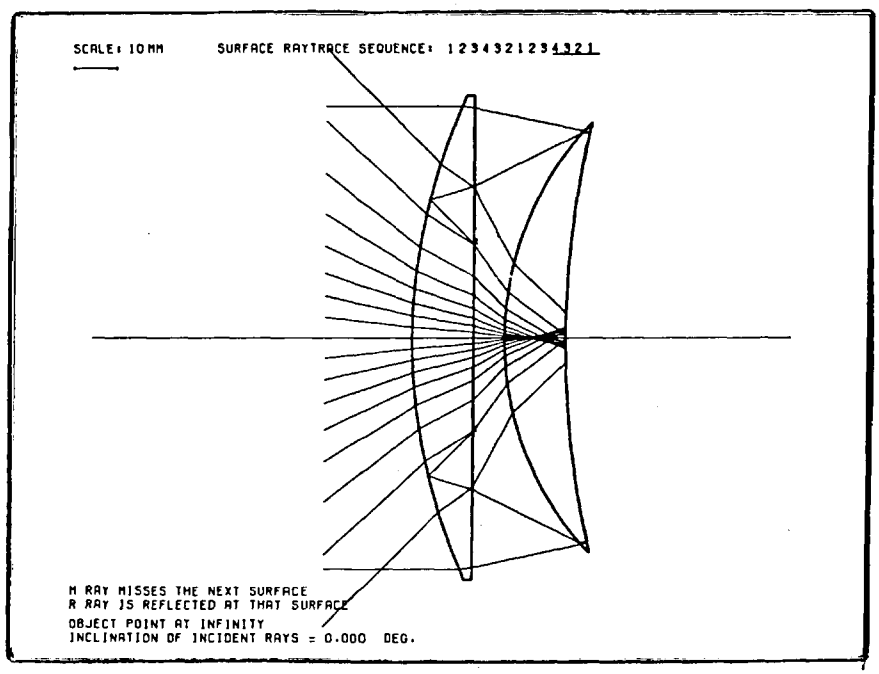


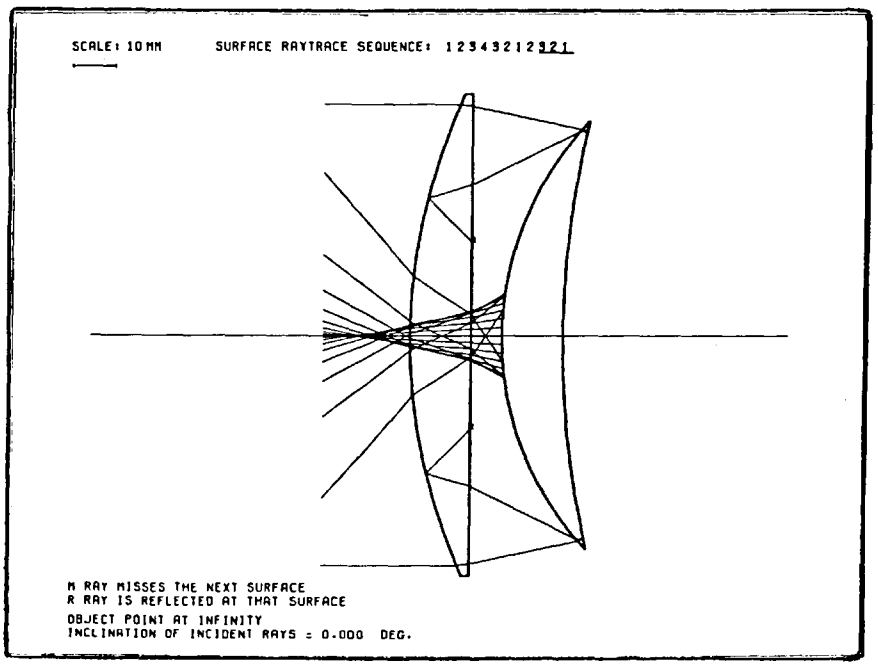
3 Reflections

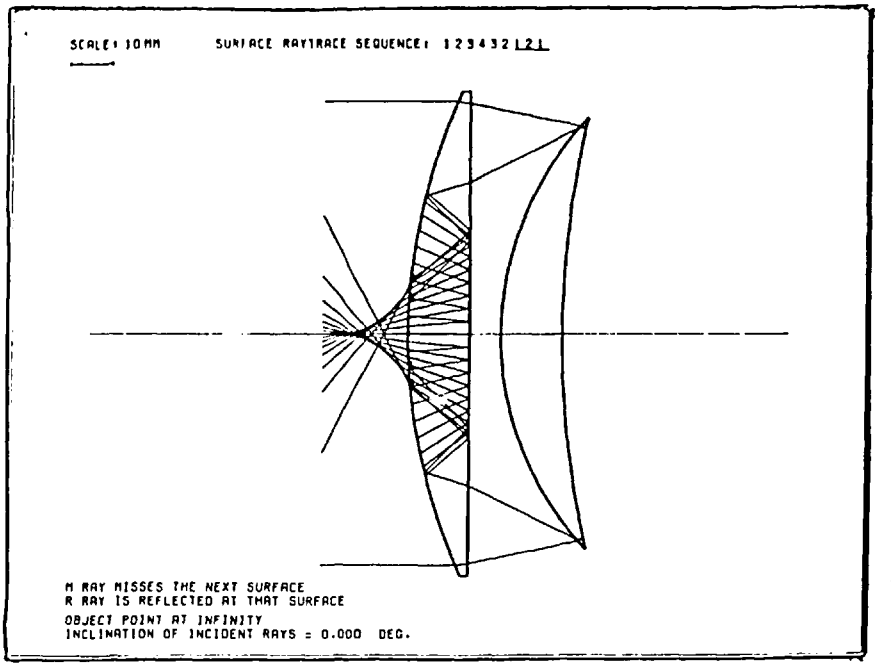


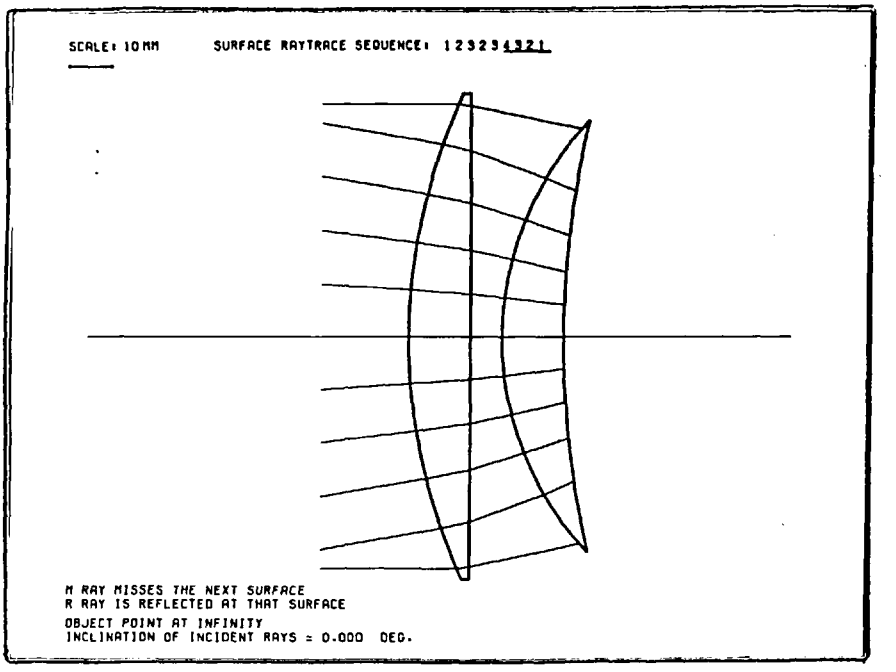


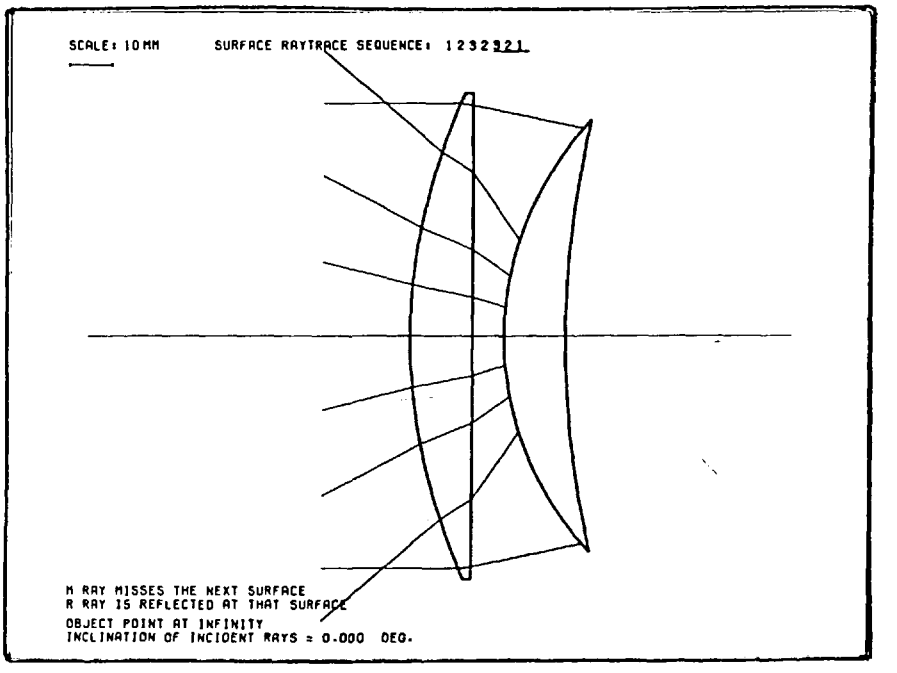

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 4 3 2 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

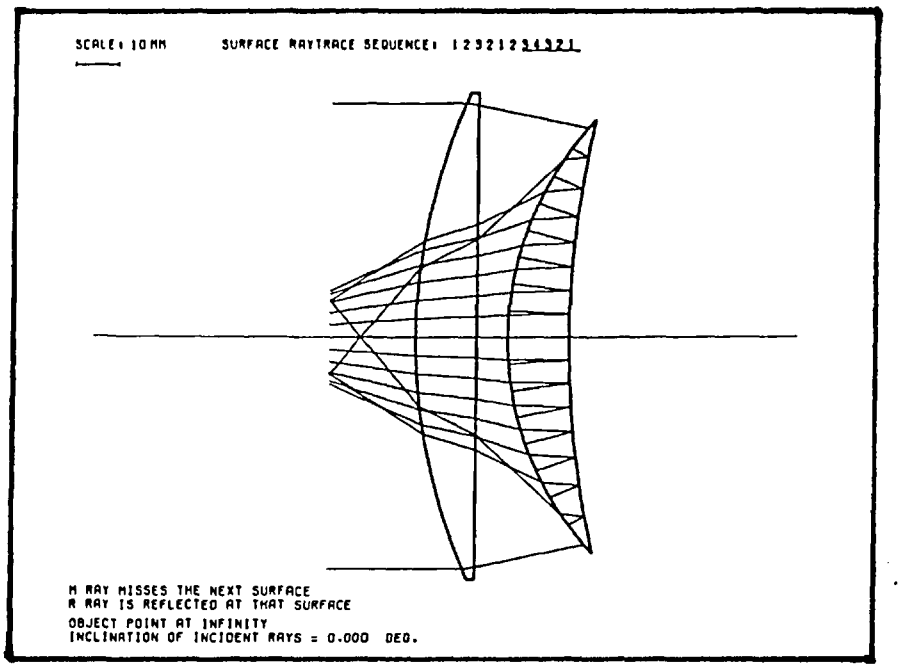

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 4 3 2 1 2 3 4 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

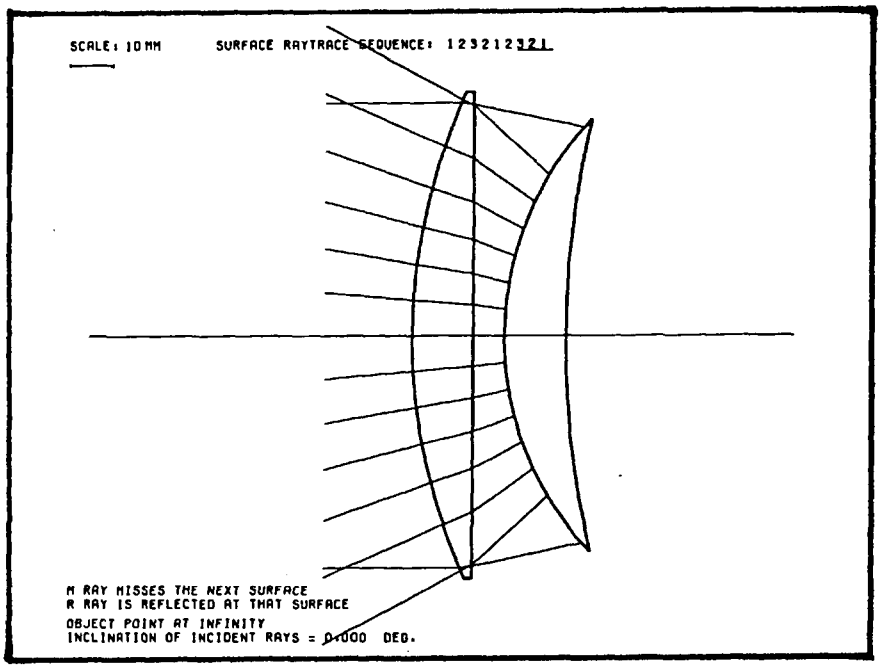

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 4 3 2 1 2 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

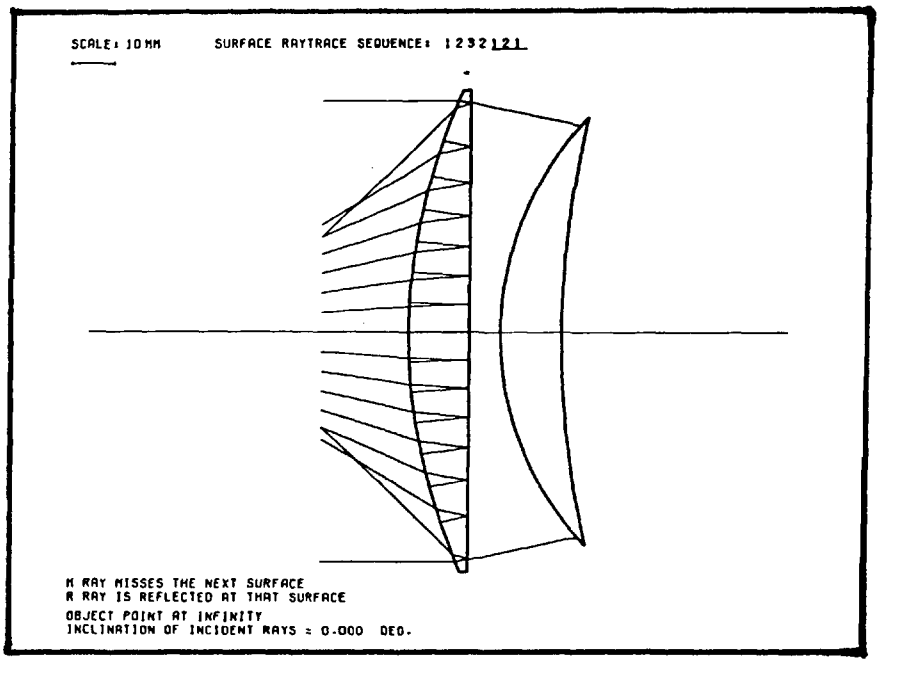

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 4 3 2 1 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

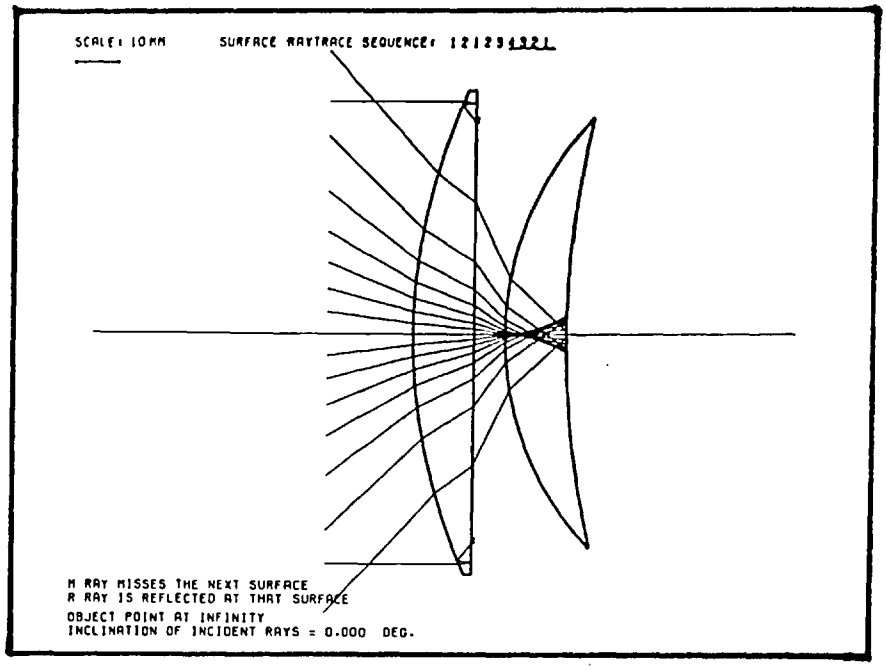

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 3 4 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

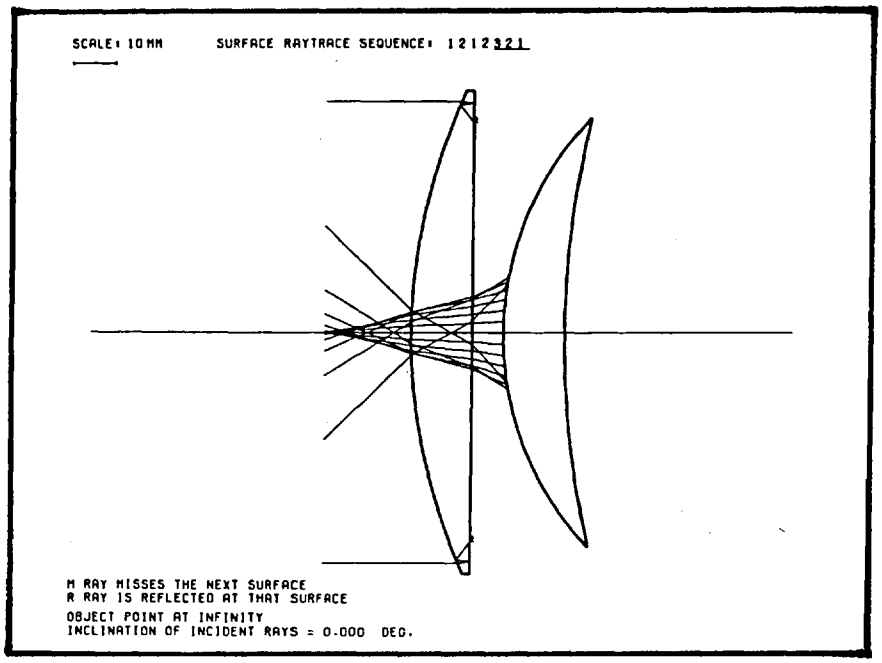

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

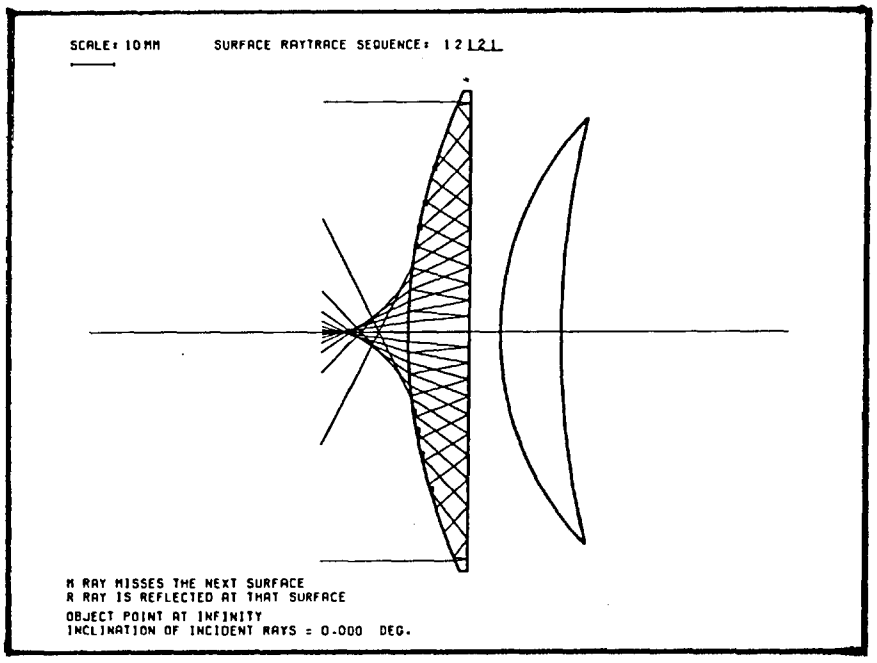

SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 1 2 3 4 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 1 2 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 3 2 1 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 1 2 3 4 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 1 2 3 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.


SCALE: 10 MM
SURFACE RAYTRACE SEQUENCE: 1 2 1 2 1
M RAY MISSES THE NEXT SURFACE
R RAY IS REFLECTED AT THAT SURFACE
OBJECT POINT AT INFINITY
INCLINATION OF INCIDENT RAYS = 0.000 DEG.

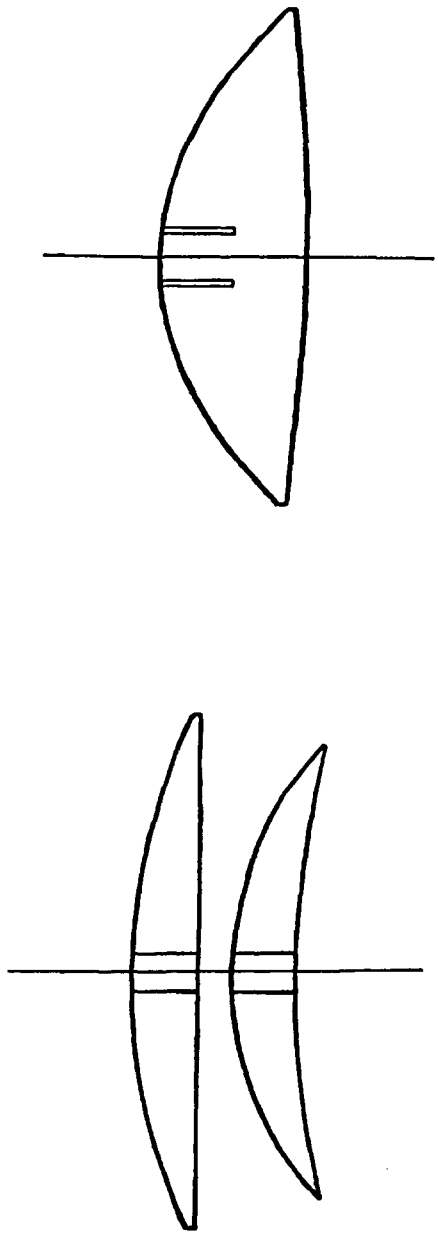


Fig. (7.3) The method for preventing damage from the ghost images is to cut a central hole through each element or where appropriate an annular ring thus minimizing the shadow area over the target.

this.

The only other alternative to this problem would be the use of antireflection coatings but their disadvantages have already been given. Added complications occur if we consider the dual-wavelength lenses in Chapter 6. Any coating would have to be antireflecting at two wavelengths. At least a two-layer film would actually be required (Achromatized film; A. Vasvek (21)). Thus unless a very even coating can be guaranteed, they should not be applied.

7.3. Nonlinear phase: the B-integral

The original reason for restricting the number of refracting elements was the desire to keep the so called B-integral as small as possible. This expresses the amount of nonlinear phase introduced as the beam propagates through optical glass.

$$B_{\int} = \frac{2\pi}{\lambda} n_I \int I \, ds \qquad (7.3)$$

where, to good approximation, the integral is along the unperturbed ray path and the refractive index given by (7.2).

Since the intensity of the beam is likely to vary with Y, the incident ray height, and the path length inside glass will certainly do so, then the amount of nonlinear phase will also. An incident plane wavefront will therefore emerge distorted from the shape it would normally have at low intensities. The beam will then tend to break up as it propagates down the laser. We shall therefore calculate the contribution of the focussing optics to the B-integral.

Since the greatest amount of glass is along the axis and the intensity is going to be higher here, then the B-integral along the axial ray will also be a maximum. The calculation will therefore be restricted to the axis. Thus equation (7.3) reduces to,

$$B_{\int} = \frac{2\pi n_I}{\lambda} \int I(Z)dZ \tag{7.4}$$

where I(z) is the intensity along the axis. We now write this as a summation of terms from all the different media traversed.

$$B_{\int} = \frac{2\pi}{\lambda} \sum_{i=1,2,3,} n_{Ii} \int_0^{d_i} I_i(Z)dZ \tag{7.5}$$

Here n_{Ii} is the nonlinear index of the ith medium and $I_i(Z)$ the intensity within it.

A paraxial raytrace through the lens will give the ray heights at each surface, y_i, and if the intensity at the pole of each surface is I_{0i} then by the conservation of energy,

$$\pi y_1^2 I_{01} = \pi y_2^2 I_{02} = \ldots = \pi y_i^2 I_{0i} = \ldots \tag{7.6}$$

or

$$I_{0i} = I_{01} \frac{y_1^2}{y_i^2} \tag{7.7}$$

Clearly I_{01} is also the central intensity of the incident beam, which is known. Thus we can find the other I_{0i} from (7.7).

Now in order to find $I_i(Z)$, note that

$$\pi y_i^2(Z)\, I_i(Z) = \pi y_1^2\, I_{01} \tag{7.8}$$

or

$$I_i(Z) = I_{01}\, \frac{y_1^2}{y_i^2(Z)} \tag{7.9}$$

where $y_i(Z)$ is the ray height at any plane between surfaces i and i+1 and varies linearly with Z from y_i at Z=0 to y_{i+1} at $Z=d_i$. Therefore,

$$y_i(Z) = (y_{i+1} - y_i)\frac{Z}{d_i} + y_i \tag{7.10}$$

Hence (7.9) becomes,

$$I_i(Z) = \frac{I_{01} y_1^2}{\left[(y_{i+1} - y_i)Z/d_i + y_i\right]^2} \tag{7.11}$$

and putting this in the integral of equation (7.5)

$$\int_0^{d_i} I_i(Z)dZ = I_{01} y_1^2 \int_0^{d_i} \frac{dZ}{\left[(y_{i+1}-y_i)Z/d_i + y_i\right]^2} \tag{7.12}$$

$$= \frac{-I_{01} y_1^2 d_i}{(y_{i+1}-y_i)} \left[\frac{1}{(y_{i+1}-y_i)Z/d_i + y_i}\right]_0^{d_i} \tag{7.13}$$

$$= \frac{-I_{01} y_1^2 d_i}{(y_{i+1}-y_i)} \left[\frac{1}{y_{i+1}} - \frac{1}{y_i}\right] \tag{7.14}$$

$$= \frac{-I_{01} y_1^2 d_i\, (y_i - y_{i+1})}{(y_{i+1}-y_i) y_{i+1} y_i} \tag{7.15}$$

c_i	d_i	n_i	y_i	u_i
		1•0		•0
•0077101			1•0	
	14•5	1•50621		-•002591
•0			•9624	
	7•345	1•0		-•003903
•01349528			•9338	
	14•5	1•50621		-•006826
•00455873			•8348	
		1•0		-•008355

Table (7.2) Paraxial raytrace of the doublet shown in Fig.(4.4)

$$= \frac{I_{01} y_i^2 d_i}{y_{i+1} y_i} \qquad (7.16)$$

Substituting into (7.5)

$$B_\int = \frac{2\pi y_1^2 I_{01}}{\lambda} \Sigma \frac{n_{Ii} d_i}{y_{i+1} y_i} \qquad (7.17)$$

$$i=1,2,3,.$$

Apart from the notation this result agrees with Glass and Trenholme (22) derived in a slightly different way.

As an example, consider the doublet of Fig.(4.4). With a beam radius of 55mm or 5·5 cm, a power of 500GW will produce a flux of 5·3GW/cm^2 or 5·3 x 10^{7}W/mm^2. It is below the damage threshold and therefore represents a reasonable power level. This is the value of I_{01}. Table (7.2) shows the paraxial raytrace through this lens. The value of n_I for BK7 from table (1.1) is 3·4 x 10^{-16}cm^2/W or 3·4 x 10^{-14}mm^2/W .So,

$$B_\int = \frac{2\pi\ 1^2\ 5{\cdot}3\ x10^7}{1{\cdot}064x10^{-3}} \left[\frac{3{\cdot}4x10^{-14}x14{\cdot}5}{0{\cdot}9624x1} + \frac{3{\cdot}4x10^{-14}x14{\cdot}5}{0{\cdot}9338x0{\cdot}8348} \right]$$

= 0·358 rad

= 0·06λ error.

CHAPTER 8

SUMMARY AND CONCLUSIONS

The focussing optics can be loosely divided into two classes,

1.) Those using two aspheric surfaces to satisfy two conditions:

a.) To achieve an aberration free image and a desired energy transformation (Chapter 3).

b.) To achieve an aberration free image or a desired energy transformation at two wavelengths (Chapter 6).

and,

2.) Those using only one aspheric surface to satisfy one condition:

a.) To achieve an aberration free image.

b.) To achieve a desired energy transformation; the image being selectively aberrated.

The energy transformation any particular system produces depends on the spherical aberration present and the shape of the principal surface. If the former is absent (image is diffraction limited) then the target intensity depends on the principal surface alone. Since coma is desired to be reasonably well corrected then even the principal surface is restricted to a roughly spherical shape. Thus the optics in the 1a) category above, are limited to energy transformations that also correct coma sufficiently well. The transformations of a Gaussian beam required a system with enormous coma though with a uniform beam the transformation onto the target

can be achieved with much reduced coma. This really limits the use of 1a) optics to laser systems where the beams have been smoothed out from the usual Gaussian profile.

On the other hand, if we allow spherical aberration to be present the target illumination can be optimized for any incident beam intensity while still keeping the coma at a tolerable level. This was done by either using a single aspheric that aberrates the image by just the right amount or by adjusting the spacing between the components of a doublet. The latter option having the advantage that different beam profiles, target size, target coverage and even wavelengths can be accommodated. Thus a wider set of conditions may be dealt with. The systems of this type include the wavelength tunable and dual-wavelength lenses. It appears that uniformity to within a few percent can be achieved from a lens corrected for the constant optical path condition and absolute uniformity by selectively aberrating the image. Thus there is no real need to manufacture two aspheric surfaces required for the optics in 1a) when to all intents and purposes the same results can be obtained by using one aspheric.

The uniformity provided by the wavelength tunable lenses is not surprisingly, worse. With the aspheric devoted to correcting the system at some suitable visible wavelength, we are left with less control over the aberrations and hence illumination in the infrared.

APPENDIX A

THE ALDIS THEOREM

The Aldis theorem is derived by Cox (19) and in a slightly different form by Welford (23). It expresses the total transverse ray aberration for a finite ray with respect to a paraxial principal ray. The equations used for its derivation are the following; the finite and paraxial raytracing equations,

$$\frac{X_i - X_{i-1}}{L_{i-1}} = \frac{Y_i - Y_{i-1}}{M_{i-1}} = \frac{Z_i - Z_{i-1} + d_{i-1}}{N_{i-1}} \tag{A.1}$$

$$\frac{n_i L_i - n_{i-1} L_{i-1}}{\alpha_i} = \frac{n_i M_i - n_{i-1} M_{i-1}}{\beta_i} = \frac{n_i N_i - n_{i-1} N_{i-1}}{\gamma_i} \tag{A.2}$$

where we have adopted the symbol L to represent the X direction cosine for a ray.

$$y_i = y_{i-1} + d_{i-1} u_{i-1} \tag{A.3}$$

$$n_i u_i = n_{i-1} u_{i-1} - c_i y_i (n_i - n_{i-1}) \tag{A.4}$$

$$B_i = n_i (y_i c_i + u_i) = n_{i-1} (y_i c_i + u_{i-1}) \tag{A.5}$$

The Lagrange invariant is,

$$H = n_i (\bar{y}_i u_i - y_i \bar{u}_i) = n_{i-1} (\bar{y}_i u_{i-1} - y_i \bar{u}_{i-1}) \tag{A.6}$$

where a bar ($\bar{}$) refers to the principal ray.

Also the surface equation,

$$F_i(X,Y,Z) \equiv c_i (X_i^2 + Y_i^2 + Z_i^2) - 2Z_i + f_i = 0 \tag{A.7}$$

The direction cosines of a surface normal $(\alpha_i, \beta_i, \gamma_i)$ are therefore given by,

$$(\alpha_i,\beta_i,\gamma_i) = \frac{-(F_{Xi},F_{Yi},F_{Zi})}{\sqrt{F_{Xi}^2+F_{Yi}^2+F_{Zi}^2}} \qquad (A.8)$$

$$(\alpha_i,\beta_i,\gamma_i) = \frac{-(f_{Xi}+2c_iX_i,f_{Yi}+2c_iY_i,2(c_iZ_i-1))}{\sqrt{F_{Xi}^2+F_{Yi}^2+F_{Zi}^2}} \qquad (A.9)$$

Consider the Y-component only. Rearranging (A.1),

$$(Y_i - Y_{i-1})N_{i-1} = (Z_i - Z_{i-1} + d_{i-1})M_{i-1} \qquad (A.10)$$

and multiply by $n_{i-1}u_{i-1}$,

$$n_{i-1}u_{i-1}N_{i-1}(Y_i - Y_{i-1}) = n_{i-1}u_{i-1}M_{i-1}(Z_i-Z_{i-1}+d_{i-1}) \qquad (A.11)$$

Substituting for $u_{i-1}d_{i-1}$ from (A.3),

$$n_{i-1}u_{i-1}N_{i-1}(Y_i-Y_{i-1}) = n_{i-1}u_{i-1}M_{i-1}(Z_i-Z_{i-1}) + n_{i-1}M_{i-1}(y_i-y_{i-1}) \qquad (A.12)$$

and rearranging,

$$n_{i-1}u_{i-1}N_{i-1}Y_i - n_{i-1}u_{i-1}N_{i-1}Y_{i-1}-n_{i-1}u_{i-1}M_{i-1}Z_i + n_{i-1}u_{i-1}M_{i-1}Z_{i-1} - n_{i-1}M_{i-1}y_i + n_{i-1}M_{i-1}y_{i-1} = 0 \qquad (A.13)$$

Now rearranging (A.2) and multiplying by y_i,

$$(n_iM_i - n_{i-1}M_{i-1})y_i\gamma_i = (n_iN_i - n_{i-1}N_{i-1})y_i\beta_i \qquad (A.14)$$

Substituting for β_i and γ_i from (A.9)

$$(n_iM_i - n_{i-1}M_{i-1})y_i\cdot 2(c_iZ_i-1) = (n_iN_i - n_{i-1}N_{i-1}) \cdot y_i(f_{Yi} + 2c_iY_i) \qquad (A.15)$$

Using (A.5) to substitute for,

$$n_i y_i c_i = B_i - n_i u_i \tag{A.16}$$

and

$$n_{i-1} y_i c_i = B_i - n_{i-1} u_{i-1} \tag{A.17}$$

will give,

$$n_{i-1}u_{i-1}N_{i-1}Y_i - n_i u_i N_i Y_i - n_{i-1}u_{i-1}M_{i-1}Z_i + n_i u_i M_i Z_i - n_{i-1}M_{i-1}y_i + n_i M_i y_i = B_i\{-Y_i(N_i-N_{i-1})+Z_i(M_i-M_{i-1})\}-\frac{y_i f_{Y_i}}{2}(n_i N_i - n_{i-1}N_{i-1}) \tag{A.18}$$

Subtracting (A.18) from (A.13) we can see that half the terms on the left-hand side cancel out,

$$(n_i u_i N_i Y_i - n_{i-1}u_{i-1}N_{i-1}Y_{i-1})-(n_i u_i M_i Z_i - n_{i-1}u_{i-1}M_{i-1}Z_{i-1}) -(n_i M_i y_i - n_{i-1}M_{i-1}y_{i-1}) = B_i\{Y_i(N_i-N_{i-1})-Z_i(M_i-M_{i-1})\}+\frac{y_i f_{Y_i}}{2}(n_i N_i - n_{i-1}N_{i-1}) \tag{A.19}$$

Now summing over all the surfaces, i=1 to i=K+1 where K+1 refers to the image plane and 0 to the object plane,

$$(n_{K+1}u_{K+1}N_{K+1}Y_{K+1} - n_o u_o N_o Y_o)-(n_{K+1}u_{K+1}M_{K+1}Z_{K+1}-n_o u_o M_o Z_o) -(n_{K+1}M_{K+1}y_{K+1} - n_o M_o y_o) = \sum_{i=1}^{K+1} B_i\{Y_i(N_i-N_{i-1})-Z_i(M_i-M_{i-1})\}+\frac{y_i f_{Y_i}}{2}(n_i N_i - n_{i-1}N_{i-1}) \tag{A.20}$$

Since there is no refraction at the image plane then,

$$n_{K+1} = n_K, u_{K+1} = u_K, N_{K+1} = N_K, M_{K+1} = MK \qquad (A.21)$$

Further, the object and image surfaces are planes, so

$$Z_{K+1} = Z_0 = 0 \qquad (A.22)$$

and the paraxial marginal ray defined by (y_i, u_i) has,

$$y_{K+1} = y_0 = 0 \qquad (A.23)$$

Hence (A.20) reduces to,

$$n_K u_K N_K Y_{K+1} - n_0 u_0 N_0 Y_0 = \sum_{i=1}^{K} B_i\{Y_i(N_i - N_{i-1}) - Z_i(M_i - M_{i-1})\} + \frac{y_i f_{Yi}}{2}(n_i N_i - n_{i-1} N_{i-1}) \qquad (A.24)$$

Note that the upper summation limit is now K. Now subtract the term $n_K u_K N_K \bar{y}_{K+1}$ from both sides of (A.24) and writing the Y-component of the transverse ray aberration as,

$$T_Y = Y_{K+1} - \bar{y}_{K+1} \qquad (A.25)$$

$$n_K u_K N_K T_Y = n_0 u_0 N_0 Y_0 - n_K u_K N_K \bar{y}_{K+1} + \sum_{i=1}^{K} B_i\{Y_i(N_i - N_{i-1}) - Z_i(M_i - M_{i-1})\} + \frac{y_i f_{Yi}}{2}(n_i N_i - n_{i-1} N_{i-1}) \qquad (A.26)$$

If we paraxialize this,

$$T_Y \to 0,\ N_i \to 1,\ Z_i \to 0,\ f_{Yi} \to 0$$

and (A.26) reduces to,

$$0 = n_0 u_0 Y_0 - n_K u_K \bar{y}_{K+1} \qquad (A.27)$$

or

$$n_0 u_0 Y_0 = n_K u_K \bar{y}_{K+1} \qquad (A.28)$$

Here Y_0 is equal to its paraxial value ($\neq y_0$ which refers to a different ray altogether) since the finite and paraxial rays are to originate from the same object point.

Now from (A.6),

$$H = n_0(\bar{y}_0 u_0 - y_0 \bar{u}_0) \tag{A.29}$$

But $y_0 = 0$ so

$$H = n_0 u_0 \bar{y}_0 \tag{A.30}$$

$$= n_0 u_0 Y_0 \cdot \frac{\bar{y}_0}{Y_0} \tag{A.31}$$

If $\bar{y}_0$ is the maximum ray height in the object plane while Y_0 the actual ray height considered, then

$$\eta = \frac{Y_0}{\bar{y}_0} \tag{A.32}$$

is the normalized field (height) so,

$$\eta H = n_0 u_0 Y_0 \tag{A.33}$$

Using (A.28) and (A.33) in (A.26),

$$n_K u_K N_K T_Y = \eta H(N_0 - N_K) + \sum_{i=1}^{K} B_i\{Y_i(N_i - N_{i-1}) - Z_i(M_i - M_{i-1})\}$$

$$+ \frac{y_i f_{Yi}}{2}(n_i M_i - n_{i-1} N_{i-1}) \tag{A.34}$$

But we can write,

$$\eta H(N_0 - N_K) = -\sum_{i=1}^{K} \eta H(N_i - N_{i-1}) \tag{A.35}$$

so the final result is,

$$n_K u_K N_K T_Y = \sum_{i=1}^{K} (B_i Y_i - \eta H)(N_i - N_{i-1}) - B_i Z_i (M_i - M_{i-1})$$

$$+ \frac{y_i f_{Yi}}{2} (n_i N_i - n_{i-1} N_{i-1}) \qquad (A.36)$$

For the X-component we have a similar expression. Since the object point can always be considered to lie in the meridonal plane the field in this expression will be zero ($\eta \equiv 0$),

$$n_K u_K N_K T_X = \sum_{i=1}^{K} B_i \{X_i (N_i - N_{i-1}) - Z_i (L_i - L_{i-1})\}$$

$$+ \frac{y_i f_{Xi}}{2} (n_i N_i - n_{i-1} N_{i-1}) \qquad (A.37)$$

For an axial object point $\eta = 0$ and (A.36) and (A.37) reduce to the same form, equation (4.14).

Although the second term containing f_{Yi} (or f_{Xi}) will be absent for a spherical surface, it should be noted that the first term is not due entirely to spherical surfaces. There will, in general, be some aspheric contribution arising out of the first term. For example, in calculating the fifth-order aberration, $T_Y^{(5)}$, one of the terms on the right-hand side will be $B_i Z_i^{(4)} (M_i^{(1)} - M_{i-1}^{(1)})$ and since $Z_i^{(4)} = (c_i^3/8 + a_4)(Y_i^{(1)})^4 + c_i Y_i^{(1)} Y_i^{(3)}$, the aspheric term is introduced via the a_4.

APPENDIX B

RUNGE-KUTTA SOLUTION OF FIRST-ORDER DIFFERENTIAL EQUATIONS

The Runge-Kutta method for numerically solving differential equations is one of the most popular techniques. The method for solving an equation of the form,

$$\frac{dZ}{dY} = f(Y,Z) \tag{B.1}$$

is based on a simulation of the Taylor series,

$$Z(Y+h) = Z(Y) + h \left.\frac{dZ}{dY}\right|_Y + \frac{h^2}{2!} \left.\frac{d^2Z}{dY^2}\right|_Y + \ldots. \tag{B.2}$$

by replacing the higher order derivatives with a combination of the first derivative at several points. It was found completely satisfactory for the application required in Chapter 3.

With initial conditions (Y_0,Z_0) equation (B.1) is solved using the following scheme. Calculate for i=0,1,2,3,...

$$K_1 = f(Y_i,Z_i) \tag{B.3}$$

$$K_2 = f(Y_i+\tfrac{1}{2}h,\ Z_i+\tfrac{1}{2}hK_1) \tag{B.4}$$

$$K_3 = f(Y_i+\tfrac{1}{2}h,\ Z_i+\tfrac{1}{2}hK_2) \tag{B.5}$$

$$K_4 = f(Y_i+h,\ Z_i+hK_3) \tag{B.6}$$

where h is the step length of Y. The next point is then,

$$Y_{i+1} = Y_i+h \tag{B.7}$$

$$Z_{i+1} = Z_i + \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4) \tag{B.8}$$

By decreasing the size of h any accuracy can be obtained, though a greater number of steps are required to span a given interval of integration. The greatest disadvantages appear to be that an estimate of the error and a procedure for checking are very difficult.

The above scheme is a "fourth-order" method, meaning that the error is of $O(h^4)$. A proof will not be given here but can be found (more often a second-order derivation) in most books on numerical analysis or differential equations. See Morris (24) for example.

The corresponding scheme for the solution of a pair of simultaneous equations,

$$\frac{dZ}{dY} = f(Y,Z,Z') \tag{B.9}$$

$$\frac{dZ'}{dY} = g(Y,Z,Z') \tag{B.10}$$

is a natural extension of equations (B.3) to (B.8).

$$K_{Z_1} = f(Y_i, Z_i, Z_i') \tag{B.11}$$

$$K'_{Z_1} = g(Y_i, Z_i, Z_i') \tag{B.12}$$

$$K_{Z_2} = f(Y_i+\tfrac{1}{2}h,\ Z_i+\tfrac{1}{2}hK_{Z_1},\ Z_i'+\tfrac{1}{2}hK'_{Z_1}) \tag{B.13}$$

$$K'_{Z_2} = g(Y_i+\tfrac{1}{2}h,\ Z_i+\tfrac{1}{2}hK_{Z_1},\ Z_i'+\tfrac{1}{2}hK'_{Z_1}) \tag{B.14}$$

$$K_{Z_3} = f(Y_i+\tfrac{1}{2}h,\ Z_i+\tfrac{1}{2}hK_{Z_2},\ Z_i'+\tfrac{1}{2}hK'_{Z_2}) \tag{B.15}$$

$$K'_{Z_3} = g(Y_i+\tfrac{1}{2}h,\ Z_i+\tfrac{1}{2}hK_{Z_2},\ Z_i'+\tfrac{1}{2}hK'_{Z_2}) \tag{B.16}$$

$$K_{Z_4} = f(Y_i+h,\ Z_i+hK_{Z_3},\ Z_i'+hK_{Z_3}') \tag{B.17}$$

$$K_{Z_4}' = g(Y_i+h,\ Z_i+hK_{Z_3},\ Z_i'+hK_{Z_3}') \tag{B.18}$$

and,

$$Y_{i+1} = Y_i + h \tag{B.19}$$

$$Z_{i+1} = Z_i + \frac{h}{6}(K_{Z_1} + 2K_{Z_2} + 2K_{Z_3} + K_{Z_4}) \tag{B.20}$$

$$Z_{i+1}' = Z_i' + \frac{h}{6}(K_{Z_1}' + 2K_{Z_2}' + 2K_{Z_3}' + K_{Z_4}') \tag{B.21}$$

The initial conditions are $(Y_0,\ Z_0,\ Z_0')$.

APPENDIX C

FORTRAN LISTING OF ROUTINES THAT WILL CALCULATE THE ABERRATION COEFFICIENTS, $T^{(i)}$

Listed here are three subroutines that program the calculation given in section(4.4) for the aberration coefficients. They were taken from a much larger program so that some of the declarative statements will bear no relevance to the calculation itself.

The routines are named INTI, FIRST and HIGHER. INTI sets the initial conditions and therefore only needs to be called once. FIRST calculates first (paraxial) and second order terms and HIGHER the terms of order 3 and higher. The reader should not have much trouble relating the list of equations in section (4.4) to the corresponding statements.

The meaning of the indices in the arrays should be clear from a few examples; $c_i \equiv$C(I), $\beta_i^{(J)} \equiv$ BETA(I,J). $(Y_i^l)^{(J)}$ $\equiv$Y(I,J,L); ie the first represents the surface or medium, the second the order and the third the power. An exception is $T^{(J)} \equiv$ TCOEF(J).

The 2-D arrays are dimensioned in the second order index to 13, so the maximum order that can be placed in the array is 25 or 26, depending on whether it is an odd or even array respectively. Thus up to 13 of the aspheric coefficients, a_{Ji}, can be found.

Also note that in the text the subscript 0 was used to denote the object space. Since the arrays do not have a zeroth location we store n_i in RI(I+1) for example, and to be consistent c_i in c(I+1) etc.

FIRST (and HIGHER) also include a calculation of wavefront aberration terms and defocussing effects given in section (5.2).

$$T^{(J)} \equiv \text{TCOEF(J)},\ T_i^{(J)} \equiv \text{T(I,J)},\ D^{(J)} = \text{TDEF(J)}$$

MINOR and MAXOR (line 5910 in higher) represent the minimium and maximium orders to be calculated respectively, and may be in the range.

$$3 \leqslant \text{MINOR} \leqslant \text{MAXOR} \leqslant 25$$

All terms whose order is less than MINOR must be known before HIGHER is called.

```
2720 SUBROUTINE INT1
2730 REAL M(7,13),N(7,13)
2740 COMMON/ARRAYS/Y(7,26,13),M,N,Z(7,13),ASP(7,13),BETA(7,13),GAMMA(7,13),
2750+ZY(7,13),FY(7,13),T(7,13),TCOEF(13),WCOEF(13),TDEF(13),WDEF(13),
2760+COEF(13),TN(13),TA1(13),TA2(13)
2770 COMMON/ONE/C(7),D(7),RI(7),TYPM(7),A(7),DFC(6,6)
2780 COMMON/SINGLE/MINOR,MAXOR,MXST,NSURF NSF1,NSEPV,NTN,ID,PIF,WVL,AN4,AN8,
2790+RMX,RRSPH,DEXIT,FL,BFL,SCF
2800C ** SET INITIAL CONDITIONS **
2810 DO 757 I=1,13
2820 I2=2*I-1
2830 Y(1,I2,1)=M(1,I)=Z(1,I)=N(1,I)=0.0
2840 757 CONTINUE
2850 Y(1,1,1)=RMX
2860 D(1)=0.0
2870 RETURN
2880 END
```

```
5500 SUBROUTINE FIRST
5510 REAL M(7,13),N(7,13)
5520 COMMON/ARRAYS/Y(7,26,13),M,N,Z(7,13),ASP(7,13),BETA(7,13),GAMMA(7,13),
5530+ZY(7,13),FY(7,13),T(7,13),TCOEF(13),WCOEF(13),TDEF(13),WDEF(13),
5540+COEF(13),TN(13),TA1(13),TA2(13)
5550 COMMON/ONE/C(7),D(7),RI(7),TYPM(7),A(7),DFC(6,6)
5560 COMMON/SINGLE/MINOR,MAXOR,MXST,NSURF,NSF1,NSEPV,NTN,ID,PIF,WVL,AN4,AN8,
5570+RMX,RRSPH,DEXIT,FL,BFL,SCF
5580C ** CALCULATE FIRST AND SECOND ORDER QUANTITIES **
5590C
5600 DO 900 I=2,NSF1
5610 Y(I,1,1)=Y(I-1,1,1)+D(I-1)*M(I-1,1)
5620 Y(I,2,1)=Y(I,1,1)*Y(I,1,1)
5630 Z(I,1)=C(I)*Y(I,2,1)/2.0
5640 ZY(I,1)=C(I)*Y(I,1,1)
5650 BETA(I,1)=-2.0*ZY(I,1)
5660 M(I,1)=(RI(I-1)*M(I-1,1)+BETA(I,1)*(RI(I)-RI(I-1))/2.0)/RI(I)
5670 GAMMA(I,1)=-2.0*C(I)*Z(I,1)
5680 N(I,1)=-M(I,1)*M(I,1)/2.0
5690 T(I,1)=FY(I,1)=0.0
5700 A(I)=RI(I)*(C(I)*Y(I,1,1)+M(I,1))
5710 900 CONTINUE
5720 BFL=-Y(NSF1,1,1)/M(NSF1,1)
5730 FL=-Y(1,1,1)/M(NSF1,1)
5740 RRSPH=BFL-DEXIT
5750 TCOEF(1)=WCOEF(1)=0.0
5760 TDEF(1)=TN(1)=M(NSF1,1)
5770 WDEF(1)=-RI(NSF1)*RMX*M(NSF1,1)/(2.0*RRSPH)
5780 RETURN
5790 END
```

```
5810 SUBROUTINE HIGHER
5820 REAL M(7,13),N(7,13)
5830 COMMON/ARRAYS/Y(7,26,13),M,N,Z(7,13),ASP(7,13),BETA(7,13),GAMMA(7,13),
5840+ZY(7,13),FY(7,13),T(7,13),TCOEF(13),WCOEF(13),TDEF(13),WDEF(13),
5850+COEF(13),TN(13),TA1(13),TA2(13)
5860 COMMON/ONE/C(7),D(7),RI(7),TYPM(7),A(7),DFC(6,6)
5870 COMMON/SINGLE/MINOR,MAXOR,MXST,NSURF,NSF1,NSEPV,NTN,ID,PIF,WVL,AN4,AN8,
5880+RMX,RRSPH,DEXIT,FL,BFL,SCF
5890 INTEGER P,Q,S
5900C ** CALCULATE HIGHER ORDER QUANTITIES **
5910 DO 950 S=MINOR,MAXOR,2
5920 IS=(S+1)/2
5930 DO 1000 I=2,NSF1
5940 V1=0.0
5950 DO 100 P=1,S-2,2
5960 IP=(P+1)/2
5970 ISP=(S-P)/2
5980 V1=V1+M(I-1,IP)*(Z(I,ISP)-Z(I-1,ISP))-(Y(I,P,1)-Y(I-1,P,1))*N(I-1,ISP)
5990 100 CONTINUE
6000 Y(I,S,1)=Y(I-1,S,1)+D(I-1)*M(I-1,IS)+V1
6010C
6020 DO 101 Q=3,S,2
6030 IQ=(Q+1)/2
```

```
6040 Y(I,S,IQ)=0.0
6050 DO 101 P=1,S-Q+1,2
6060 Y(I,S,IQ)=Y(I,S,IQ)+Y(I,P,1)*Y(I,S-P,IQ-1)
6070 101 CONTINUE
6080 DO 102 Q=2,S+1,2
6090 IQ=Q/2
6100 Y(I,S+1,IQ)=0.0
6110 DO 102 P=1,S-Q+2,2
6120 Y(I,S+1,IQ)=Y(I,S+1,IQ)+Y(I,P,1)*Y(I,S+1-P,IQ)
6130 102 CONTINUE
6140C
6150 ZY(I,IS)=0.0
6160 DO 103 P=2,S+1,2
6170 IP=P/2
6180 ZY(I,IS)=ZY(I,IS)+FLOAT(P)*(ASP(I,IP)+COEF(IP)*C(I)**(P-1))*Y(I,S,IP)
6190 103 CONTINUE
6200C
6210 V2=0.0
6220 DO 104 P=1,S-2,2
6230 ISP=(S-P)/2
6240 IP=(P+1)/2
6250 V2=V2+ZY(I,IP)*Z(I,ISP)
6260 104 CONTINUE
6270 FY(I,IS)=2.0*(ZY(I,IS)-C(I)*Y(I,S,1)-C(I)*V2)
6280C
6290 BETA(I,IS)=-2.0*C(I)*Y(I,S,1)-FY(I,IS)
6300C
6310 V3=0.0
6320 DO 105 P=1,S-2,2
6330 ISP=(S-P)/2
6340 IP=(P+1)/2
6350 V3=V3+BETA(I,IP)*(RI(I)*N(I,ISP)-RI(I-1)*N(I-1,ISP))-
6360+(RI(I)*M(I,IP)-RI(I-1)*M(I-1,IP))*GAMMA(I,ISP)
```

```
6370 105 CONTINUE
6380 M(I,IS)=(2.0*RI(I-1)*M(I-1,IS)+BETA(I,IS)*(RI(I)-RI(I-1))+V3)/(2.0*RI(I))
6390C
6400 IF(S.EQ.MAXOR) GOTO 1000
6410C
6420 Z(I,IS)=0.0
6430 DO 106 P=2,S+1,2
6440 IP=P/2
6450 Z(I,IS)=Z(I,IS)+(ASP(I,IP)+COEF(IP)*C(I)**(P-1))*Y(I,S+1,IP)
6460 106 CONTINUE
6470C
6480 GAMMA(I,IS)=-2.0*C(I)*Z(I,IS)
6490C
6500 V4=0.0
6510 DO 107 P=1,S,2
6520 IP=(P+1)/2
6530 IPP=(S-P+2)/2
6540 V4=V4+M(I,IP)*M(I,IPP)
6550 107 CONTINUE
6560 V5=0.0
6570 DO 108 P=2,S-1,2
6580 IP=P/2
6590 IPP=(S-P+1)/2
6600 V5=V5+N(I,IP)*N(I,IPP)
6610 108 CONTINUE
6620 N(I,IS)=-(V4+V5)/2.0
6630C
6640 1000 CONTINUE
6650 V7=0.0
6660 TN(IS)=M(NSF1,IS)
6670 DO 112 P=1,S-2,2
6680 IP=(P+1)/2
6690 IPP=(S-P)/2
```

```
6700 V7=V7+TDEF(IP)*N(NSF1,IPP)
6710 TN(IS)=TN(IS)-TN(IP)*N(NSF1,IPP)
6720 112 CONTINUE
6730 TDEF(IS)=M(NSF1,IS)-V7
6740C
6750C
6760C    CALCULATE ABERRATION COEFFICIENTS FOR EACH SURFACE AND THEIR SUM.
6770C
6780 TCOEF(IS)=0.0
6790 DO 1050 I=2,NSF1
6800 V4=V5=0.0
6810 DO 110 P=1,S-2,2
6820 ISP=(S-P)/2
6830 IP=(P+1)/2
6840 V4=V4+A(I)*(Y(I,P,1)*(N(I,ISP)-N(I-1,ISP))-(M(I,IP)-M(I-1,IP))*Z(I,ISP))
6850 V5=V5+Y(I,1,1)*FY(I,IP)*(RI(I)*N(I,ISP)-RI(I-1)*N(I-1,ISP))/2.0
6860+-RI(NSF1)*M(NSF1,1)*N(NSF1,ISP)*T(I,IP)
6870 110 CONTINUE
6880 V6=V4+V5+Y(I,1,1)*FY(I,IS)*(RI(I)-RI(I-1))/2.0
6890 T(I,IS)=V6/(RI(NSF1)*M(NSF1,1))
6900 TCOEF(IS)=TCOEF(IS)+T(I,IS)
6910 V8=-RI(NSF1)*RMX/(RRSPH*FLOAT(S+1))
6920 WCOEF(IS)=V8*TCOEF(IS)
6930 WDEF(IS)=V8*TDEF(IS)
6940C
6950 1050 CONTINUE
6960C
6970C
6980 950 CONTINUE
6990 RETURN
7000 END
```

APPENDIX D

COMA, THE SINE RELATION, THE SINE CONDITION AND OFFENSE AGAINST THE SINE CONDITION (23),(25)

The wavefront aberration for linear coma is of the form,

$$W(X_p,Y_p) = \sum_{J=1,2,.} C_J(X_p^2 + Y_p^2)^J Y_p \eta' \qquad (D.1)$$

where (X_p,Y_p) are the pupil coordinates and η' the field coordinate in the image plane. If R' is the radius of the reference sphere, then the transverse ray aberrations will be,

$$T_X = \frac{-2R'\eta'}{n'} \sum_{J=1,2,.} JC_J X_p Y_p (X_p^2+Y_p^2)^{J-1} \qquad (D.2)$$

$$T_Y = \frac{-R'\eta'}{n'} \sum_{J=1,2,.} C_J(X_p^2 + (1+2J)Y_p^2)(X_p^2 + Y_p^2)^{J-1} \qquad (D.3)$$

For a ray in the sagittal section, $Y_p=0$ and

$$T_X = 0, \quad T_Y=T_{sag}= \frac{-R'\eta'}{n'} \sum_{J=1,2....} C_J X_p^{2J} \qquad (D.4)$$

Such a ray hits the image plane at a different height to the principal ray. Since this height varies with the zone, linear coma can be considered as a zonal variation in magnification, provided we restrict attention to a single azimuth angle.

To evaluate some measure of linear coma we first derive the so called "sine relation". Referring to Fig. (D.1), consider a small object OP of height η imaged to $O'P'$ with

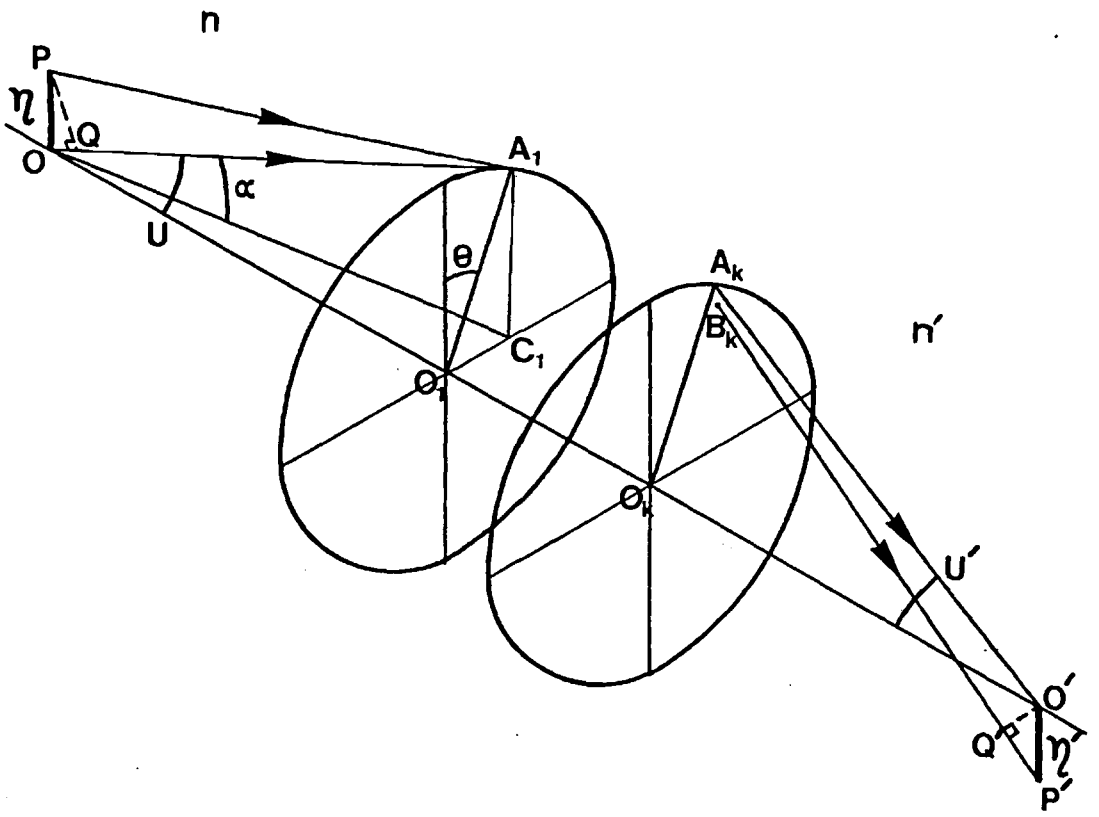


Fig. (D.1) Imaging of a small object OP to O′P′.

height η' by a system of K surfaces. η and η' are assumed to be small enough such that $\eta^2, \eta^3, ..$ etc are negligible. The ray $OA_1 ... A_K O'$ images the axial point and lies in a plane with azimuth angle θ. Its inclination to the axis in object and image space is U and U' respectively. The ray $PA_1 ... B_K P'$ images the point P at the edge of the object. If both points are to be imaged accurately then the optical path lengths along each ray must be equal. It can be shown that,

$$\left[OA_1 ... A_K O'\right] = \left[PA_1 ... B_K P'\right] \qquad (D.5)$$

or,

$$\left[OQ\right] + \left[QA_1 .... A_K O'\right] = \left[PA_1 .... B_K Q'\right] + \left[Q' P'\right] \qquad (D.6)$$

where the points Q and Q' result from dropping a perpendicular from one ray onto the other, as shown. Now if the object is small, these two rays are neighbouring rays and the optical path lengths between these perpendicular lines are equal. PQ and $O' Q'$ just represent a small section of the wavefront between which the optical path length must be the same, i.e.

$$\left[QA_1 .... A_K O'\right] = \left[PA_1 ... B_K Q'\right] \qquad (D.7)$$

Subtracting (D.7) from (D.6) gives,

$$\left[OQ\right] = \left[Q' P'\right] \qquad (D.8)$$

To evaluate OQ, note that

$$OQ = OP \sin \widehat{OPQ} \qquad (D.9)$$

and $\widehat{OPQ} = A_1\widehat{O}C_1 = \alpha$, $OQ = \eta \sin\alpha$ (D.10)

Now,

$$\sin \alpha = \frac{C_1A_1}{\overline{OA_1}} \tag{D.11}$$

and

$$\sin U = \frac{O_1A_1}{\overline{OA_1}} \tag{D.12}$$

Dividing (D.11) by (D.12),

$$\frac{\sin \alpha}{\sin U} = \frac{C_1A_1}{\overline{O_1A_1}} = \cos \theta \tag{D.13}$$

Hence by equation (D.10),

$$\left[OQ\right] = n\eta \sin U \cos\theta \tag{D.14}$$

Similarly for $\left[Q'P'\right]$,

$$\left[Q'P'\right] = n'\eta'\sin U'\cos\theta \tag{D.15}$$

Equating (D.14) and (D.15) according to (D.8) gives the sine relation,

$$n\eta \sin U = n'\eta' \sin U' \tag{D.16}$$

or

$$\eta' = \frac{n\eta \sin U}{n' \sin U'} \tag{D.17}$$

For a given object size, η, the image size produced by rays passing through any zone is given by (D.17). In the presence of linear coma , η' will vary from zone to zone. If coma is absent, however, η' will remain constant and in particular will be equal to the paraxial value given by the Lagrange invariant,

$$\eta' = \frac{H}{n'u'} = \frac{nu\eta}{n'u'} \qquad (D.18)$$

Equating (D.17) and (D.18),

$$\frac{\sin U'}{u'} = \frac{\sin U}{u} \qquad (D.19)$$

This is the "sine condition" and will be satisfied by every zone if linear coma is not present. Furthermore, this is valid only when there is no spherical aberration.

To obtain some measure of the coma that may actually be present we refer to Fig.(D.2). Here, spherical aberration is also taken into account. O_1 is the paraxial image point for an axial object and O_2 the focus for a finite ray with convergence angle U'. The distance O_1O_2 is the longitudinal spherical aberration, LA.

A principal ray from an off-axis point hits the paraxial image plane at P_1 and the defocussed plane at P_2. When coma is present, however, a pair of sagittal rays will not focus at P_2 but slightly above or below it at S. The ray aberration, T_{sag} is given by equation (D.4).

The reason for using sagittal rays rather than tangential is particularly important when considering the wavefront error. For a point $X_p=a$, (D.4) gives T_{sag} as,

$$T_{sag} = \frac{-R'\eta'}{n'} \sum_{J=1,2,.} C_J a^{2J} \qquad (D.20)$$

But for $X_p=0$ and $Y_p=a$, (D.1) gives the wavefront aberration,

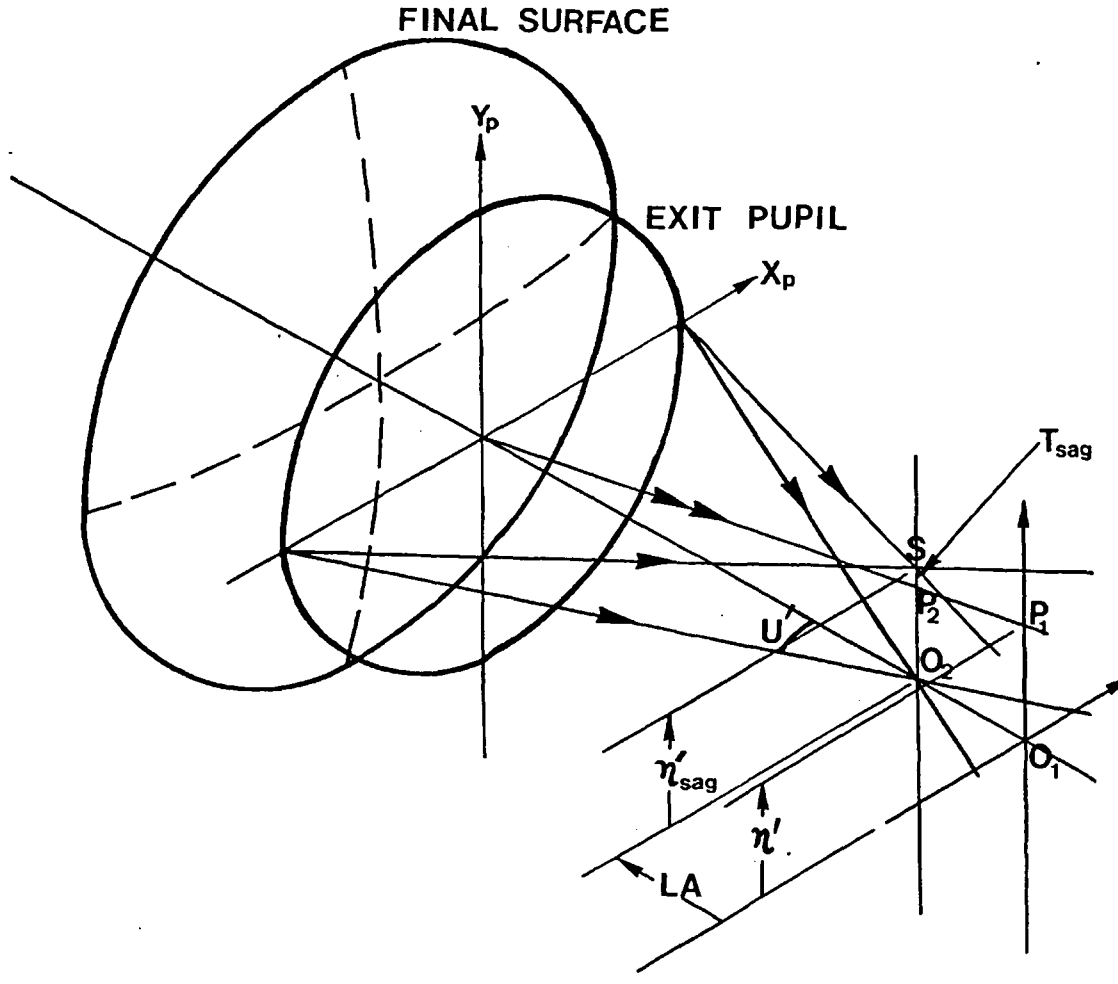


Fig. (D.2) Effect of linear coma. A pair of rays from an axial object point focus at O_2 rather than O_1 due to spherical aberration. A principal ray crosses the planes containing these points at P_2 and P_1 respectively. Due to linear coma, a pair of sagittal rays from the same off-axis object point focus at S rather than P_2; a height T_{sag} above P_2.

$$W(X_p=0,\ Y_p=a) = \eta' \sum_{J=1,2,.} C_J a^{2J+1} \tag{D.21}$$

T_{sag} can then be written,

$$T_{sag} = \frac{-R'}{n'a} W(X_p=0,\ Y_p=a) \tag{D.22}$$

and on multiplying by $\cos\theta$,

$$W(X_p=a\sin\theta,\ Y_p=a\cos\theta) = W(X_p=0,\ Y_p=a)\cos\theta \tag{D.23}$$

$$= \frac{-n'a\ T_{sag}\cos\theta}{R'} \tag{D.24}$$

So the wavefront error can be easily found from tracing a sagittal ray. If $T_{sag} = 0$, $W = 0$ and vice versa. The ray aberration for a tangential ray is instead,

$$T_{tan} = \frac{-R'\eta'}{n'} \sum_{J=1,2,.} C_J(1+2J)a^{2J} \tag{D.25}$$

and the factor of $(1+2J)$ prevents us from relating this to the wavefront error. As a result T_{tan} need not vanish even though T_{sag} and W may.

Here though, we use a quantity that Conrady (9) called the "offense against the sine condition," OSC. This is defined as,

$$OSC = \frac{T_{sag}}{\eta'_{sag} - T_{sag}} \tag{D.26}$$

Now referring again to Fig.(D.2), we see that

$$\frac{O_1P_1}{R'} = \frac{O_2P_2}{R' + LA} \tag{D.27}$$

or

$$\frac{\eta'}{R'} = \frac{\eta'_{sag} - T_{sag}}{R' + LA} \tag{D.28}$$

Using this for T_{sag} in (D.26),

$$OSC = \frac{\eta'_{sag}R'}{\eta'(R' + LA)} - 1 \tag{D.29}$$

Here we now apply the sine relation, (D.16), to the sagittal ray,

$$n\eta \sin U = n'\eta'_{sag} \sin U' \tag{D.30}$$

so,

$$OSC = \frac{n\eta R' \sin U}{n'\eta'(R'+LA)\sin U'} - 1 \tag{D.31}$$

and finally using the Lagrange invariant, (D.18),

$$OSC = \frac{u'}{\sin U'} \cdot \frac{\sin U}{u} \cdot \frac{R'}{(R'+LA)} - 1 \tag{D.32}$$

If OSC = 0 and LA = 0, this reduces to the sine condition, (D.19).

This expresses the off-axis imaging characteristics using data from an axial pencil only. The case of an infinite object conjugate is of interest here and in such circumstances the factor sin U /u becomes Y/y, the ratio of the incident ray heights. Usually the two rays are traced at the same height

so this factor becomes unity,

$$OSC = \frac{u'}{\sin U'} \cdot \frac{R'}{(R'+LA)} - 1 \qquad (D.33)$$

Also, when LA=0,

$$OSC = \frac{u'}{\sin U'} - 1 \qquad (D.34)$$

REFERENCES

(1) C.R.Phipps, S.E.Bodner and J.W.Shearer. Appl. Opt. 14 985 (1975).

(2) W.L.Smith, D.Milan and M.J.Weber. Laser Program Annual Report. Lawrence Livermore Lab. 2.278 (1976).

(3) C.E.Thomas. Appl. Opt. 14 1267 (1975).

(4) A.J.Glass. Laser Program Annual Report. Lawrence Livermore Lab. 234 (1974).

(5) D.J.Nicholas and W.T.Welford. "Fast Optics for the Rutherford Laser Compression Expts." Central Laser Facility Rutherford Lab. (Dec. 1976).

(6) J.L.Emmett, J.Nuckolls and L.Wood. Sci. Amer. 230 24 (1974).

(7) J.E.Howard. Appl. Opt. 16 2764 (1977).

(8) W.A.Kleinhans. Appl. Opt. 15 14 (1976).

(9) A.E.Conrady. "Applied Optics and Optical Design." Dover Vol. 1 (1960).

(10) K.A.Brueckner and J.E.Howard. Appl. Opt. 14 1274 (1975).

(11) E.M.Vaskas. J.Opt. Soc. Amer. 47 669 (1957).

(12) A.Greenbaum, A.J.Glass and J.B.Trenholme. Appl. Opt. 15 2579 (1976).

(13) A.K.Head. Proc. Phys. Soc. (London) 70B 731 (1957).

(14) L.C.Martin. Proc. Phys. Soc. (London) 56 104 (1944).

(15) B.R.Frieden. Appl. Opt. 4 1400 (1965).

(16) P.W.Rhodes and D.L. Shealy. Appl. Opt. 19 3545 (1980).

(17) D.J.Nicholas, C.Pataky and W.T.Welford. Appl. Opt. 17 3368 (1978).

(18) B.Jurek. "Optical Surfaces" Elsevier (1977).

(19) A.Cox. "A system of optical design." Focal Press. (1964).

(20) H.G.Ahlstron. Appl. Opt. 20 1902 (1981).

(21) A.Vasvek. "Optics of thin films" 4.10. North-Holland Pub. Co. (1960).

(22) A.J.Glass and J.B. Trenholme. Laser Program Annual report Lawrence Livermore Lab. 155 (1975).

(23) W.T. Welford. " Aberrations of the symmetric optical system" Academic Press (1974).

(24) W.D. Morris "Differential equations for engineers and applied scientists" McGraw-Hill (1974).

(25) Edited by E.Wolf. "Progress in Optics" North-Holland Pub. Co. Vol. 13 (1976).

ACKNOWLEDGMENTS

I wish to thank Professor W.T. Welford for his advice and encouragement during the last three years and in the preparation of this thesis. Also to Ms. Janice Lewis for undertaking the task of typing the thesis, especially Chapter 4.

Finally I would like to thank the Science Research Council for their financial support.