AN EXPERIMENTAL STUDY OF PROPAGATION IN MULTIMODE OPTICAL WAVEGUIDES USING SPATIALLY INCOHERENT PROBE TECHNIQUES

by

PETER JOHN STEVENS, B.Sc.

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of the Loughborough University of Technology, May, 1976

Supervisor : C. Wilson, M.Sc., Ph.D.

Department of Electronic and Electrical Engineering.

ABSTRACT

The analysis of propagation in waveguides is generally based upon Maxwell's Equations and yields solutions in the form of sets of orthogonal fields called characteristic modes of the waveguide. If numerous modes are excited simultaneously by a single monochromatic source, the intensity distribution measured within the waveguide will be the coherent superposition of all the excited modal fields. The analysis of this coherent superposition is cumbersome for all but a small number of modes. If the source is polychromatic or spatially incoherent the increased complexity of the superposition procedure, which must now consider the relative coherence of modes, suggests that this method is inappropriate.

An alternative basic approach is through the scalar representation of wave propagation known as the geometrical ray theory. This theory is applied to the propagation of polychromatic light in stepped refractive index profile, multimode, optical waveguides. The study is based upon observations of cross sectional variations of intensity in waveguides of this type which appeared to be more appropriately analysed in terms of the ray theory. It is shown that the variations in intensity are caused by microscopic perturbations of the waveguide core from a nominally circular cross section and are only visible when the waveguide is excited by the polychromatic or spatially incoherent source. The dimensions and format of the intensity variations are shown to be simply related to the cross sectional geometry and dimensions of the waveguide and this suggests a useful method of determining the length dependent variations of these parameters.

STATEMENT OF ORIGINAL WORK

Claims to originality are made in the conclusions to certain chapters, and the following is a summary of the work described in this thesis which to the best of the author's knowledge is original.

1. Preparation of single unmounted fibres, Section 2.5.
2. Determination of fibre orientation and angle of incidence, Section 2.8.
3. Entrance aperture diffraction, Sections 4.3.1 and 4.3.2.
4. Ramp refractive index profiles, Sections 4.4.1. and 4.4.2.
5. The remainder of the thesis is concerned with the study of the patterns observed in optical fibres and unless otherwise acknowledged, is thought to be original work.

CONTENTS

CHAPTER 1

1.1 Introduction

There are a number of books[1-5] available containing comprehensive studies of optical waveguides and associated subjects and only those aspects directly related to the subject of this thesis are derived in later chapters.

The rapid progress now being made in research and development of optical communications systems is liable to make any attempt at a comprehensive review of the field incomplete in many aspects. The following references[6,7] indicate general progress in the most important fields.

The rapid development of optical waveguides with characteristics previously thought unobtainable suggests that the potentially most useful waveguides have yet to be developed. At the beginning of this work the most favoured contender for communications applications was a glass waveguide, of core-cladded construction where the core glass has a higher refractive index than the cladding glass. This was the form of waveguide described by Kao and Hockham[8] in their original study of the potential communications applications of optical waveguides, and has received by far the most attention both theoretically and experimentally. The selection of this type of fibre may have been due in part to the earlier studies of Hopkins[9] et al and the subsequent widespread use of fibres of this form for image transfer and remote illumination, embodied in the now well-established subject of 'Fibre Optics'[3-5].

A major achievement of the optical communications waveguide manufacturer has been the reduction of the loss of waveguides from the 1000dB/km normally associated with 'Fibre Optics' applications to the nominal 20dB/km now obtained, sometimes with glass or liquid filled fibres, but more generally with Silica. Lower losses than 20dB/km have been obtained[7], but the significance of this figure is the 2km inter-repeater loss of 40 dB which is the maximum acceptable for a long distance communications application. Any further loss reduction enhances the potential of the waveguides, and clearly establishes their place in future communications networks. The most recent loss figure is quoted in Appendix A.

The low loss silica waveguides are manufactured by the vapour deposition method which in theory allows an arbitrary radial variation of refractive index. This leads naturally to the production of a radially parabolic refractive index waveguide, first developed by the Japanese[10] using a different manufacturing method. This form of waveguide was called by them 'SelFoc' for self focussing. The attraction of this waveguide is its theoretical low modal dispersion[11,12] as opposed to the length dependent modal dispersion of the core cladded waveguide[13]. Low dispersion is essential for digital communications systems to achieve high data rates without intersymbol interference.

The loss of an optical waveguide is associated with its material absorption,whereas the dispersion is a function of its design. The low loss silica fibre with a radially parabolic index profile would appear to form an optimum combination for communications applications. These waveguides are not yet generally available and the work in this thesis is based upon experiments using core-cladded waveguides.

In most physical situations, an understanding of observed phenomena may be obtained on two intellectual levels. A simple approach which contains large approximations, but which gives a macroscopically correct interpretation of the observations will provide a useful conceptual aid for the derivation of a rigorous analysis based on the fundamental laws of physics. The geometrical ray theory of optics is an excellent example of a macroscopic theory. Rigorous optical theory is based upon Maxwell's Equations which in turn are derived, though not necessarily rigorously, from the fundamental laws of physics.

The macroscopic approach is particularly successful for the analysis of optical systems, where the geometrical ray theory provides sufficient analysis for many applications, especially in the design of imaging systems. Its success is directly related to the ratio of the size of the system components to the wavelength of light. As this ratio increases the effects of the assumptions of the macroscopic theory diminish and the accuracy of the theory increases.

The use of the macroscopic theory for analysis in Fibre Optics applications has been successful where the rigorous analysis of the propagation of light within single fibres is not vitally significant. Furthermore the incoherent light or polychromatic (white) light used for imaging and remote illumination is incompatible with a rigorous analysis which considers propagation of monochromatic, spatially coherent waves. A rigorous analysis would have to superpose the solutions for each elementary monochromatic wave emitted by the source. According to Kapany and Burke[1] (Chap. 2, p7) such a procedure is 'generally unnecessary', "because they (the elementary waves) would have randomly related amplitudes and phases and would produce no observable interference effects".

On the basis of experimental observations described in the next section of this chapter, this thesis examines under what conditions observable phenomena are produced by spatially incoherent light when it propagates in optical waveguides of core cladded construction.

There is considerable interest in the geometrical ray interpretation of waveguide theory[14] as applied to both optical and general waveguide problems. Both Kapany[1] and Marcuse[2] use ray analysis of slab waveguides to introduce their rigorous waveguide theory. Gambling[15] and others[16,17] have used, with some success, ray propagation models to calculate pulse dispersion in round core cladded optical waveguides.

However, ray theory is not generally used to calculate the intensity distribution within a waveguide. Such calculations are reserved for the rigorous waveguide analysis in which the individual waveguide modes present at the point of observation are superposed and interfere to produce the observed intensity distribution. The modes present are calculated from a knowledge of the source fields at the entry port of the waveguide and determination of the characteristic waveguide modes excited by such fields. It will be suggested in Chapter 8 that a characteristic mode of a waveguide is a complex field configuration composed of elementary fields which interfere. The propagation of each elementary wave may be represented by a single ray whose propagation obeys the laws of geometrical ray theory. To calculate intensity distributions using ray analysis due account must be taken of any foci or caustics formed where the ray analysis is inappropriate and diffraction theory must be applied.

This thesis examines this apparently complex relationship between the macroscopic and rigorous optical analytical methods as applied to optical waveguides.

1.2 Initial Experiments

This section describes the initial experiments which revealed the unexpected cross sectional variations of intensity in round core cladded optical waveguides when excited by a white light source.

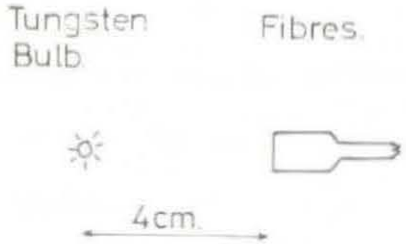


Figure 1a. White light illumination of fibres.

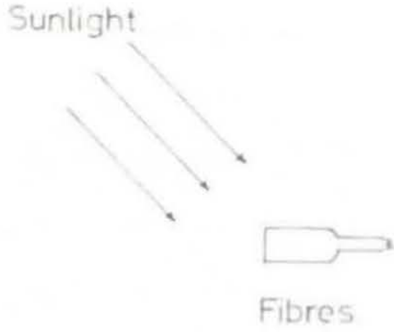


Figure 1b. Sunlight illumination of fibres.

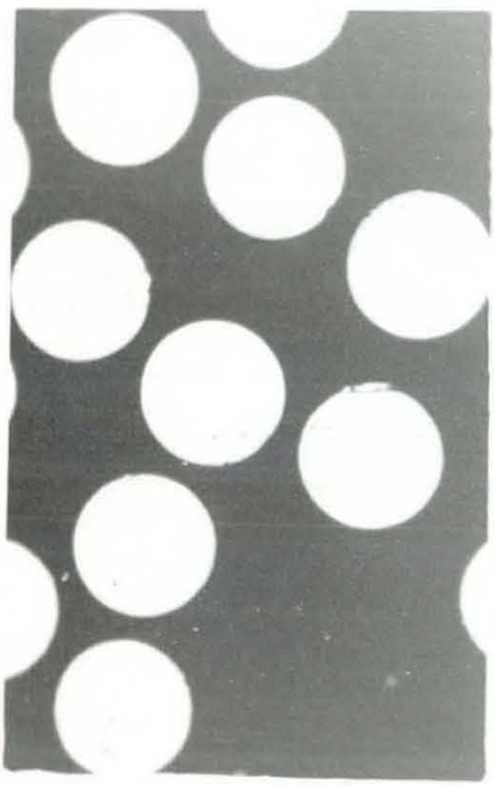


Figure 2a.

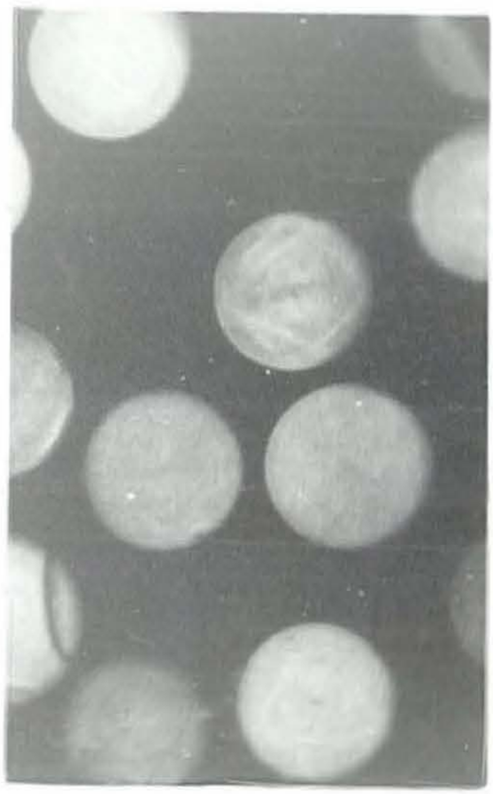


Figure 2b.

Figure 2. Microphotographs of the near field radiation patterns of the ≈ 50μ diameter fibres illuminated as shown in figure 1.

The first observation in company with other observations of classical optical phenomena (Newton's solar spectra) was made with the use of sunlight as a source. At the time the ends of a 50 cm length of commercially available fibre bundle* were being prepared for experimental use in loss measurements. The effectiveness of the polishing process described in Chapter 2 was being examined using a microscopic observation of the polished end of the fibre bundle whilst illuminating the other end of the bundle (already polished) with a white light source positioned as shown in Figure 1a. The illuminated end of the bundle was hand held and a chance movement placed the ends of the fibres in a bright beam of sunlight, Figure 1b. The resultant photomicrograph of the viewed end of the bundle is shown in Figure 2b. This may be compared with Figure 2a which is the photomicrograph corresponding to illumination of the bundle as shown in Figure 1a.

In Figure 2a all the fibres of the bundle appear to be equally illuminated and there is no observable variation of intensity within each fibre cross section. The only apparent difference between fibres is a variation in dimension of the illuminated area which corresponds roughly to the core diameter of each fibre. Such variations are to be expected as a result of the manufacturing method[4 pp.63] in which all fibres are pulled simultaneously from their own preforms of core and cladding glass. Any variations in preform dimensions, pulling rates or furnace temperatures will result in variations of fibre size.

In Figure 2b marked variations in intensity are visible within the cross section of certain fibres. The precise form of the variations and their contrast is seen to vary from fibre to fibre. Of the

*Rank Kershaw "FIBROFLEX", See Appendix B for specification.

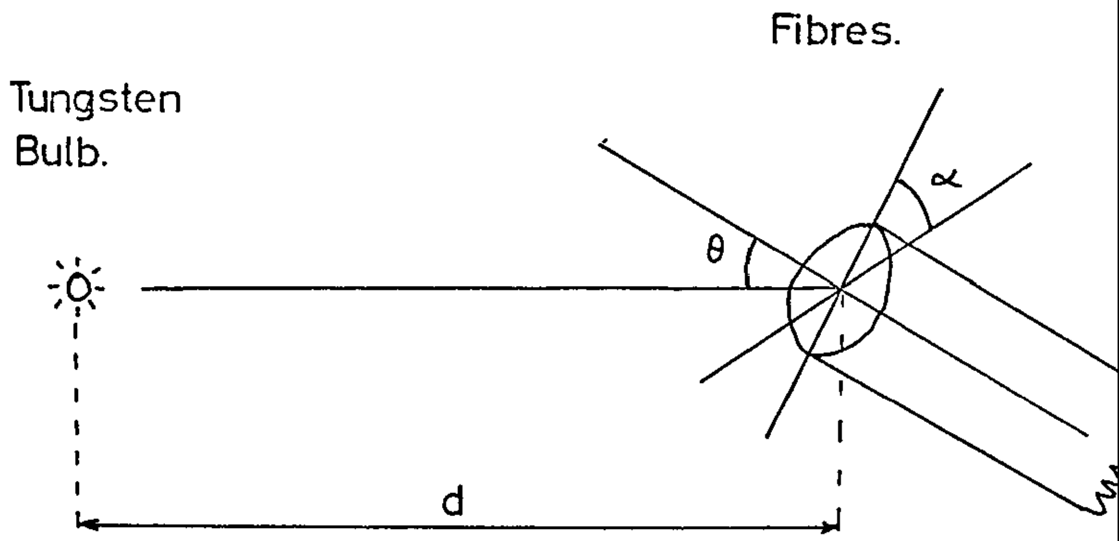


Figure 3. The experimental arrangement for illuminating a fibre bundle with white light.

approximately 400 fibres in the bundle, 220 fibres exhibited good contrast variations, 100 showed poor contrast, the remainder showing no observable variations of intensity within their cross section.

Two questions are posed by these observations. What is the source of the variations of intensity and why does their contrast or visibility vary from fibre to fibre? A literature search revealed that although white light has been used in previous experimental investigations of propagation in optical waveguides[18,19] the current observations have not been reported elsewhere. A paper[20] published during the course of the work independently confirmed certain aspects of the experimental results and also provided a useful impetus for the analytical method derived for the explanation of the observed phenomena.

To aid further investigation, the experimental arrangement of Figure 1b was replaced by that shown in Figure 3. The sun is replaced by a 15 watt tungsten projection bulb positioned a distance $\dot{d}$ from the end of the fibre bundle, such that for $d \geqslant 30$ cms approximately plane polychromatic waves are incident upon the ends of the fibres. These plane waves are incident at an axial angle θ defined as the angle between the normal $\bar{n}$ to the wavefront and the longitudinal axis of the fibres and at azimuth angle α. The azimuth angle is defined as the angle between the longitudinal plane containing the wave normal and an arbitrary fixed longitudinal plane. The normal $\bar{n}$ of a plane wave coincides with the direction of propagation of the wave.

The fibre bundle is clamped in the rotational mount shown in detail in Figure 4. This mount allows independent variation of θ and α for a fixed source position. The two components of θ available on the rotational mount denoted θ_x, θ_y add vectorially to give θ. With θ_y clamped at 0^0 the θ and α values required now correspond linearly to θ_x and α of the rotational mount although there may be an offset required

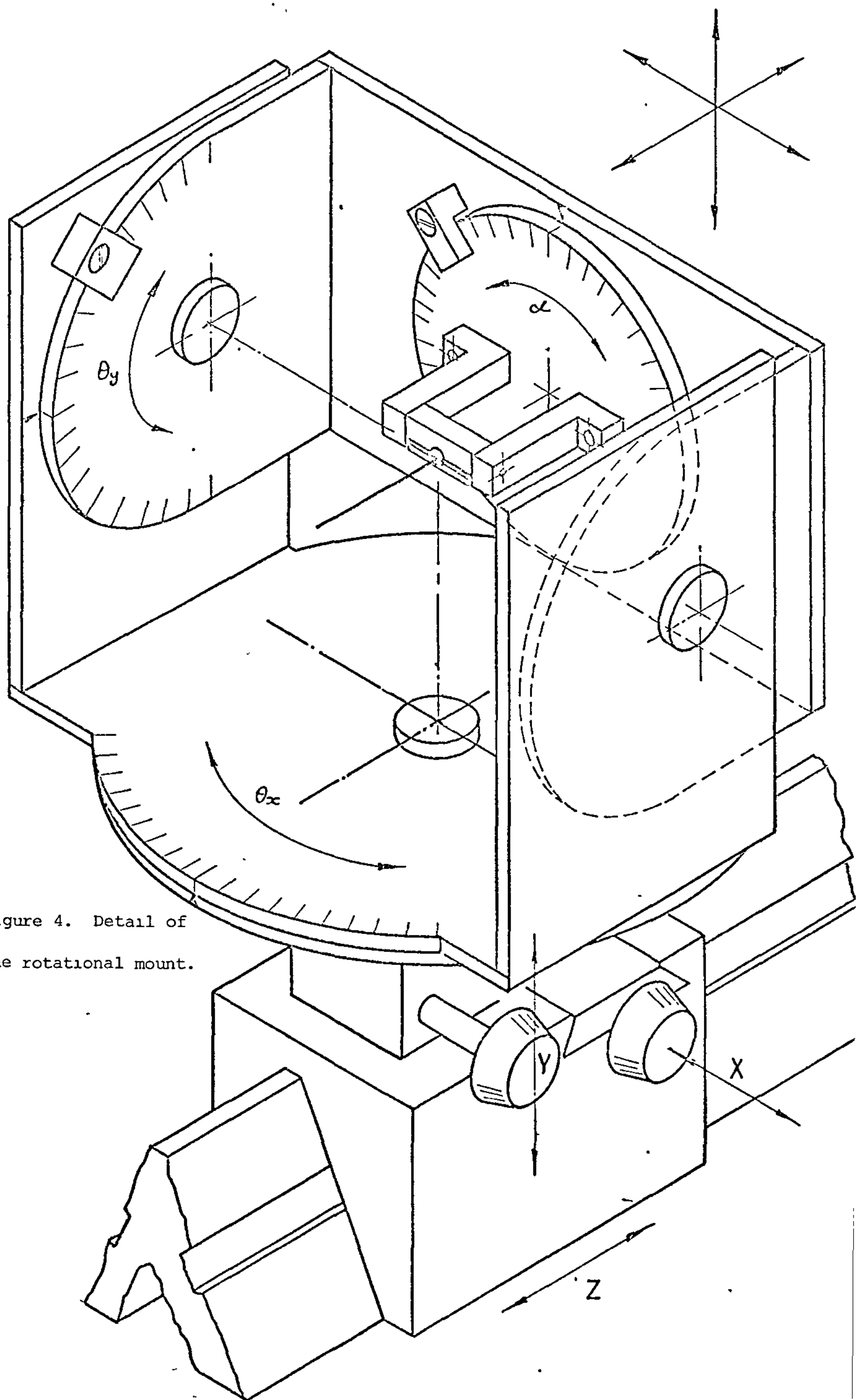


Figure 4. Detail of the rotational mount.

due to the off axis position of individual fibres within the bundle termination. The method of determining the offset required is considered later.

Neglecting for the time being any off set corrections necessary, an indication of any dependence of the "variations of intensity" hereafter referred to as the patterns, upon θ or α was obtained by selecting individual "good contrast" fibres and noting the form and visibility of their patterns as θ and α were varied for a given source distance d. When $d \geq 5$ cms it was found that the form of the patterns had a pronounced α dependence and an increasing visibility for increasing θ when $\theta \geq 10^{\circ}$. When $\theta < 10^{\circ}$ no patterns were visible. When $d < 5$ cms the patterns became generally blurred and for d increasing and greater than 10 cms no dependence upon d was observed. These observations agree with the results of the experimental arrangements of Figure 1. Figure 1a corresponds to $d < 5$ cms where the patterns are generally indistinct whereas Figure 1b corresponds to $d >> 10$ cms and $\theta > 10^{\circ}$, which is an optimum visibility condition.

This next procedure is designed to test the hypothesis that the patterns are a function of the fibre bundle terminations. A single fibre radiating good contrast, well defined patterns,was selected at the viewing end of the bundle. The patterns together with their α dependence were recorded using a video tape recorder as described in Chapter 2. The greatest magnification microscope objective available (x 100) was used in conjunction with the microscope overhead illuminator to illuminate the end of this single fibre only. The crosstalk between this fibre and adjacent fibres was observed to be small by viewing the second end of the bundle through another microscope. The termination was removed from the second end of the fibre bundle and the protective sleeving removed over the whole length of the bundle up to within 2 cm of the other end termination. The single illuminated fibre was now identified by the high intensity of its radiation field. The free end of this single fibre was fixed in the rotational mount and the α dependence of the patterns was tested as before. This end of the fibre was a broken end and the α dependence test was repeated five times with a freshly broken end each time. The form and α dependence of the patterns was found to be unchanged, although some variations of visibility were observed over the five tests.

This single fibre was now broken away from the remaining end termination and both ends ground and polished using the procedures described in Chapter 2 for the preparation of unmounted single fibres. This single fibre now 45 cm in length replaced the bundle shown in Figure 3 and the original θ and α dependence tests were repeated. The results for this unmounted fibre showed no variations from those obtained when it was a member of a bundle. This suggests that the end termination and the other fibres of the bundle play no part in the pattern formation process of an individual fibre. This result, together with the previous observation of very weak dependence upon polished flat fibre ends (from the "broken ends" experiment), suggests that the formation of the patterns is primarily a result of propagation in the fibre. The variability of patterns from fibre to fibre within a bundle suggests variations in the fibre characteristics in addition to the core diameter variations already discussed.

An experimental identification and study of individual fibre patterns is clearly facilitated by the use of robust fibre bundles as opposed to fragile single unmounted fibres. To confirm the equivalence of the bundle mounted fibre to the single unmounted fibre for observation of patterns differing from those of the fibre previously tested, three further fibres were tested. Each fibre was selected to show patterns

Figure 5a. $\alpha = 0^{\circ}$

Figure 5b. $\alpha = 52^{\circ}$

Figure 5c. $\alpha = 60^{\circ}$

Figure 5d. $\alpha = 230^{\circ}$

Figure 5e. $\alpha = 260^{\circ}$

Figure 5. Microphotographs of near field radiation patterns observed when a 40 cm length of cladded fibre was illuminated with white light at an axial angle of incidence of 25° and at azimuthal angles α as indicated.

of different forms whilst in the bundle and each was tested as an unmounted fibre with both ends broken at lengths of 20 cm, 30 cm and 40 cm. No variation in the results previously established was obtained. With prior knowledge of the theoretical results established in later chapters, the fibre patterns selected for detailed description here appear to have been formed as the result of propagation in an elliptic cross section fibre. Since this simple perturbation of a circular cross section is a major feature of the theoretical analysis of the pattern formations and illustrates many of the results achieved, it will form a useful experimental basis for the introduction of the remainder of the thesis.

The experimental arrangement of Figure 3 was used to obtain the photomicrographs shown in Figure 5 of the patterns of the selected fibre. The axial angle θ was fixed at 25° and the azimuthal angle α is indicated against each picture of Figure 5. The pronounced dependence of the patterns upon the azimuthal angle α is clearly demonstrated. The patterns are the regions of high or low intensity which appear within the grey background of the fibre cross section. Figure 5a has a high intensity band forming a diameter, and low intensity regions at selected azimuthal positions and at radii $r/2$ and $3r/4$, where r is the radius of the illuminated core. These low and high intensity regions become high and low intensity regions respectively at other azimuthal positions of the source, for example in Figure 5b where the $3r/4$ radius pattern is now high intensity. The low intensity region corresponding to the high intensity diameter band of Figure 5a is the dark region at the centre of the fibre shown in Figure 5e.

If straight lines are drawn within and along the high intensity patterns and extended to the circumference of the fibre core as shown in Figure 6, simple closed geometrical figures are obtained. The high intensity diameter pattern (Figure 5a) corresponds to a straight line

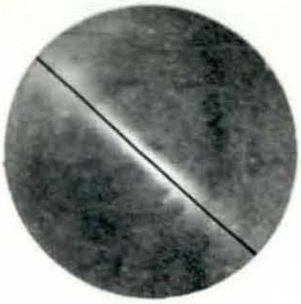


Figure 6a.

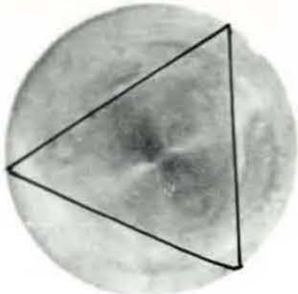


Figure 6b.

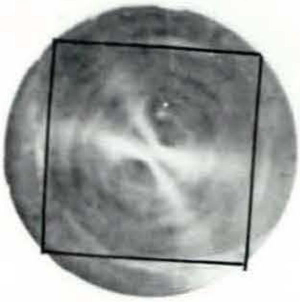


Figure 6c.

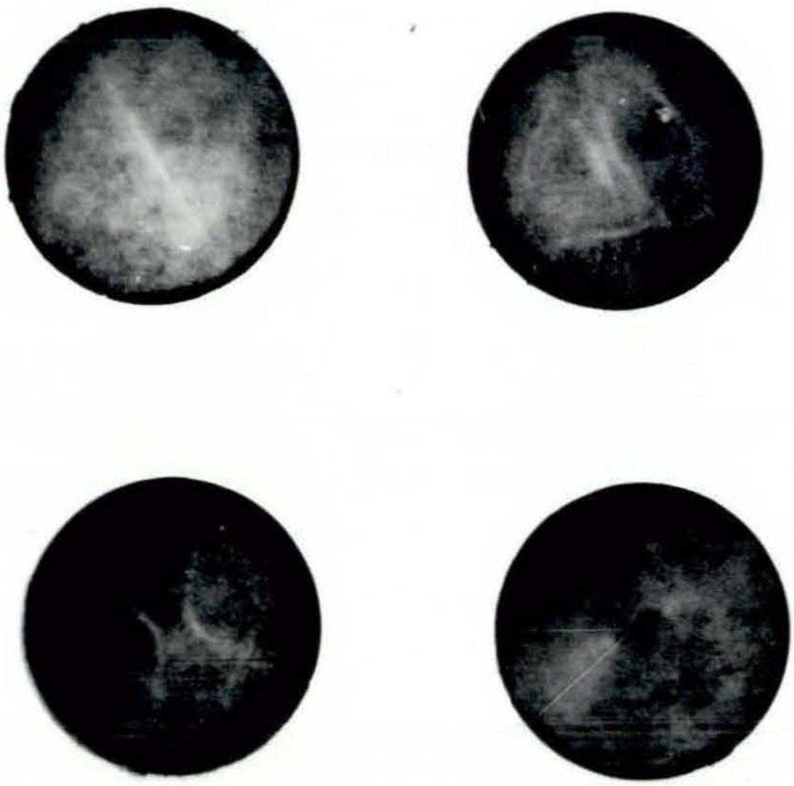


Figure 7. Microphotographs of the near field patterns observed when a 40 cm. length of single fibre was illuminated by white light.

(Figure 6a), the r/2 radius pattern (Figure 5c) produces a triangle (Figure 6b) and the 3r/4 radius patterns (Figure 5b) produces a square (Figure 6c). The similarity between these closed figures and the closed figures of the skew ray paths, which are introduced in Chapter 5, forms the basis of the analysis of the patterns given in Chapter 6. The patterns of other fibres contain components similar in form to those of Figure 5 but which appear in different combinations. This suggests that the analysis based upon the closed figures may be generally applicable to fibres of the core cladded type.

A hypothesis which may be set against this suggestion is that the patterns observed are a unique property of propagation in core cladded fibres manufactured by the method used for the production of these bundles. To test this hypothesis a 40 cm length of single core cladded fibre manufactured by a different process was ground and polished at both ends as an unmounted fibre and tested as before using the experimental arrangement of Figure 3. The observation of the patterns within this fibre, shown in Figure 7 which appear similar to those of Figure 5, suggest that this hypothesis may be discounted. Apart from the obvious difference in the number of fibres pulled simultaneously, the manufacturing processes of the bundle fibres and this single fibre differ in the production of the preform from which the fibres are pulled. The bundle fibre preform consists of a glass rod placed inside a glass tube, whereas the single fibre preform is drawn from two glasses using the double crucible method. The double heating required to produce the single fibre is likely to cause diffusion of the glasses at the core cladding boundary, which in turn may produce a ramp refractive index profile as opposed to the theoretical step profile. The measurement of the characteristics of a ramp profile is discussed in Chapter 4.

1.3 Conclusion

This study is concerned with a previously unreported phenomenon which occurs as the result of propagation of white light in core cladded waveguides. The opening comments of the chapter suggest that although such phenomena may be of little interest to the "fibre optics" user, they may well be of interest to the optical communications engineer.

The potential complexity of analysing these phenomena has meant that each stage of analysis has been accompanied by, or preceded by, experiment. To avoid lengthy diversions into experimental detail in later chapters most of the experimental procedures are presented in the next chapter.

CHAPTER 2

2.1 Introduction

The experimental requirements of this study fall into two categories. The first concerns the preparation of fibre samples. Some considerable time was spent on developing a technique for the precision termination of single unmounted fibres and which is the subject of a short paper[21]. This technique was evolved from the methods developed for terminating fibre bundles using gelatine moulds and the subsequent grinding and polishing of the fibre ends. The problems of terminating and jointing optical waveguides has been the subject of considerable research and a short review of current methods is included.

The second experimental requirement concerns the illumination of the fibre samples and observation of the various phenomena described in the previous and later chapters. The nominal diameter of the fibre core is 50 micron and the observation and measurement of the microscopic phenomena within this diameter poses experimental difficulties with generally expensive solutions. In the absence of any suitable facilities within the department a single capital expenditure was made at the end of the first year to obtain equipment to undertake the study. The limit of the budget was such that novel solutions to certain measurement and observation problems were developed and implemented.

2.2 Review of Fibre Preparation and Jointing Methods

The usual method[3,4] of terminating fibre optic components is to embed the ends of the fibres in a resin cement which when cured is as hard as the glass of the fibres. This forms a continuous surface which may be ground into any desired contour and then polished using the proven lens production procedures[22]. The optical quality of the finished surface depends to a great extent upon the polishing time, the longer the polishing time the higher the quality of the surface.

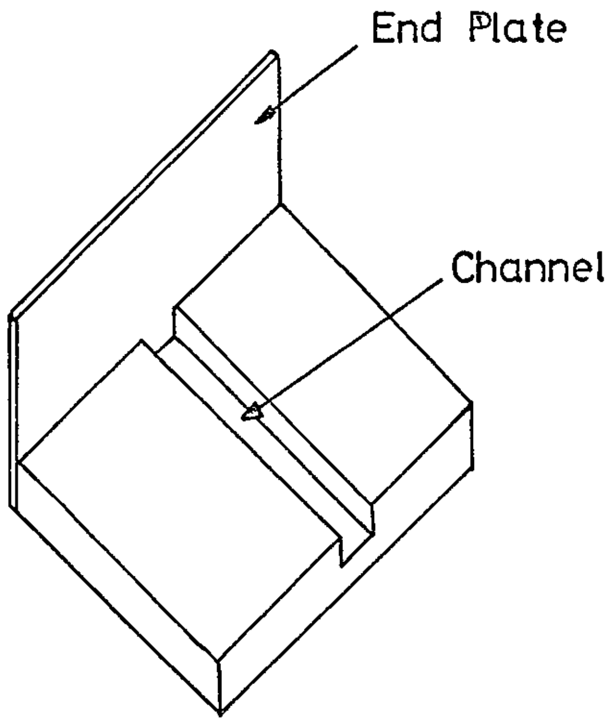


Figure 8. The immersion tank for terminating single fibres.

The usual polishing compound is ferric oxide or jewellers rouge. Emery powders of different grades are used in the grinding process. The two processes differ in that only the grinding stages remove material from the surface whereas polishing is a smoothing operation, hence its time dependent effect. The rough starting surface is progressively ground into the required contour using increasingly finer emery powders. The scratch marks made by each grade of emery are removed by the following finer grade. The polishing process then smooths the scratches of the final grade of emery. A more detailed description of the preparation of optical components is contained in reference 22.

These procedures may also be used to prepare the ends of single fibres. The required length of fibre is cemented into a capillary tube and the conventional grinding and polishing machines are used to produce a flat polished surface at each end of the fibre and tube combination. In communications applications long lengths of fibre are used and a short length of capillary tube is then cemented to each end of the fibre. This method has three major disadvantages :- the fibre length is fixed, the grinding and polishing procedure is laborious and the cementing of a comparatively heavy length of capillary tubing onto the fragile fibre greatly increases the likelihood of accidental breakage of the fibre during routine handling.

Early experimenters using single fibres removed the necessity for any preparation of the ends of the fibre by using the immersion technique shown in Figure 8. The fibre is broken and the rough end placed in the channel, butting up to the glass plate which is then clearly normal to the fibre's longitudinal axis. The channel is filled with a liquid matching the refractive index of the core glass, (for core cladded type fibres). The glass plate now acts as the fibre end and for a lossless termination is of the same refractive index as the core glass.

The cutting of sheet glass is facilitated by the property of the glass to shear along a fault line induced by scoring the surface with a suitably hard point, normally a diamond tipped tool. An adaptation of this procedure is used by Gloge et al[23] in a very successful device for shearing single unmounted fibres to produce fibre ends of unsurpassed optical quality. This is the method now used by most communications laboratories.

The methods for terminating single fibres for use in commercial communications systems are generally based upon the embedding technique, with inbuilt alignment facilities. A typical alignment requirement is perhaps optimistically quoted[24] to be within 5 micron for a 100 micron diameter fibre. To position and fix the fibre within these limits using an initially liquid cement clearly presents considerable engineering problems.

Dalgleish and Lukas report one method[25], though the successful use of the technique outside of the laboratory is doubtful. The unreliability of plug and socket connections and the difficult alignment problems these pose may be side stepped by devising a method for quasi-permanent connections between fibres and components, or the butt joining of fibres. The quasi-permanent connection would be made as a permanent joint, but the methods used would allow the break and remake of the joint with the minimum of effort. Such a method requires a simple alignment procedure such as that described by Someda[26] for the butt joining of fibres, and the application of a quick setting cement. This procedure will be essential anyway for the emergency piecing of optical waveguide cables in the event of breakage during service.

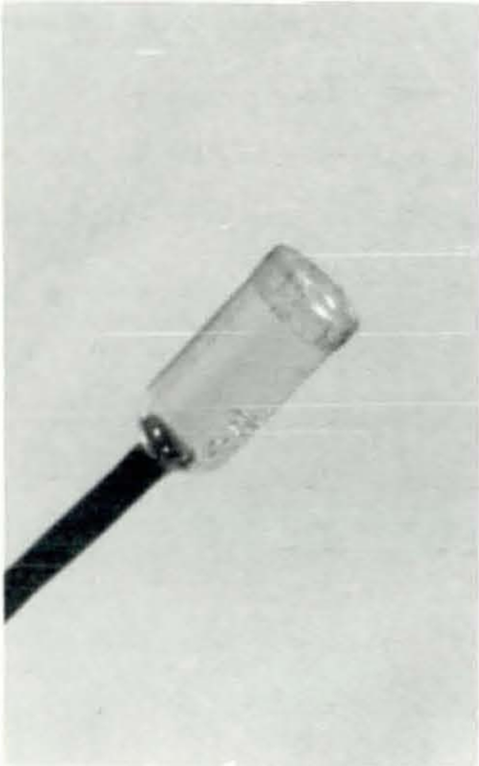


Figure 9. A completed fibre bundle termination.

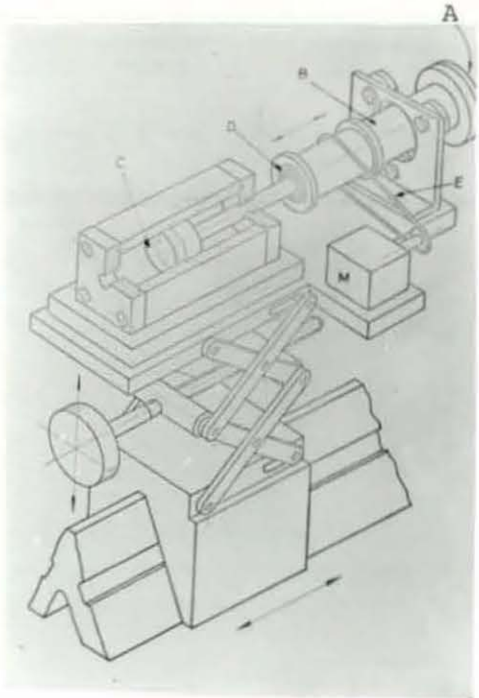


Figure 10a. Schematic of the grinding and polishing machine.

Figure 10b. The grinding and polishing machine.

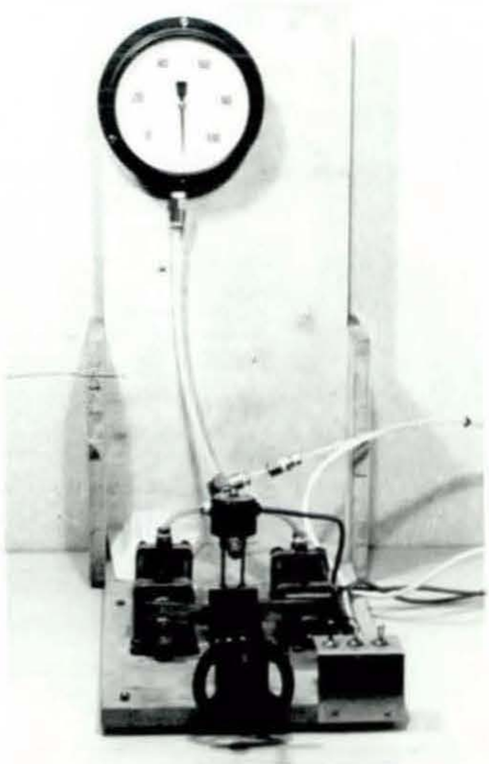


Figure 10c. Manual air pistons.

An alternative to the cementing of the butt (fibre to fibre) joint is to cut the fibre ends using the shear method, position the two ends a predetermined distance apart and then to fuse the two fibres by controlled heating. Bisbee[27] and Dyott[28] report encouraging results using this method. There is an interesting parallel between this method and the braze jointing of co-axial cables. If the terminal and repeater units of the optical communication system are precision fitted with fibre tails during the manufacturing stages, then only butt joining of these tails to the cables would be required in the field.

Most of the jointing experiments described above have used core cladding type fibres. It is not clear whether such procedures will be directly applicable to the low loss, parabolic refractive index, silica fibre described in Chapter 1.

2.3 Development of Fibre Bundle Preparation Methods

The methods used to prepare the ends of the fibre bundles used in this study differ from the usual method described in Section 2.2 in detail only. The moulds used to contain the resin cement around the end of the bundles are 7 mm diameter gelatine pill capsules. The shape of these capsules is ideal for clamping in the rotational mount of Figure 4 for subsequent grinding and polishing. A completed termination is shown in Figure 9.

The absence of any existing facilities encouraged the development of the special purpose grinding and polishing machine shown in Figure 10. The following extract from reference 21 describes its operation with reference to Figure 10a. The various grinding and polishing compounds are mounted on aluminium plates (A). These plates are fixed in turn to the rotating chuck (B). The chuck is mounted on the piston rod of the pneumatic cylinder (C) via two ball races (D). The motor (M) rotates the chuck via the belt drive (E). The plates (A) are thus rotated by the motor (M) and moved along the z axis by the pneumatic piston action. When the plates are in contact with the workpiece, pressure between the two may be adjusted by varying the air pressure in the pneumatic cylinder (C). The motor is a single pole induction motor and is driven by a variable frequency supply. Its speed is variable between 0 and 3000 rpm. The air supply is obtained from a large manually actuated piston, Figure 10b, which forms a closed system with cylinder (C).

Coarse grade emery paper is used for rough grinding and fine diamond paste (as used for the preparation of metallurgical specimens) provides a rapid final grinding compound. When polishing is required the ferric oxide is applied using a felt pad mounted on plate (A). The two rough grinding and two diamond grinding steps required for each termination normally take five minutes. This time includes the time to make four changes of the (A) plates. The polishing time depends upon the finish required and is typically a further five minutes for the bundles used in this study. The resin cement* used is softer than the fibre glass and the diamond grinding leaves the glass fibres proud of the resin casting. The resultant small surface area of the fibre ends alone enjoys enhanced attention from the polishing pad and this may account for the surprisingly short polishing time required for the good optical finish obtained.

An attempt to use a 7 micron grade diamond paste to increase the speed of the rough grinding step was found to cause excessive chipping of the fibre ends. The finer grades of paste tend to pick up particles of glass which also cause chipping and these pastes have to be frequently replaced. However, only minute quantities are required and the cost is negligible.

*CIBA-GIEGY, ARALDITE MY753, HARDENER HY956.

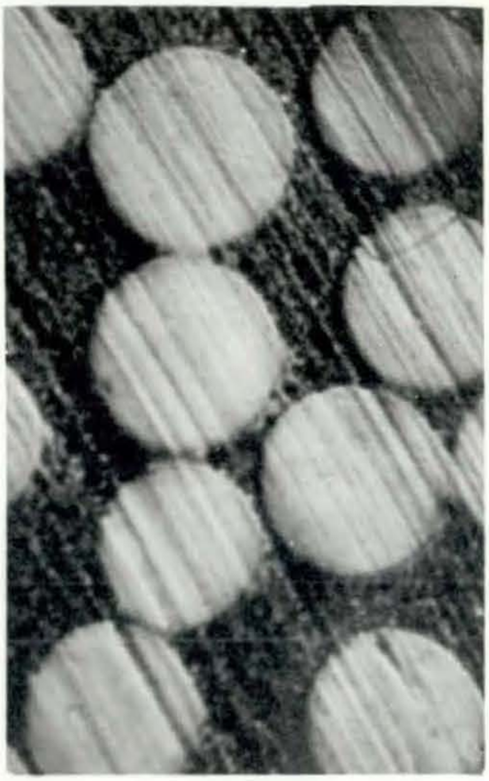


Figure 11a. 5μ grade paste

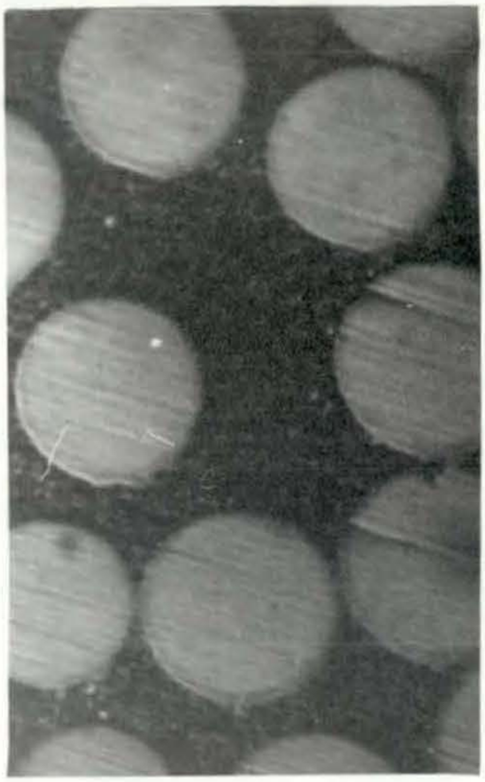


Figure 11b. 3μ grade paste

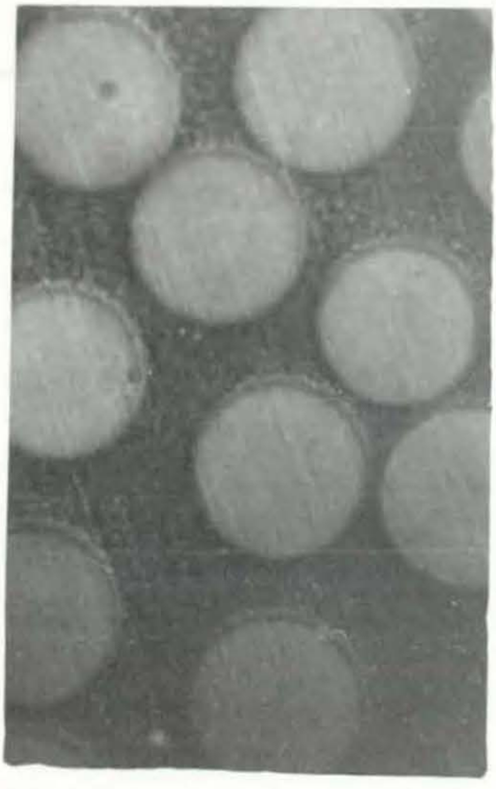


Figure 11c. 1μ grade paste.

Figure 11. Microphotographs of fibre bundle terminations after grinding with diamond grinding pastes of the grades shown.

The completion of each stage of grinding is established by microscopic observation of the fibre ends. The scratch marks are observed using overhead illumination and chipping of the fibre is best observed by illuminating the opposite end of the bundle, (the experimental arrangement of Figure 1a). The fibres should appear as complete "circles" of light and with the scratch marks associated with the previous grade of grinding visible over the whole of the fibre ends. This is illustrated in Figure 11 where the scratch marks for three grades of diamond paste are visible, the middle grade is not normally used.

2.4 Preparation of Single Unmounted Fibres

At the time the method to be described was developed there was no other published method for terminating single fibres without permanently embedding them in resin compounds, with the accompanying disadvantages already discussed. The shearing technique[23] has since provided a method suitable for use in communication fibre laboratories. Unlike the shear method, the current method is not limited to fibres greater in length than the shearing machine (typically 20 cm), and may also terminate single fibres with a sloping end face. The grinding and polishing machine[21] built for preparing fibre bundle terminations is used without modification to prepare the single fibres.

A set of vice jaws was manufactured from an embedding resin casting, cut as shown in Figure 12a. The mould was a gelatine pill capsule as used for the bundle terminations. The fibre to be prepared is passed through the hole in the base of the jaws and then clamped between the jaws by the mounting vice shown in Figure 12b. The mounting vice is secured to the rotational mount of Figure 4 and grinding and polishing proceeds as for the bundle terminations. A fibre end, after the initial diamond paste grinding, is shown in Figure 13a and the curvature of the finished end is

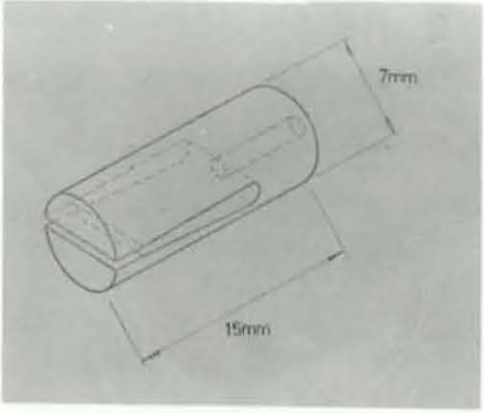


Figure 12a. Vice jaws for holding single fibres during grinding and polishing.

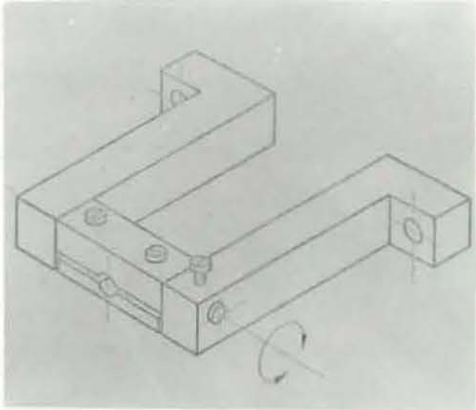


Figure 12b. Vice associated with the jaws shown in figure 12a.

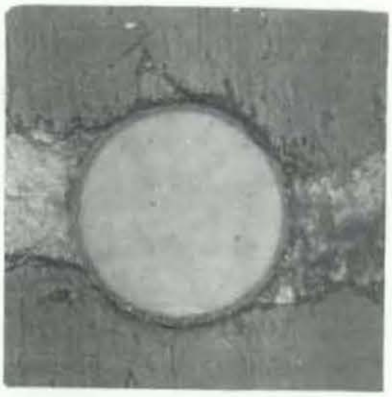


Figure 13a.

Figure 13b. (Overall surface curvature ≈ 1μ)

Figure 13c.

demonstrated in Figure 13b. The fringes were obtained using a standard interference microscope objective and a sodium light source. The fibre was adjusted to give the minimum number of fringes across the end face. The curvature observed is due primarily to the polishing process because as previously discussed, the diamond paste grinding leaves the fibre proud of the softer jaws and the polishing compound tends to round off the edges of the fibre.

The finished fibre is removed from the jaws by relieving the vice pressure until the jaws are well open, pushing the fibre forwards to release it from the jaws and then pulling it out backwards. Any deposit on the end face may be removed by gently swabbing the fibre end with alcohol. The clamping pressure applied to the fibre is critical, excessive pressure encourages chipping of the fibre during grinding and insufficient pressure obviously allows the fibre to slide back beneath the jaw surface. An indication of the vice pressure upon the fibre is obtained by observing the deformation of the jaws adjacent to the fibre as the jaws are closed. The correct pressure is demonstrated in Figure 13a where the fibre is clamped but there is no indentation of the jaws.

2.5 Preparation of Short Fibres

The observation of single reflection caustics and entrance aperture diffraction effects as described in Chapters 4 and 5, requires a prepared fibre of $< 20r$ length, where r is the radius of the fibre core. ($20r \approx 0.5$ mm for the fibres used in this study). The equivalence of the bundle mounted fibre to the single unmounted fibre for the observation of the phenomena of major interest, which was established in Chapter 1, suggested that these short lengths could be produced from thin slices taken from a bundle termination. The procedure is analogue to the microtome technique used in biological and botanical section microscopy.

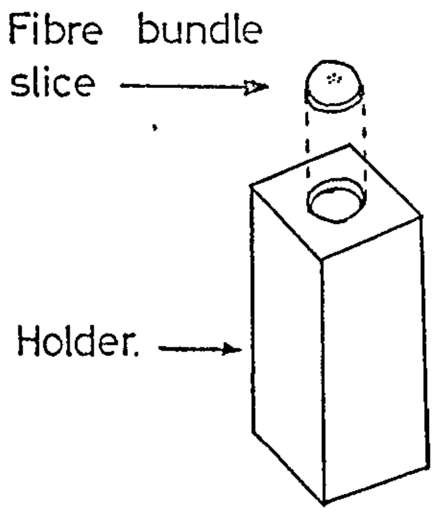


Figure 14. Holding device for short fibre sample preparation.

Eyepiece	Objective.	Scaling. 1 mm =
x15	x90	.878μ
x15	x45	1.76μ
x15	x20	3.96μ

Figure 16. Magnification calibration table.

A thin section (typically 1 mm thickness) is sawn from a prepared end of a fibre bundle and using the holding device shown in Figure 14, the sawn end is hand ground until the section thickness approaches the required length of fibre. The section is then diamond ground and polished by holding the section against the A plates using the device of Figure 14. The shortest length of fibre prepared in this manner is $\approx 12r \approx 0.3$ mm. Below this length the fibres tend to be pulled out of the resin compound by the diamond grinding operation.

A similar procedure for the preparation of short fibre specimens has recently been published[29] in connection with the viewing of refractive index profiles of fibres.

2.6 Microscopic Measuring and Recording Techniques

All the microscopic observations reported in this thesis, with the exception of those made during fibre preparation, were made using the experimental arrangement of Figure 15. The microscope system is based upon an Ealing Beck Epimax microscope sub assembly. The Tetraver Intermediate Unit allows permanent access to the image by the Polaroid Instrument camera and access either to the binocular eyepiece or the T.V. camera. Relevant information about the eyepieces and objectives is given in Table 1. Calculation of the overall magnification of an image is avoided by using the calibration chart shown in Figure 16 where the measured magnification of a calibration graticule for each combination of eyepiece and objective used in the study is given. These figures apply to measurements on photographs taken with the polaroid camera, but the same calibration system is used for the T.V. aided measurements described later.

Focussing is performed by moving the object carrier and its movement is measured using the dial gauge positioned as shown in Figure 17. A measurement is made with reference to the focussed position

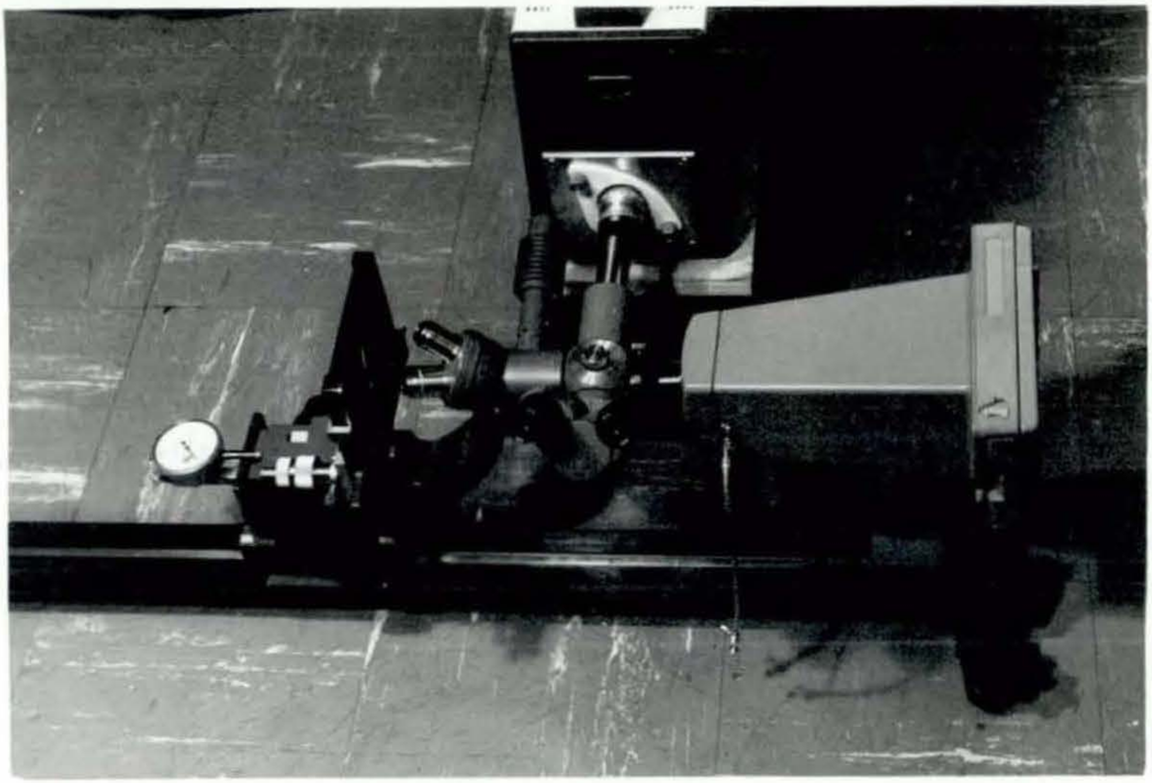


Figure 15. Microscope experimental arrangements.

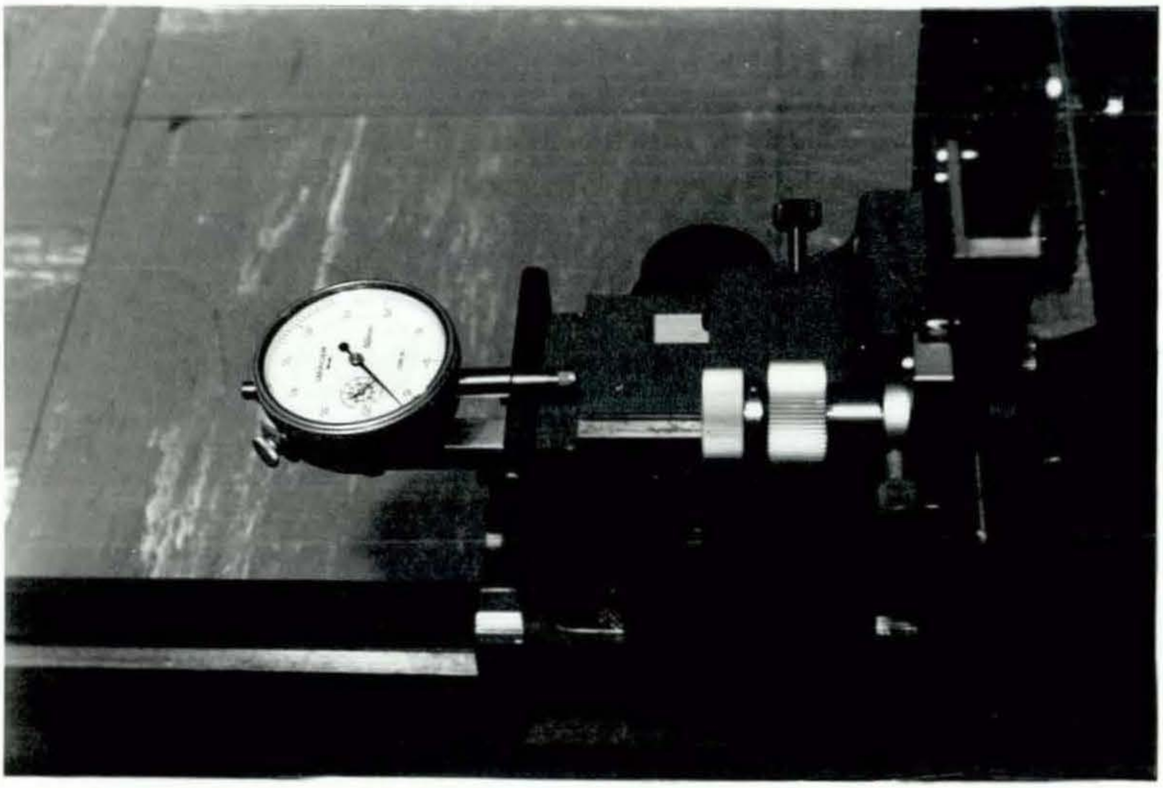


Figure 17. Detail of the dial gauge.

of the object and a positive movement increases the distance between the object and the objective lens.

The T.V. camera facility allows projection of the image onto a T.V. monitor which gives a small (2 to 3 times) magnification but more important allows the superposition of reference diagrams upon the image to aid the observation and measurement of the patterns. An extension of this facility is the use of a Video Tape Recorder (VTR) which also allows the recording of an audio commentary. This permits direct comparison of patterns using two monitors without the use of expensive photographic records.

The measurement of intensity within the microscopic image plane is usually accomplished using a microphotometer. This consists of a single optical fibre which scans the image plane within a special eyepiece and the fibre output is fed to a photomultiplier. The resultant measurement is related to the fibres XY position in the form of a chart recording. These instruments are expensive (>£1000) and clearly outside the budget of this project.

However, a T.V. picture is already a continuously scanned intensity measurement and by sampling the video waveform at the relevant time, the intensity at any point within the T.V. picture may be obtained. Further, the construction of the T.V. video waveform as a series of X direction scans is already in the most usual form of scanned intensity measurements. In the television engineers language, a suitable intensity measuring scheme will extract and display a single selected line from the video waveform and provide a marker on the monitor picture to show which line has been selected. A further refinement which will allow point intensity measurements, is a second marker which gives an X position along the line on both the line display and the monitor.

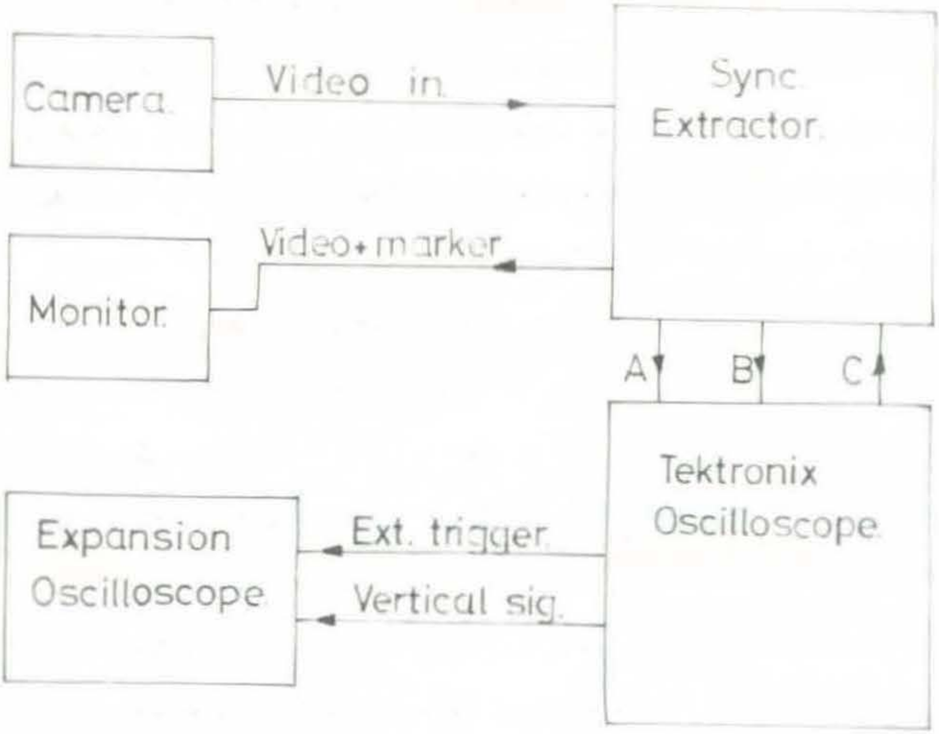


Figure 18a. The block diagram of the line extractor and marker circuit.

A - Line Sync to A Trigger input

B - Frame sync to B Trigger input.

C - A gate out to Delay trigger input.

Figure 18b. Monitor picture of 80μ graticule in 2μ steps.

Figure 18d. The selected line. (the white line in Figure 18b)

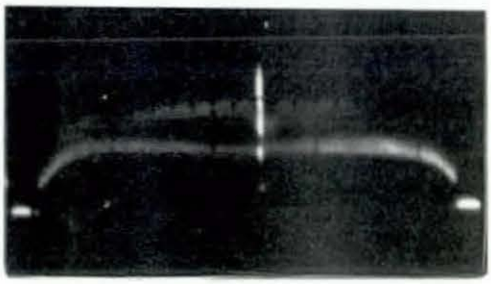


Figure 18c. The frame waveform showing the selected line.

A simple system to achieve these facilities was constructed and is shown in the block diagram of Figure 18a. The major task of extracting the desired line from the video waveform is virtually a standard facility of the Tektronix Oscilloscope and is described in the relevant handbook*. A Sync Extractor and Line Marker circuit was designed and built and is described in detail in Appendix C. This circuit provides stable frame and line pulses for the triggering of the Tektronix delay circuits. The Line Marker brightens the selected line on the monitor (Figure 18b) with a variable width pulse which commences on receipt of the line Sync pulse. The point marker is the end of this pulse and is positioned along the selected line dependent upon the pulse width. A portion of the selected line waveform (complete line waveform shown in Figure 18c) is expanded and displayed on a second oscilloscope using a variable delay facility. The resultant calibration waveform of the monitor picture of Figure 18b is shown in Figure 18d, with the 20 micron markers adjusted to coincide with the 1 cm oscilloscope graticule markers.

The complexity of the intensity amplification system with the numerous automatic gain control facilities, suggests that comparisons of intensity may be made only between points in the same image and no attempt is made to relate these measurements to measurements obtained at different times. This does not apply when the intensity measurements are experimentally normalised to obtain positional information about the image under variable illuminating conditions, which obviously must occur at different times. The normalised procedure is discussed further in Chapter 4 in connection with the measurement of the ramp refractive index profile of core cladded fibres.

*TEKTRONIX OSCILLOSCOPE TYPE 545A

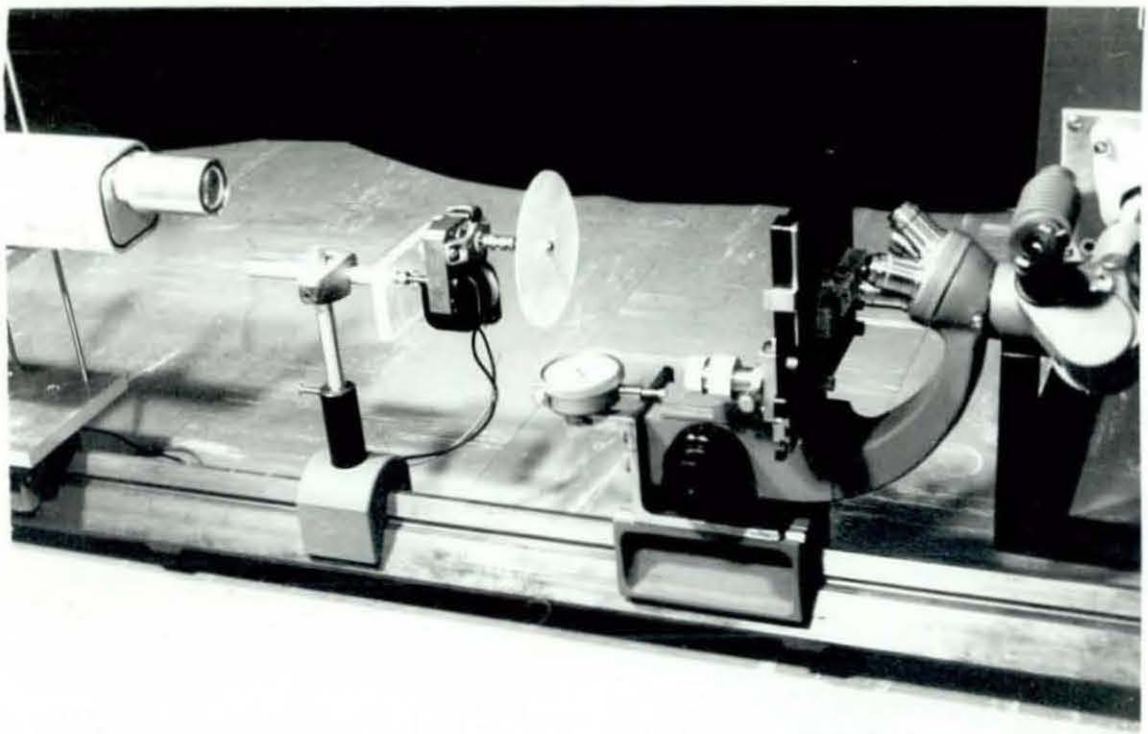


Figure 19. The experimental arrangement for varying the spatial coherence of laser light.

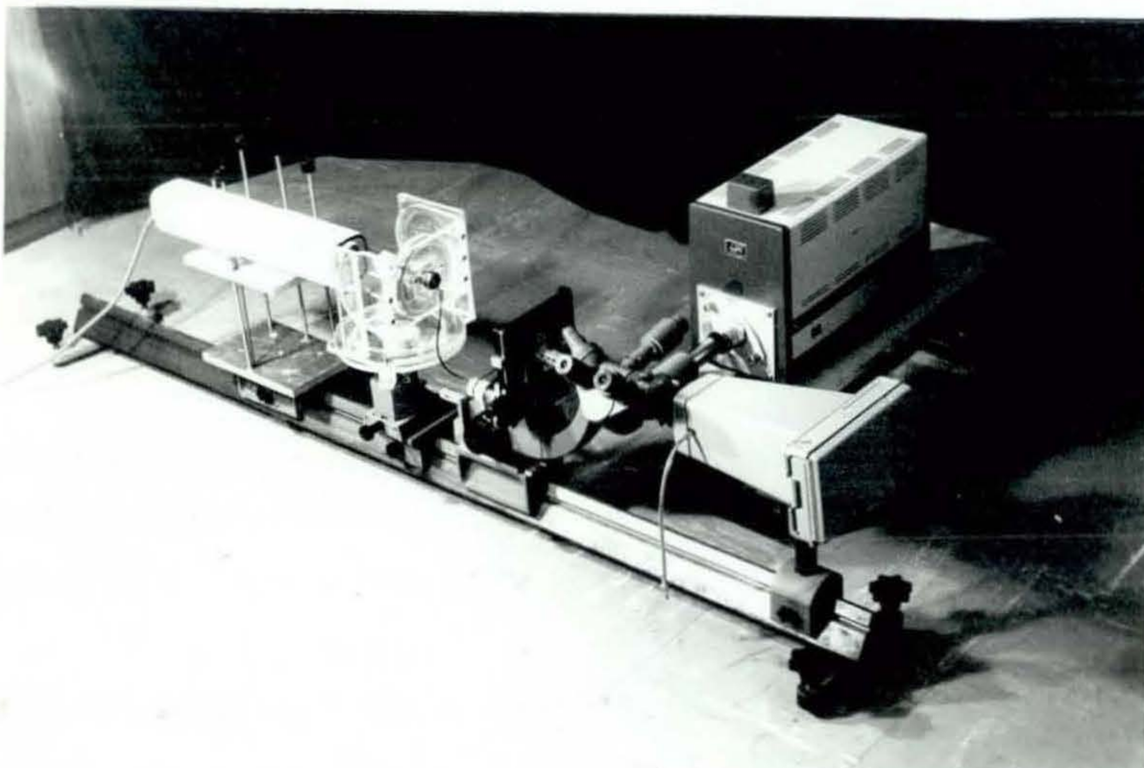


Figure 20. The rotational mount experimental arrangement.

2.7 Light Sources and Illuminating Arrangements

This section is concerned with the experimental details of the illuminating systems used in this study and their analysis is deferred until the discussion of the experimental results in later chapters.

Three sources are used to cover the required range of coherence and bandwidth characteristics. A tungsten filament bulb provides a polychromatic light source and a sodium lamp provides quasi-monochromatic light, both of these sources having spatially incoherent characteristics. A Metrologic Model 920 Helium-Neon laser provides a source of monochromatic spatially coherent light at a wavelength of .6328 micron.

The source areas of the tungsten filament and sodium lamps are adjusted using diaphragms, although the entrance apertures of the fibres are generally the significant diaphragms in the illuminating system.

The spatial coherence of the laser light is varied by inserting a rotating ground glass screen in the laser beam. The degree of spatial coherence of the beam at the entrance aperture of the fibres is a function of the divergence of the beam, the characteristic of the screen and the distances between the various components. This experimental arrangement is shown in Figure 19.

The main experimental difficulty encountered in this study is in arranging the fibres and the illuminating system in a manner which allows the angles of incidence of the light, as defined in Chapter 1, to be varied independently whilst simultaneously observing the effects of such variations on the light leaving the opposite end of the fibres. When the length of the fibres permit the use of the rotational mount of Figure 4, the illuminating system, the rotational mount and the microscope system are mounted on a common axis as shown in Figure 20. In this case the desired degrees of freedom are obtained by virtue of the flexibility of the fibre or fibre bundle.

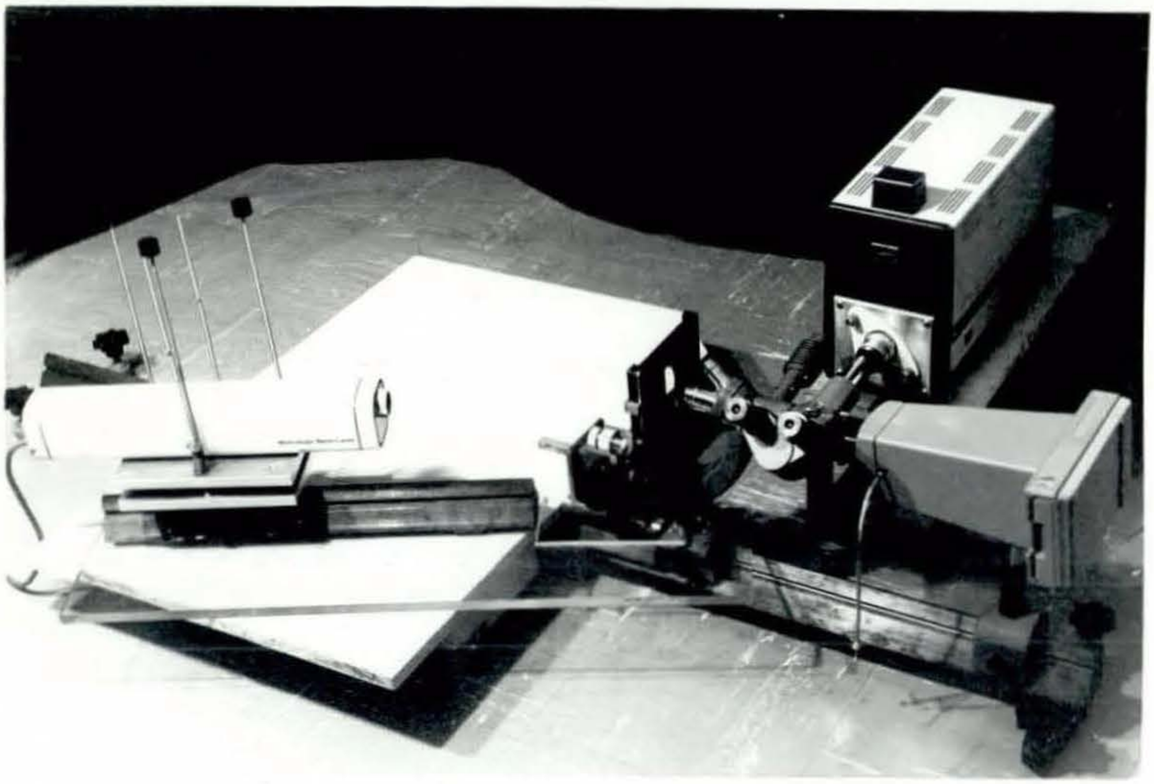


Figure 21. The swivelling optical bench experimental arrangement.

The short fibre specimens have either limited flexibility or no flexibility and their positioning is dictated by the requirements of the microscope system. With the fibres fixed in position the illuminating system must now be moved within a solid angle centred upon the entrance apertures of the fibres. It is not easy to provide the bulky illuminating systems with both the axial and azimuthal degrees of freedom. A simpler solution is to arrange for the fibres to retain a rotational degree of freedom, thus giving the azimuth angle adjustment. The axial degree of freedom is provided by mounting the illuminating system on the swivelling optical bench as shown in Figure 21.

The fibre rotational adjustment refers to the rotation of the entrance end of fibres, with limited flexibility, their other ends being fixed, whereas the completely embedded fibres will rotate at both ends simultaneously. The rotation of the microscopically viewed end may result in the displacement of the fibre and at high magnification the fibre of interest may pass out of the field of view. The azimuthal adjustment is therefore made slowly at low magnification to allow the position of the fibre to be followed, and this makes the adjustment cumbersome. There is, however, no easy alternative. The previous comments about the offset correction for the rotational mount of Figure 4 apply equally to this system and the method for determining the correction is now discussed.

2.8 Determination of the Fibre Orientation and Angle of Incidence

The grinding and polishing machine described in Section 2.3 produces a surface which is normal to the optical bench axis. To utilize this property for fibre terminations requires a method of aligning the embedded fibres at the desired angle to this surface prior to grinding. A procedure which positions the fibres with their longitudinal axis parallel to the optical bench axis is sufficient since the rotational mount will permit the insertion of a measured off axis angle, resulting in a sloping end termination.

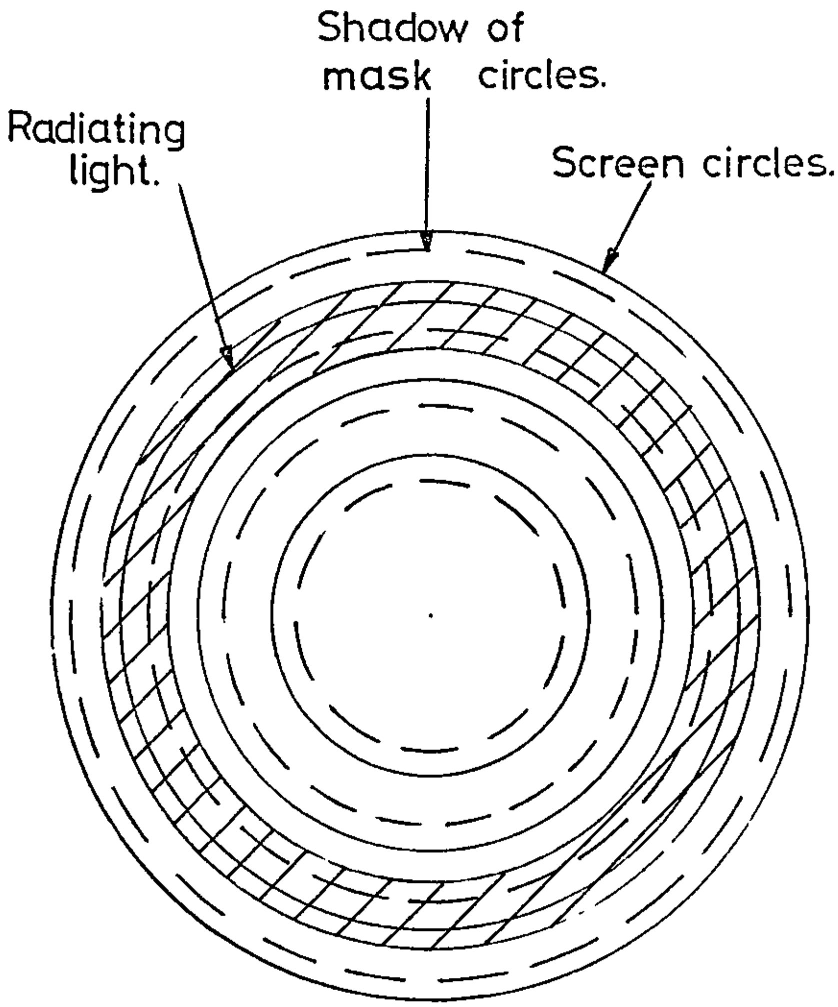


Figure 23. The picture observed on the screen of Figure 22 when the fibre bundle is correctly positioned.

The procedure developed for this purpose uses a well known property[18] of core cladded fibres to radiate a hollow cone of light when illuminated at a single angle of incidence. This property is derived in Chapter 5. If the radiated cone of light is projected onto a screen it produces a distribution of light centred on the extended longitudinal axis of the fibre. When the screen is normal to this axis the distribution forms an annulus whose mean diameter depends upon the angle of incidence and whose width depends upon the divergence of the radiating beam. The required procedure is thus reduced to erecting a screen normal to the optical bench axis and positioning the fibre such that its radiated cone of light produces an annular distribution of light on the screen.

The experimental arrangement is shown in Figure 22 where the single fibre may be replaced by a fibre bundle if the individual fibre axes are parallel. Both ends of the bundle are ground nominally flat and normal. The end to be reground at the measured angle is placed in the rotational mount. The opaque screen S and the transparent mask M are both marked with a series of concentric circles at radial increments of 1 cm. S and M are placed with their centres coaxial with the mounted fibre ends and normal to the optical bench axis. The other end of the fibres are illuminated by the He-Ne laser at incident angles θ, α. The radiating cone of light passes through the mask M and illuminates the screen S.

If the fibre ends are coaxial with the screen and mask, the shadows of the mask rings will be concentric with the screen rings. The fibre is rotated in the θx, θy, and α directions until the annulus of light is also concentric with the screen rings, the concentricity of the shadow mask rings being maintained by using the x, y adjustments of the rotational mount. The correct alignment situation is shown in Figure 23.

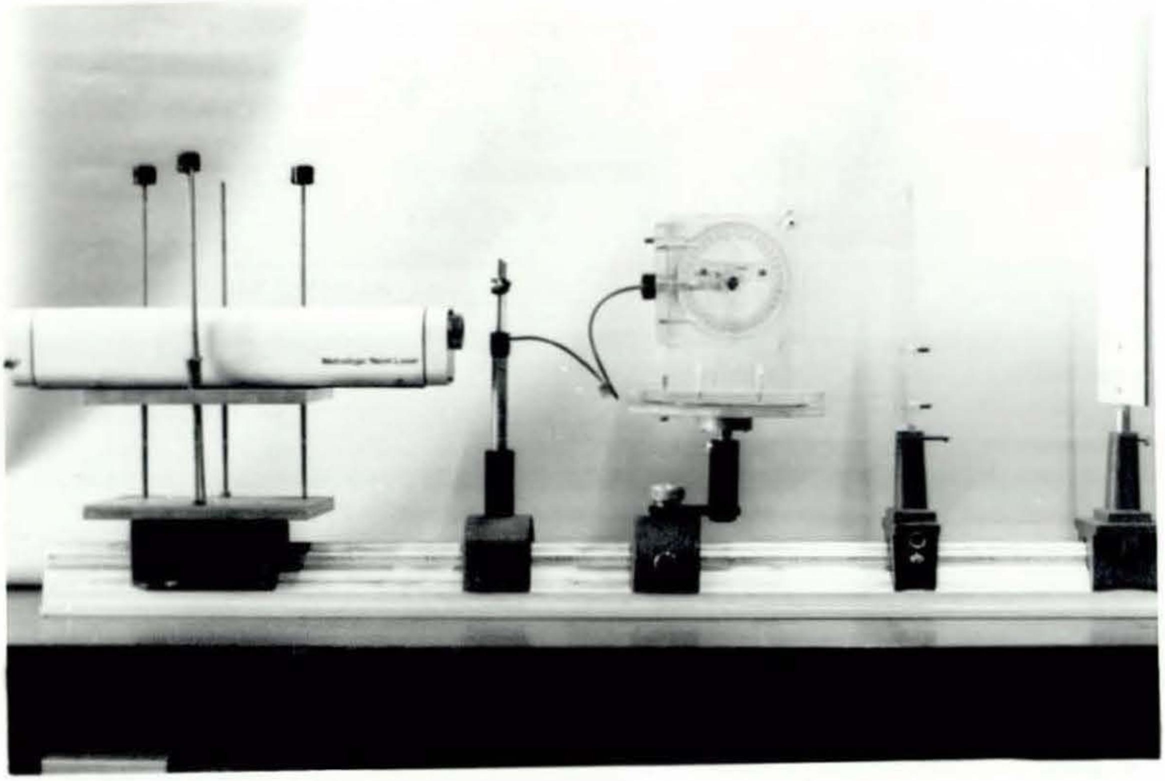


Figure 22. The experimental arrangement for the alignment of fibres.

Figure 25.

$\theta_{r/2} = 60^{\circ}$

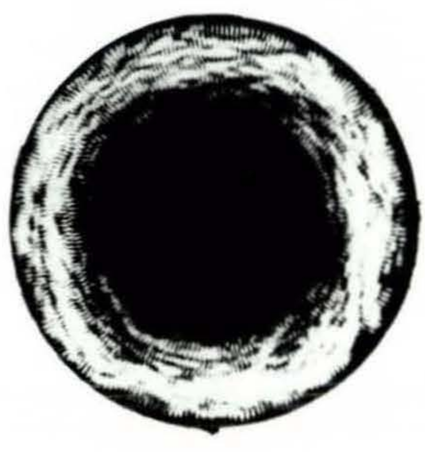


Figure 24.

No account has been taken of the effects of any existing slopes on the radiating ends of the fibre. These are discussed in Chapter 5 where it is shown that the repeated application of this alignment procedure and its associated grinding process will produce a surface which is progressively nearer to the required normal surface.

An alternative procedure is to relate the actual angle of incidence of the light on the fibre to the rotational mount readings thus taking account of any sloping end effects. The axial angle of incidence is measured using an experimental interpretation of the "Black Band" effect different from that first reported by Potter[18]. This effect is now more generally known to be the propagation of skew rays with incident angles above the meridional cut off value, or in waveguide nomenclature as leaky modes[30]. These effects are discussed in more detail in Chapter 5. It is sufficient here to observe that if light is launched at an axial incident angle $\theta_{r/2}$ given by

$$\theta_{r/2} = \sin^{-1}\left(\frac{2}{\sqrt{3}} \sin \theta_c\right) \qquad 1.$$

where θ_c is the meridional critical angle, then in short lengths of fibre (<2m) a black hole of radius r/2 appears in the cross sectional intensity distribution of the fibre, (r is the radius of the fibre core). An example of this condition is shown in Figure 24 where the fibre is illuminated by He Ne laser at $\theta_{r/2} = 40^{o}$ using the experimental arrangement of Figure 20. The dependence of the radius of the black hole on the axial angle of incidence given by equation 189 suggests that this angle may be measured by the appearance of the r/2 radius hole to within $\pm 2^{o}$.

2.9 Conclusion

The techniques described in this chapter for the preparation of single-fibre or multiple-fibre terminations are capable of producing high quality optical surfaces on the ends of the fibres at any desired angle to their longitudinal axes. Since the phenomena of major interest in this thesis are not significantly dependent upon the quality of the end terminations, the techniques developed have proved more than satisfactory for use in this study.

The limited budget available for the purchase of equipment has encouraged the use of novel experimental arrangements, and the apparent simplicity of the phenomena under investigation is reflected in the simplicity of these arrangements. However it is well known that simple optical experiments require complicated analytical techniques for their rigorous analysis. This is apparent from the contents of the next chapter which contains the basic optical theory used in the subsequent analysis of the experiments described in later chapters.

CHAPTER 3

3.1 Introduction

This chapter is intended to provide the basic theory of optics and wave phenomena which is pertinent to this study. The treatment of many of the topics is necessarily brief in order to cover the field in this single chapter. This is considered to be justified in view of the many available textbooks covering similar subject matter[(31-34)].

All electromagnetic waves satisfy Maxwell's Equations and this provides a convenient starting point for the chapter. The characteristics of optical waves are presented in the sections entitled Bandwidth and Coherence, Polarisation, and Intensity. The effects produced when waves are obstructed by apertures or obstacles are analysed using Diffraction Theory. The interaction of reflected or diffracted waves with themselves is considered in the section on Interference Phenomena. The study of the behaviour of waves at dielectric interfaces leads to the derivation of the important Fresnel Equations. The derivation of a scalar wave equation from Maxwell's equation is the foundation of the geometrical ray theory of light propagation.

The selection of certain experimental arrangements to illustrate some of the theory in this chapter is based upon their close analogy with experimental arrangements which arise in the analysis of light propagating in the optical fibres.

3.2.1 Maxwell's Equations and the Wave Equation

The form of Maxwell's equations used in this chapter is that used in reference 31 which is specifically concerned with light waves.

The derivation of these equations from the fundamental observations of moving charges is contained in the extensive literature on Electromagnetism[(35-38)] and only a statement of the equations is given here.

The electric vectors are denoted by E, the field intensity, and D, the displacement. The corresponding magnetic vectors are H, the field intensity and B the flux density. The following four equations are Maxwell's equations for linear isotropic media in the absence of currents or charges.

$$\mathrm{curl}\ H - \frac{\partial D}{\partial t} = 0 \tag{2.}$$

$$\mathrm{curl}\ E + \frac{\partial B}{\partial t} = 0 \tag{3.}$$

$$\mathrm{div}\ D = 0 \tag{4.}$$

$$\mathrm{div}\ B = 0 \tag{5.}$$

The operators curl and div are standard vector operators and $\frac{\partial}{\partial t}$ denotes differentiation with respect to time. Two auxiliary equations, known as the material equations, relate to the two electric vectors and the two magnetic vectors in terms of the material constants, ε the dielectric permittivity and μ the magnetic permeability.

$$D = \varepsilon E \tag{6.}$$

$$B = \mu H \tag{7.}$$

The energy propagated by an electromagnetic field is measured in terms of the Poynting vector S which is the power flow density vector.

$$S = E \times H \tag{8.}$$

The wave behaviour of electromagnetic fields becomes apparent by modifying equations (2) and (3) using equations (4) to (7) to form two independent wave equations in terms of E and H, signifying that each of these field components propagates as a wave.

The wave equations are :-

$$\nabla^2 E = \varepsilon\mu \frac{\partial^2 E}{\partial t^2} \qquad 9a.$$

$$\nabla^2 H = \varepsilon\mu \frac{\partial^2 H}{\partial t^2} \qquad 9b.$$

These are standard wave equations where the velocity, v, of propagation is given by

$$v = (\varepsilon\mu)^{-\frac{1}{2}} \qquad 10.$$

The solutions to this wave equation are many and just two are considered here, those for a plane wave and a spherical wave. It is also convenient to introduce here the notation used to represent non-monochromatic waves.

A solution to the wave equation

$$\nabla^2 V = \frac{1}{v^2} \frac{\partial^2 V}{\partial t^2} \qquad 11.$$

will be of the form $V(x,y,z,t)$ where V may be either the E or H component. If $r(x,y,z)$ is a position vector then V may be represented by its time dependent component $F(t)$ and space dependent component $U(r)$

$$V(x,y,z,t) = F(t)\ U(r) \qquad 12.$$

The space dependent component of the wave will also satisfy the time independent wave equation

$$(\nabla^2 + k^2)U = 0 \qquad 13.$$

where k is called the wave number and is given by

$$k = \frac{\omega}{v} \qquad 14.$$

The angular frequency, ω, is related to the frequency, ν, of the time dependent function

$$\nu = \frac{\omega}{2\pi} \qquad 15.$$

Using the exponential function the general space dependent component may be expressed as

$$U(r) \quad = \quad R\{a(r)\ e^{-ig(r)}\} \tag{16.}$$

The amplitude a(r) and g(r) are both real functions of position and R denotes the real part of a function.

A monochromatic harmonic wave has a time dependence of the form

$$F(t) \quad = \quad \cos \omega t \tag{17.}$$

which in exponential form becomes

$$F(t) \quad = \quad R\{e^{-i\omega t}\} \tag{18.}$$

A non-monochromatic wave may be represented as the superposition of monochromatic waves (using the Fourier Theorem)

$$F(t) \quad = \quad R\int_0^\infty a(\omega)\ e^{-i(\omega t-g(\omega))} \quad \partial\omega \tag{19.}$$

where $a(\omega)$ are the amplitudes of the Fourier components and $g(\omega)$ their phase function.

The amplitudes and phase functions of the time and space components may be combined to give the general form of the wave.

$$V(rt) \quad = \quad R\int_0^\infty a(\omega,r)e^{-i(\omega t-g(\omega,r))} \quad \partial\omega \tag{20.}$$

3.2.2 <u>Wave Solutions for the Space Dependent Component</u>

A plane wave has a non varying amplitude in planes perpendicular to its direction of propagation. If $\overline{r}$ represents the direction of propagation then a(r) and g(r) of equation (16) may be shown to be

$$g(r) \quad = \quad (\overline{k}.\overline{r} - \delta) \qquad a(r) = a \tag{21.}$$

where $\overline{k}$ is the wave number in the direction $\overline{r}$, and δ is a phase constant.

A spherical wave radiates isotropically from a point source and thus has a propagation constant dependent upon $|r|$ where $|r| = \sqrt{x^2 + y^2 + z^2}$ and may be shown to have an amplitude dependence of $\frac{1}{|r|}$

$$\text{i.e.} \qquad g(r) = (k.|r| - \delta) \qquad a(r) = \frac{a}{|r|} \qquad 22.$$

It is clear that a small section of a spherical wave, i.e. x,y << z may be represented as a plane wave if the variation of amplitude is of no significance.

3.3 Bandwidth and Coherence of Optical Waves

The spectral range of an optical wave is represented by the amplitudes $a(\omega)$ of its Fourier components. The term monochromatic normally implies an idealised source whose frequency spectrum consists of a single component. A sodium lamp emits light with a narrow spectral width which, for the purpose of this thesis, is termed quasi-monochromatic.

The laser radiation, which has a much narrower (but finite) spectral width than the sodium light, is termed monochromatic in this thesis purely to distinguish it from sodium light. The white light source emits radiation with significant amplitudes over a large spectral range compared with the sodium source spectral range, and is termed polychromatic.

The spatial dependence of the amplitudes of the Fourier components within the source aperture for each of the three sources is negligible, but their spatial phase relationship is extremely significant. A monochromatic wave implies that the wave has existed for all time and there will therefore be no changes of phase of the wave with time. If this is true over the whole of the source aperture, then the phase at one point in the aperture will have a constant phase relationship with the wave at any other point. Such a relationship is called spatial coherence. The converse condition called spatial incoherence implies a random phase relationship between waves at different points within the source aperture. The absence of both of these conditions is called partial spatial coherence.

If the monochromatic condition is relaxed then the continuous wave is truncated in time to form a wavelet, whose temporal length is inversely proportional to the width of the frequency spectrum of the source. (This leads to the interesting definition of white light as an impulse in the time domain). Individual wavelets will arrive at a point of observation at random times and because any optical detection process involves a long integration time relative to the length of the wavelets, they will appear to have random phase relationships with each other. This condition is termed temporal incoherence. The corollary must be that temporal coherence is observed when the observation process has a shorter duration than the time taken for a single wavelet to pass the point of observation.

These coherence properties are measured and conversely are significant in interference experiments where they affect the visibility of the interference fringes, and for this reason their analysis is deferred to the section on Interference.

3.4 Polarisation

The definition of the Poynting vector, Equation 8, specifies that the E, H, vectors and the direction of propagation of every y form a right handed triad of vectors. Thus for a plane wave propagating in the z direction the E and H vectors are orthogonal and lie in the xy plane.

The E components of this plane wave may be denoted as

$$E_x = a_1 \cos(\omega t - kz + \delta_1)$$

$$E_y = a_2 \cos(\omega t - kz + \delta_2)$$

$$E_z = 0$$

The locus of the end of the E vector

$$E = \sqrt{E_x^2 + E_y^2} \qquad 23.$$

represents the path taken by the E vector as the field propagates. The general case will give an elliptic locus which may degenerate into a straight line or a circle. These represent elliptical, linear and circular polarisation respectively and the elliptic and circular loci may be left or right handed. The phase difference $(\delta_2 - \delta_1)$ and the relative amplitudes (a_1, a_2) of the components determine the state of polarisation observed at a point in the field.

Linear polarisation $(\delta_2 - \delta_1) = n\pi (n = 0, \pm 1 \pm 2)$ a_1, a_2 arbitrary

Circular polarisation $(\delta_2 - \delta_1) = \frac{n}{2}\pi (n = \pm 1, \pm 3, \pm 5)$ $a_1 = a_2$

all other conditions give rise to elliptic polarisation.

The significance of the polarisation property of electromagnetic waves appears when considering the behaviour of waves at refracting or reflecting interfaces. Any state of polarisation may be considered in terms of two orthogonal components which may be arbitrarily selected. Normally one component is selected in the plane of incidence of the wave at the interface, making the second component normal to this plane.

The three sources used in this study emit light with no preferred state of polarisation, and which, therefore, exhibits any of the possible states in random manner. Light with this characteristic is known as natural light and may be represented as a superposition of two beams linearly polarised at right angles to each other. The total intensity of the beam is equally divided between the two components and their field components are orthogonal.

3.5 Intensity of Electromagnetic Waves

The intensity of light waves is defined[31] as the time average of the Poynting Vector.

$$I = \overline{s} = \overline{R\{\tfrac{1}{2}(E \times H^*)\}} \qquad 24.$$

The overbar indicates the time average over the period T, and * indicates a complex conjugate operation. A linearly polarised plane wave will have mutually orthogonal E and H components which may be denoted by E_x, H_y and which are related through Maxwell's Equations.

$$|H_y| = \sqrt{\frac{\varepsilon}{\mu}}\,|E_x| \qquad 25.$$

Substitution of equation (25) into (24)

$$I = R\left\{v\sqrt{\frac{\varepsilon}{\mu}} \quad \overline{E}_x^2\right\} \qquad 26a.$$

$$= R\left\{v\sqrt{\frac{\mu}{\varepsilon}} \quad \overline{H}_y^2\right\} \qquad 26b.$$

Since absolute values are not required, the intensity may be given by

$$I = \overline{E}^2 \qquad 27.$$

A monochromatic plane wave may be represented by an E field of the form

$$E(r,t) = R(u(r)\ e^{-i\omega t}) = \tfrac{1}{2}\left(u(r)\ e^{-i\omega t} + u(r)^*\ e^{i\omega t}\right) \qquad 28.$$

The complex spatial component u(r) may be expressed in terms of its rectangular cartesian components

$$u_x = a_1(r)\ e^{ig_1(r)} \qquad u_y = a_2(r)\ e^{ig_2(r)} \qquad 29.$$

where a_j are the constant amplitude terms and g_j are phase factors of the form of Equation 21.

From (28)

$$E^2 = \tfrac{1}{4}\left(u^2\, e^{-2i\omega t} + u^{*2}\, e^{2i\omega t} + 2\, u\, u^*\right) \tag{30.}$$

If the time averaging interval, T, is large compared with the interval $t = \frac{2\pi}{\omega}$ of the first two terms of Equation (30) contribute zero and

$$I = \overline{E^2} = \tfrac{1}{2}\, u \,.\, u^* = \tfrac{1}{2}\left(|u_x|^2 + |u_y|^2\right)$$

$$= \tfrac{1}{2}(a_1^2 + a_2^2) \tag{31.}$$

3.6.1 Interference

When two monochromatic plane waves E_1, E_2 at the same frequency are superposed, the total electric field at a point P will be given by

$$E = E_1 + E_2 \tag{32.}$$

The intensity at this point

$$I = \overline{E}^2 = \overline{E}_1^2 + \overline{E}_2^2 + \overline{2E_1E_2}$$

$$= I_1 + I_2 + J_{12} \tag{33.}$$

where the intensities of the individual waves, I_1, I_2, are given by Equation (31) and the term J_{12} is called the interference term, or the mutual intensity.

Let $u_1(r)$ and $u_2(r)$ be the complex amplitudes of the two waves E_1 and E_2 where

$$\begin{aligned} u_{1x} &= a_1(r)\, e^{ig_1(r)} \qquad & u_{2x} &= b_1(r)\, e^{ih_1(r)} \\ u_{1y} &= a_2(r)\, e^{ig_2(r)} \qquad & u_{2y} &= b_2(r)\, e^{ih_2(r)} \end{aligned} \tag{34.}$$

It will be assumed that the phase difference between the two waves is the same for each rectangular component and is denoted by δ.

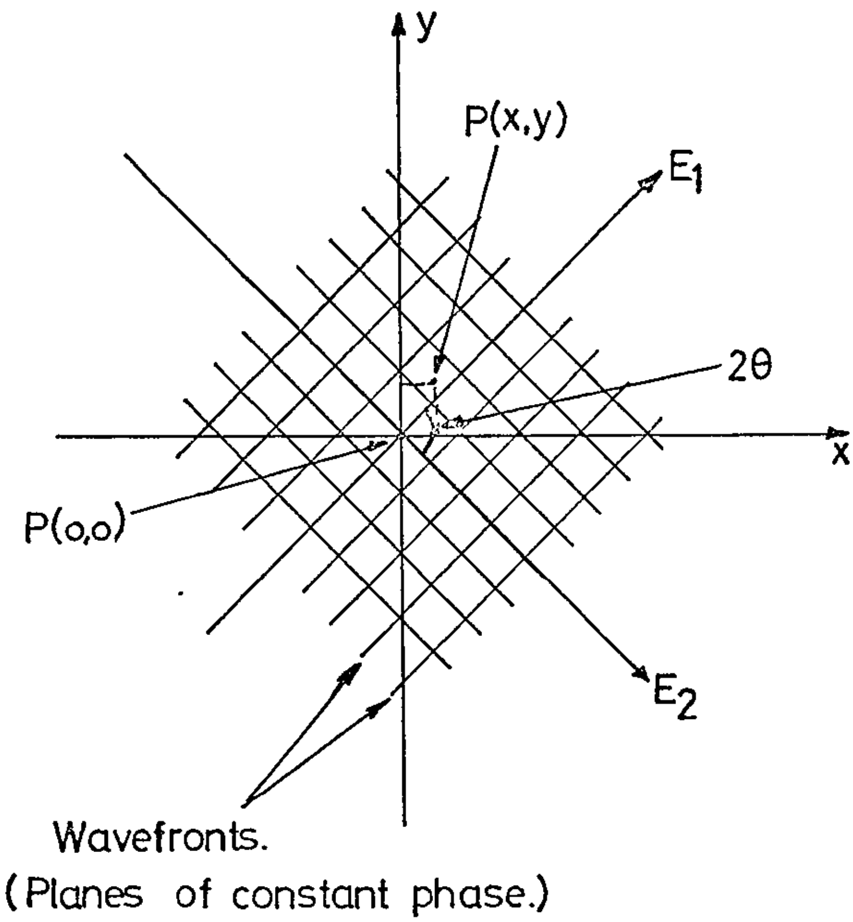


Figure 26. Two plane waves intersecting at angle 2θ.

It may be shown that

$$J_{12} = (a_1 b_1 - a_2 b_2) \cos \delta \qquad 35.$$

For light polarised linearly in the x direction $a_2 = b_2 = 0$, and using Equation (31)

$$J_{12} = a_1 b_1 \cos \delta = 2\sqrt{I_1 I_2} \cos \delta \qquad 36.$$

The total intensity at P is given by Equation (33) as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \delta \qquad 37.$$

The dependence of the total intensity upon the phase difference δ causes the spatial fringes which are the characterising feature of interference phenomena.

The discussion of natural light suggests that it may be represented by components of the form of Equation (34) where $a_1 = a_2$, $b_1 = b_2$ and under the assumption that the phase difference is the same for both linearly polarised components, Equation (37) will apply for fringes formed with natural light.

When the two beams are of equal intensity, i.e. $a_1 = b_1$, then the total intensity is given by

$$I = 4I_1 \cos^2 \frac{\delta}{2} \qquad 38.$$

Consider the case when two plane linearly polarised monochromatic waves are inclined to each other at an angle of 2θ, such that the planes of constant phase intersect as shown in Figure 26. The origin of the coordinate system is positioned at a point of equal phase of the two waves. The phase difference between the two waves at a point P(x,y) is given by

$$\delta = k\, 2y \sin \theta \qquad 39.$$

and the intensity at P, if the two waves are of equal intensity I

$$I(P) = 4\ I\ \cos^2\ (ky\ \sin\ \theta) \qquad 40.$$

The maxima of intensity will occur at

$$y = \frac{n\pi}{k\ \sin\theta} \qquad n = 0,\ 1,\ 2,\ \ldots\ldots \qquad 41.$$

and minima (zero intensity) when $n = \frac{1}{2}, \frac{3}{2}$ ..... in Equation 41.

Substituting $k = \frac{2\pi}{\lambda}$ where λ is the wavelength of the waves in the transmission media, the fringe spacing is given by

$$t = \frac{\lambda}{2\ \sin\ \theta} \qquad 42.$$

The fringe at P(0,0) is defined as the zero order fringe, which implies that it will be a point of maximum intensity irrespective of the wavelength of the waves. The m^{th} order fringe is a distance $y = mt$ from the zero order fringe.

If the waves are quasi-monochromatic, that is they have a wavelength range $\Delta\lambda$ which is small compared with the mean wavelength λ', then the m^{th} order fringe will be displaced in the plane of observation (the y axis) by an amount $\Delta y = \frac{m\Delta\lambda}{2\sin\theta}$ 43.

If

$$m << \frac{\lambda'}{\Delta\lambda} \qquad 44.$$

then Δy is negligible compared with mt and the different wavelength fringes will be coincident and the contrast of the fringes will be the same as for the monochromatic waves.

If the monochromatic wave E_1 has its optical path length increased by ΔL prior to reaching the point P(0,0) then the zero order fringe will move in the y direction

$$yo = \frac{\Delta L}{2\sin\theta} \qquad 45.$$

and the intensity at P(0,0) will be maximum or minimum depending upon the value of n as in Equation (41) in the expression

$$n = \frac{\Delta L}{\lambda} \qquad 46.$$

The contrast of fringes at a point P is measured using a quantity called visibility which is defined as

$$\Xi|P| = \frac{I_{MAX} - I_{MIN}}{I_{MAX} + I_{MIN}} \qquad 47.$$

The fringe intensity given by Equation (39) is due to the superposition of monochromatic waves which from the discussion in Section 3.3 are known to be coherent. Substituting the relevant values of (38) into (47) yields a value of Ξ of 1 corresponding to coherent fringe visibility. The absence of any interference fringes ($I_{MAX} = I_{MIN}$) implies incoherence and the corresponding value of Ξ is zero. A value of Ξ between 0 and 1 signifies a partial coherence condition.

Consider the fringe visibility at point P(0,0) when the illumination is quasi-monochromatic and ΔL is increased until the order m of the fringe at P(0,0) is such that (44) becomes

$$m \simeq \frac{\lambda'}{\Delta\lambda} \qquad 48.$$

The different frequency fringes will now be significantly displaced from the centre frequency fringe position and Ξ(0,0) will be less than 1 and the illumination at this point is now partially coherent.

An alternative explanation for this reduction in coherence is obtained by considering the wavelet representation of quasi-monochromatic light. If it is assumed that the two interfering waves are derived from the same source, then a part of each wavelet (in terms of amplitude) will propagate in each of the waves. If ΔL exceeds the length of the wavelet, then the intensity at the point of observation will consist of the superposition of different wavelets which have random phase relationships and will therefore produce no interference effects. If ΔL is less than the wavelet length, a part of each wavelet (in terms of length) will overlap and produce interference fringes.

The wavelet length is given by

$$\ell_c = \frac{\lambda'^2}{\Delta\lambda} \qquad 49.$$

and is known as the coherence length.

Typical values of the coherence length for the sources used in this study are

$\ell_c \simeq 3\mu.$ White light source

$\simeq 3$m m. Sodium Source

> 30m. Laser

The white light and sodium sources may be considered to be composed of an ensemble of small sources each radiating incoherently but emitting light within the overall source bandwidth. By judicious experimental arrangement, as described in the next section, these ensembles of incoherent sources may exhibit coherent properties. The laser however has, as part of its light generation system, a very narrow bandwidth optical filter. The modus operandum of this filter produces a spatially coherent beam of light from what may be an ensemble of incoherent emissions of radiation from the laser gas. The contribution of the stimulated emission process to the coherence properties of the laser is still under investigation[39]. The allied problem of the enhancement of spatial coherence of waves by propagation in bounded media (waveguides, cavities etc.) has also received attention[40,41]. A small contribution to this last subject is made in Chapter 8.

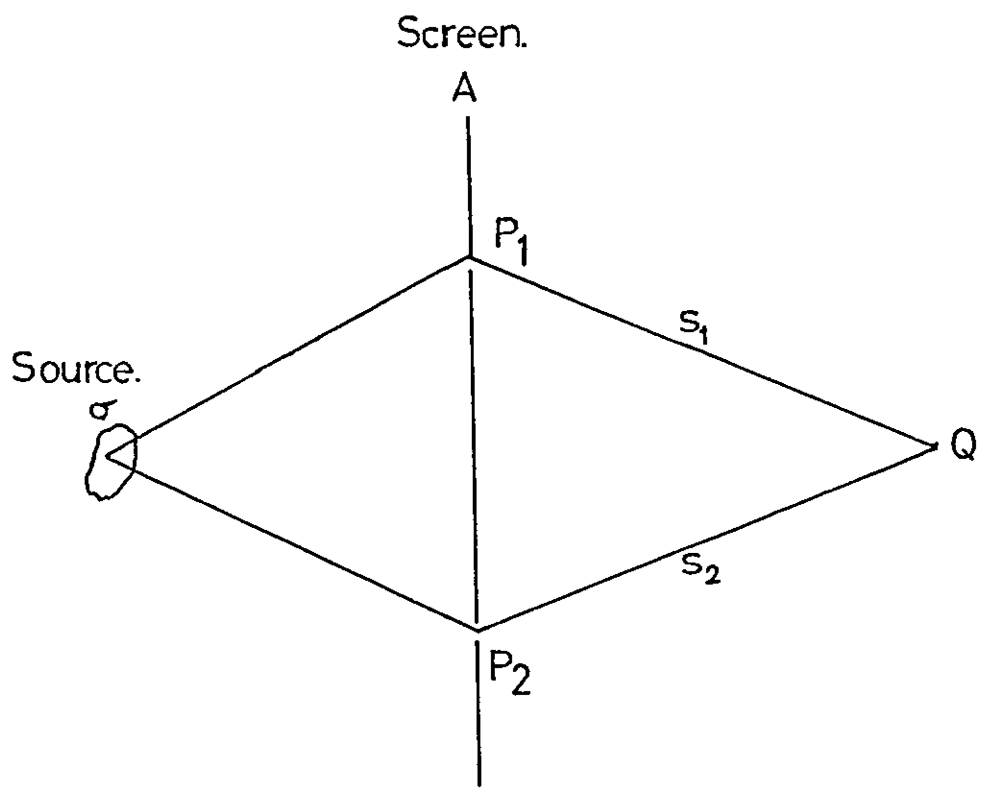


Figure 27. The source σ illuminates the two pinholes P_1,P_2 and the diffracted light contributes to the light observed at Q.

3.6.2 Interference of Non Monochromatic Waves

The electric field at a point P of a general wave may be represented by the complex function V(P,t) where the physical scalar wave is taken as the real part of V(P,t). The intensity of the wave is given by

$$I(P) = R\{\overline{V(P,t)^2}\} = \tfrac{1}{2}(\overline{V(P,t)V^*(P,t)}) \qquad 50.$$

Consider the experimental arrangement of Figure 27, where the source σ illuminates two pinholes P_1, P_2 in the opaque screen A. These pinholes behave as secondary sources and the intensity due to the superposition of these secondary waves is measured at a point Q, a distance s_1, s_2 from P_1, P_2 respectively.

The time taken for the waves to travel from P_1 and P_2 to Q are t_1, t_2 respectively where

$$t_1 = \frac{s_1}{v} \quad t_2 = \frac{s_2}{v} \qquad 51.$$

$$(t_2 - t_1) = \tau \qquad 52.$$

Denoting the waves at P_1 and P_2 by the complex functions, V_1, V_2 the total wave at Q is given by

$$V(Q,t) = K_1 V_1 + K_2 V_2 \qquad 53.$$

where K_1 and K_2 are propagating coefficients which relate the wave values at the points P_1, P_2 to their values at the point Q. These coefficients are discussed in the section on diffraction by circular apertures, Section 3.7.3. The intensity at Q is given by

$$I(Q) = |K_1^2| I_1 + |K_2|^2 I_2 + 2|K_1K_2| R(\Gamma_{12}(\tau)) \qquad 54.$$

where $R(\Gamma_{12}(\tau))$ is the real part of the function

$$\Gamma_{12}(\tau) = \overline{V_1(t+\tau)V_2^*(t)} \qquad 55.$$

The cross correlation function $\Gamma_{12}(\tau)$ is called the mutual coherence function and defines the coherence between waves at points P_1, P_2, the wave at P_1 being considered a time τ later than that from P_2. When $\tau = 0$

$$\Gamma_{12}(0) = J_{12} \qquad 56.$$

and $\Gamma_{12}(\tau)$ is clearly an intensity measurement. It may be normalised to give the complex degree of coherence.

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1 I_2}} \qquad 57.$$

and it may be shown that

$$|\gamma_{12}(\tau)| \leqslant 1 \quad \text{(Ref. 31, Chapter 10)}$$

If the complex degree of coherence is defined as

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)| \, e^{i(\alpha_{12}(\tau) - 2\pi\bar{\nu}\tau)} \qquad 58.$$

then $$R\{\gamma_{12}(\tau)\} = |\gamma_{12}(\tau)| \cos(\alpha_{12}(\tau) - \delta) \qquad 59.$$

where $$\alpha_{12}(\tau) = \arg \gamma_{12}(\tau) \qquad 60.$$

$\bar{\nu}$ is the mean frequency of the light and

$$\delta = 2\pi\bar{\nu}\tau \qquad 61.$$

Setting $$|K_1^2| \, I_1 = I_1(Q) \qquad |K_2^2| \, I_2 = I_2(Q) \qquad 62.$$

Equation 54 may be written

$$I(Q) = I_1(Q) + I_2(Q) + \sqrt{I_1(Q)\, I_2(Q)} \, |\gamma_{12}(\tau)| \cos(\alpha_{12}(\tau) - \delta) \qquad 63.$$

Comparing Equation (63) with (37) it is clear that $|\gamma_{12}(\tau)|$ determines the visibility of the fringes and $\alpha_{12}(\tau)$ will give the

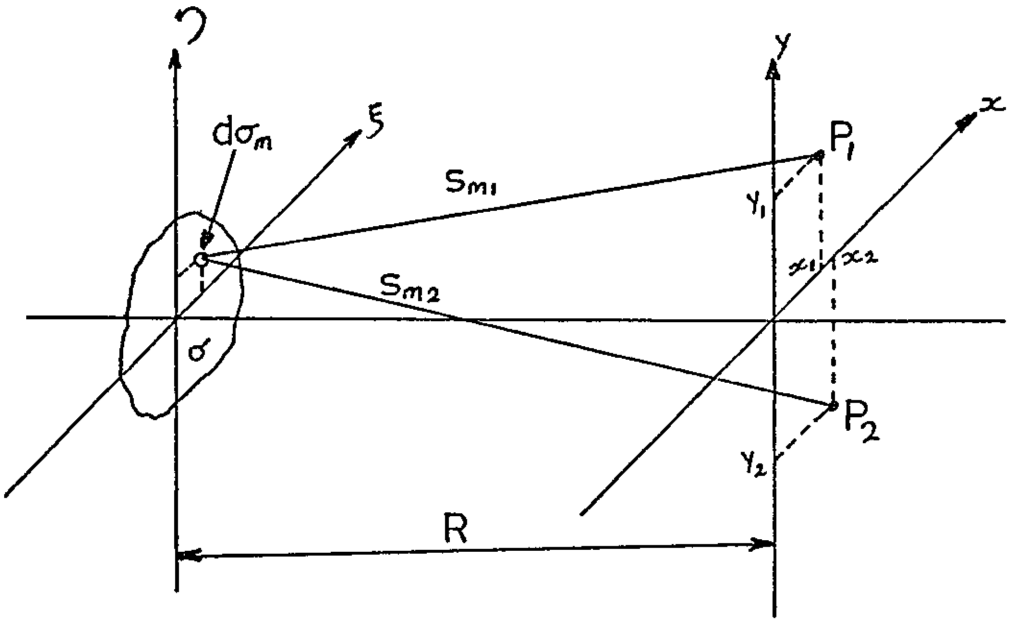


Figure 28. The light from a single element of the source σ contributes light to the two points P_1 and P_2.

displacement of the fringes. If $I_1(Q) = I_2(Q)$ then on substitution of the relevant values of Equation (63) into (47)

$$\Xi(Q) = |\gamma_{12}(\tau)| \qquad 64.$$

When $\nu\tau << \frac{\lambda^2}{\Delta\lambda}$, $\gamma_{12}(\tau)$ will change slowly in comparison with the cos δ term and this corresponds to the quasi-monochromatic condition previously considered. As $\Delta\lambda$ is increased, then the dependence upon τ becomes significant, which may be interpreted as the 'overlapping of fringes condition', but which now appears as a reduction in visibility which is implicit in the value of $|\gamma_{12}(\tau)|$.

To calculate the complex degree of coherence γ_{12} between the points P_1 and P_2 illuminated by a quasi-monochromatic source, shown in Figure 28, each element of the source is assumed to radiate spherical waves which are independent of any other element. Denoting the disturbances at P_1, P_2 due to the radiation from element $d\sigma_m$ as $V_{m1}(t)$, $V_{m2}(t)$ total disturbances at P_1 and P_2 are given by

$$V_1(t) = \sum_m V_{m1}(t) \qquad V_2(t) = \sum_m V_{m2}(t) \qquad 65.$$

Using the assumption that $\tau \simeq 0$ then from Equations (55),(56)

$$J_{12} = \overline{V_1(t)\ V_2^*(t)} = \sum_m \overline{V_{m1}(t)\ V_{m2}^*(t)} + \sum_m \overline{V_{m1}(t) V_{n1}^*(t)} \qquad 66.$$

The independence condition between radiation from different $(m \neq n)$ elements will give a zero value for the second term of (66). The radiation from the m^{th} element may be denoted $A_m\, e^{-2\pi i \bar{\nu} t}$ and, if s_{m1}, s_{m2} are the distances from this element to the points P_1, P_2, then t_{m1}, t_{m2} are given by equation (51). Using the equations for spherical waves (22),

$$V_{m1}(t) = A_m(t-t_{m1})\,\frac{e^{-i2\pi\bar{\nu}(t-t_{m1})}}{s_{m1}}\,,\; V_{m2}(t) = A_m(t-t_{m2})\,\frac{e^{-i2\pi\bar{\nu}(t-t_{m2})}}{s_{m2}} \qquad 67.$$

and $$\overline{V_{m1}(t)V^*_{m2}(t)} = \overline{A_m(t-t_{m1})A^*_m(t-t_{m2})}\,\frac{e^{i2\pi\bar{\nu}(t_{m1}-t_{m2})}}{s_{m1}s_{m2}} \qquad 68.$$

Shifting the time origin by setting $t_{m2}-t_{m1} = \tau \simeq 0$ in the argument of A^*_m

$$J_{12} = \sum_m \overline{A_m(t)\;A^*_m(t)}\,\frac{e^{ik(s_{m1}-s_{m2})}}{s_{m1}s_{m2}} \qquad 69.$$

The term $\overline{A_m(t)A^*_m(t)}$ is the intensity of the element $d\sigma_m$ and so denoting by $I(s)$ the source intensity per unit area, Equation (69) becomes

$$J_{12} = \int_\sigma I(s)\,\frac{e^{ik(s_1-s_2)}}{s_1 s_2}\,ds \qquad 70.$$

where the surface integral is taken over the source surface denoted by σ.

Normalising Equation (70)

$$\gamma_{12} = \frac{1}{\sqrt{I(P_1)I(P_2)}}\int_\sigma I(s)\,\frac{e^{ik(s_1-s_2)}}{s_1 s_2}\,ds \qquad 71.$$

where $I(P_1)$, $I(P_2)$ are the intensities at P_1 and P_2.

If, as shown in Figure 28, the co-ordinates of a typical source element are (ξ,η), the co-ordinates of $P_1 P_2$ are (x_1,y_1), (x_2,y_2) and R is the distance between the source plane and the plane containing P_1, P_2 then, using

$$p = \frac{(x_1 - x_2)}{R}\;,\quad q = \frac{(y_1 - y_2)}{R} \qquad 72.$$

and the approximations discussed in reference (31) pp 510, equation (71) becomes

$$\gamma_{12} = \frac{e^{i\psi} \iint_{\sigma} I(\xi\eta)\ e^{i\bar{k}(p\xi + q\eta)}\ d\xi d\eta}{\iint_{\sigma} I(\xi\eta)\ d\xi d\eta} \qquad 73.$$

where

$$\psi = \frac{\bar{k}\{(x_1^2 + y_1^2) - (x_2^2 + y_2^2)\}}{2R} \qquad 74.$$

and represents the phase difference $\frac{v\tau 2\pi}{\bar{\lambda}}$ which for quasi-monochromatic light may be neglected if $v\tau << \bar{\lambda}$. Equation (73) is the normalised Fourier Transform of the intensity function of the source and for a circular uniform intensity source radius ρ yields a solution in terms of Bessel functions of the first kind and first order $J_1(u)$.

$$\gamma_{12} = \left(\frac{2J_1(u)}{u}\right) e^{i\psi} \qquad 75.$$

where

$$u = \bar{k}\rho\sqrt{p^2+q^2} = \frac{2\pi\rho}{\bar{\lambda}R}\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad 76.$$

When $u = 0$, i.e. P_1 and P_2 are coincident, γ_{12} has the value 1 indicating coherence, and as P_1 and P_2 are separated γ_{12} reduces until at a separation of $\frac{0.61R\bar{\lambda}}{\rho}$ there is complete incoherence, $\gamma_{12} = 0$.

At $u = 1$, $\gamma_{12} = 0.88$ and this is taken as a measure of the diameter d, of the circular area almost coherently illuminated by a circular source of radius p, radiating incoherent quasi-monochromatic light.

$$d = \frac{0.16\bar{\lambda}R}{\rho} \qquad 77.$$

Increasing $u > 1$ increases the coherence but the maximum of γ_{12} is 0.14 and is therefore regarded as incoherent.

Equation (77) expresses the experimental conditions under which interference fringes may be observed using initially incoherent light. If it is required that an initially spatially coherent source (which implies monochromatic) be made spatially incoherent, then a suitable method is suggested by the assumption used in connection with the second term of Equation (66).

If the source σ of Figure 28 is now considered to be monochromatic, then the last term of Equation (66) may be expressed, using Equation (34)

$$\Sigma\, \overline{V_{m1}(t)\ V^{*}_{n2}(t)} \;=\; u_{m1} u^{*}_{n2} \tag{78.}$$

where

$$V_{m1}(t) = U_{m1}\, e^{-i2\pi\bar{\nu}t},\; V_{n2} = U_{n2}\, e^{-i2\pi\bar{\nu}t} \tag{79.}$$

and

$$U_{m1} = A_{m1}\, e^{gm1},\; U_{n2} = A_{n2}\, e^{gn2} \tag{80.}$$

substituting (81) into (79)

$$U_{m1}\, U^{*}_{n2} = A_{m1} \cdot A_{n2} \cos(g_{n2} - g_{m1}) \tag{81.}$$

The spatial coherence of the waves is implicit in the time independence of the phase difference $(g_{n2} - g_{m1})$ which may give a non-zero value for Equation (78).

By making either or both of the arguments of Equations (80) time dependent the phase difference and thus the spatial position of the interference fringes will become time dependent. By ensuring that the rate of movement of the fringes exceeds the temporal resolution of observation, then no fringes will be observed and the waves will be incoherent. The experimental arrangement for inserting this time

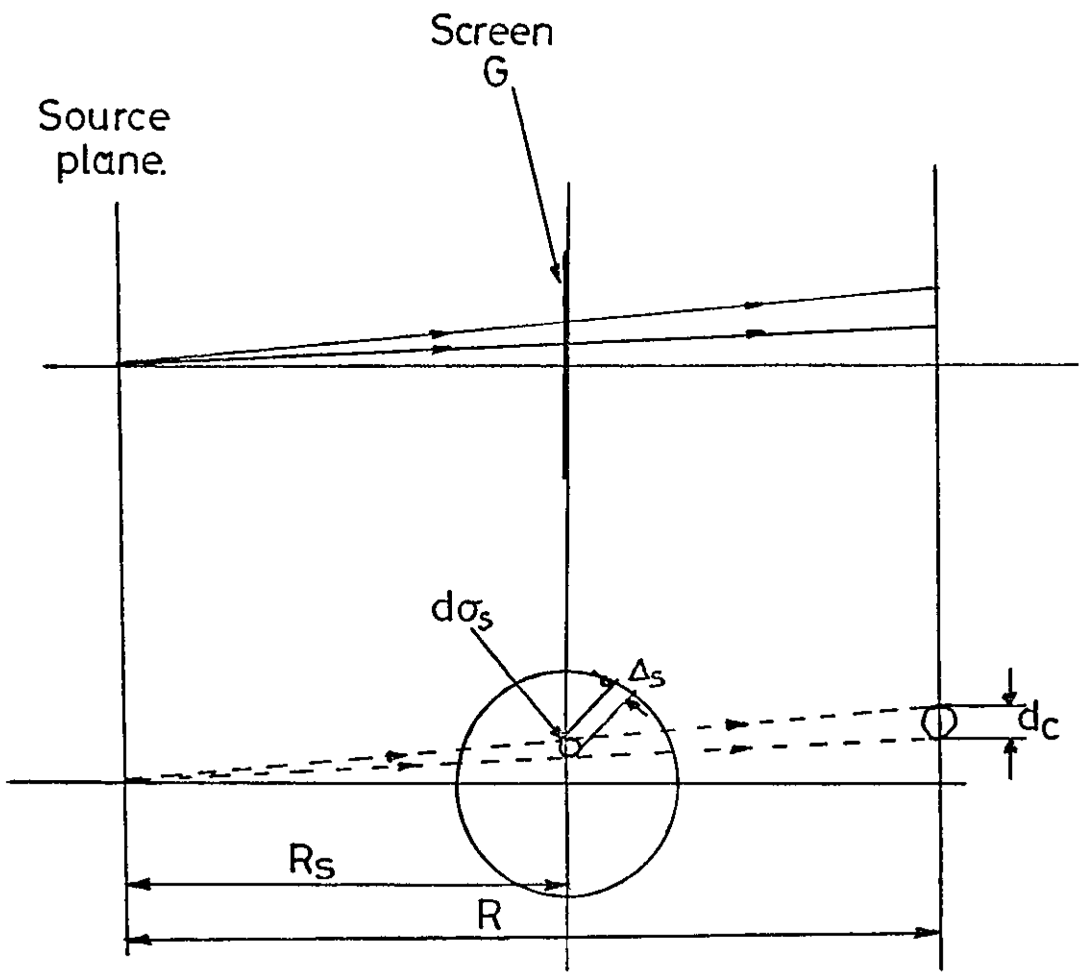


Figure 29. Diagrammatic form of the apparatus for varying the spatial coherence of the light from the source plane.

dependent phase shift is shown in Figure 19 and is shown in diagrammatic form in Figure 29.

The ground glass screen G has a spatial dependent phase characteristic which may be characterised by a unit cell of area $d\sigma_s$ of constant phase but which has a random phase relationship with all other cells. The laser and associated optics may be represented as a point source of monochromatic light a distance R_s from the screen G. The cell $d\sigma_s$ is assumed to be circular of diameter Δ_s and light from the source passing through a single cell will illuminate a circle of diameter

$$d_c = \frac{R\,\Delta_s}{R_s} \qquad 82.$$

on the plane parallel to the screen a distance R from the source.

The rotation of the screen imparts a time dependent phase shift to the light within this circle. Since the phase shift due to the ground glass screen will be random, it is assumed that the light in any other circle will be incoherent. Under these conditions the spatial coherence of the laser light is confined to a circle of diameter d_c which may be varied experimentally through the adjustment of R_s or R. The effects of the diffraction of the light by the phase variations of the screen have been neglected in this simple analysis.

3.7.1 Diffraction

An electromagnetic wave occupies a finite volume of space and the volume boundary is determined by the points in space at which the wave components E, H have negligible values. Any attempt to obstruct a wave or to confine it to within a smaller volume results in the perturbation of its wave components and produces secondary waves which give rise to the diffraction effects.

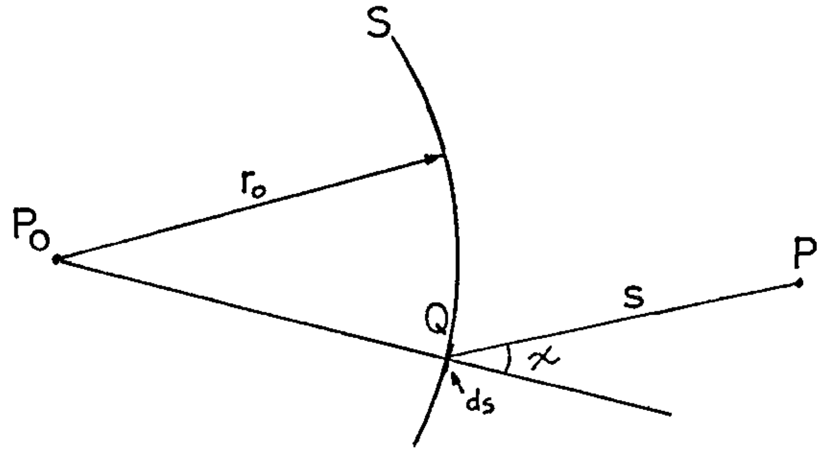


Figure 30a. S is a wavefront originating from the point source P_o and χ, s are the variables associated with the light from a section Q of the wavefront which contributes light to the point P.

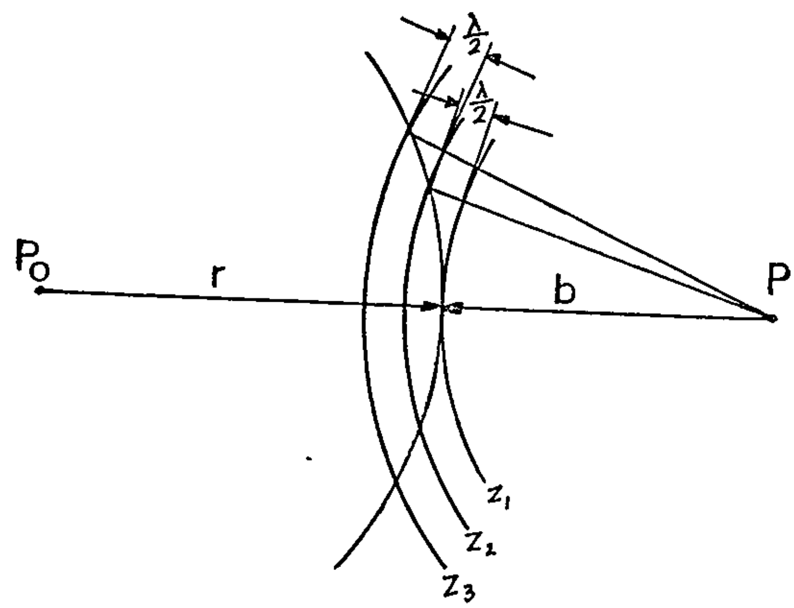


Figure 30b. The Fresnel zone construction on the wavefront of a point source.

The rigorous determination of the secondary waves involves the solution of the wave equations subject to the boundary conditions imposed on the wave components by the obstructions[31 Ch.11]. It is, however, sufficient for most optical problems to use the Fresnel-Kirchoff diffraction theory to obtain a result which is experimentally indistinguishable from the rigorous theory.

The Fresnel-Kirchoff diffraction theory was originated by Fresnel who explained diffraction phenomena in terms of the Huygens wavelet construction and interference. Huygens asserted that each element of a wavefront may be regarded as the centre of a secondary disturbance which radiates spherical waves and the wavefront at any later time is the envelope of all such wavelets. Fresnel included the effects of the mutual interference of the wavelets to predict with accuracy the diffraction effects produced by simple obstructions or apertures placed in the path of simple waves.

3.7.2 Huygens-Fresnel Diffraction Theory

Consider a point source of monochromatic waves at P_o shown in Figure 30a and let S be the instantaneous position of the wavefront, radius r_o. The disturbance at P due to the element dS at Q of the wavefront a distance s from P is given by

$$dU(P) = K(\chi)\ \frac{Ae^{ikr_o}}{r_o}\ \frac{e^{iks}}{s}\ dS \qquad 83.$$

where $K(\chi)$ is an inclination factor describing the variation with angular direction χ of the amplitude of the secondary waves originating at Q. Fresnel assumed that $K(\chi)$ is maximum for $\chi = 0$ and zero for $\chi = \frac{\pi}{2}$. The total disturbance at P will be given by

$$U(P) = \frac{Ae^{ikr_o}}{r_o} \iint_S \frac{e^{iks}}{s}\ K(\chi)\ dS \qquad 84.$$

To evaluate Equation (84) the wavefront is split into Fresnel zones z_1, z_2 .... as shown in Figure 30b, where the boundary of each zone is given by a sphere centred on P with radii b, $b + \frac{\lambda}{2}$, $b + \frac{2\lambda}{2}$ . .... Following the analysis of Born & Wolf p(31 pp.372) it is found that the contributions of adjacent zones to U(P) are approximately equal in magnitude but of opposite sign. If all zones are assumed to contribute to U(P) then

$$U(P) = i\lambda(K_1 \pm K_n) \frac{A\, e^{ik(r_o+b)}}{r_o+b} \qquad 85.$$

that is only the first and last zones make significant contributions, the intermediate zones cancelling each other out. Using the assumption of Fresnel that $K_n = K(\frac{\pi}{2}) = 0$ Equation (85) reduces to

$$U(P) = i\lambda K_1 \frac{A\, e^{ik(r_o+b)}}{(r_o+b)} . \qquad 86.$$

This will agree with the effects of a spherical wave if

$$K_1 = -\frac{i}{\lambda} . \qquad 87.$$

The contribution to U(P) by the j^{th} zone may be shown to be

$$U_j(P) = 2i\lambda(-1)^{j+1} K_j \frac{A\, e^{ik(r_o+b)}}{r_o + b} \qquad 88.$$

and Equation (86) may be written

$$U(P) = \tfrac{1}{2}U_1(P). \qquad 89.$$

That is, the total disturbance at P is equal to one half the disturbance due to the first zone alone. A similar result will be obtained if a screen with a circular aperture is placed perpendicular to P_oP with its centre on this line, such that only half of the first zone is unobstructed by the screen. Increasing the size of the opening until the whole of

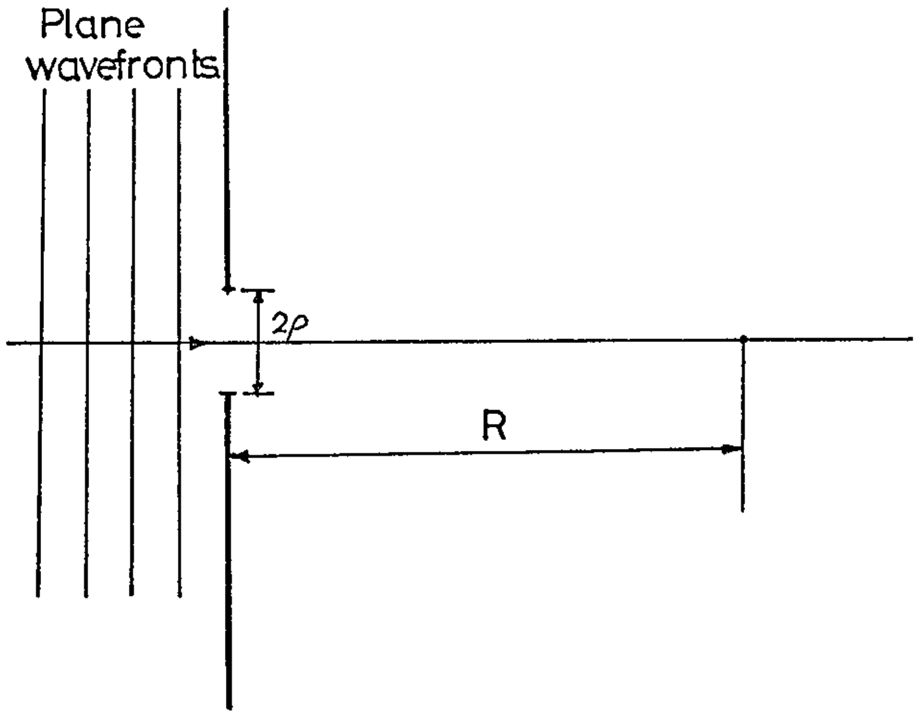


Figure 31. The circular aperture diffraction experiment.

the first zone is uncovered, will give

$$U(P) = \frac{2Ae^{ik(r_o+b)}}{r_o+b} \qquad 90.$$

and the intensity will be four times that obtained if the screen were absent.

If the first two zones are uncovered then U(P) = 0 and as more zones are uncovered the intensity at P will pass through maxima and zero values. In the experimental arrangement of Figure 31 a circular aperture, radius ρ, is illuminated by a plane monochromatic wave, wavelength λ. The distance R along the axis at which the maxima and zero values of intensity are observed are given by

$$R = \frac{\rho^2}{n\lambda} \qquad 91.$$

where $n = 1, 3, 5 \ldots..$ for maxima

$n = 2, 4, 6 \ldots..$ for zeros

Equation 91 is subject to the condition $R >> n\lambda$ which ensures that adjacent zones contribute equal (and opposite) contributions to the total disturbance at the point of observation.

The justification for Equation (87) is obtained from the generalisation of the Huygens-Fresnel diffraction theory due to Kirchoff. The Kirchoff diffraction theory also yields the correct form for the inclination factor $K(\chi)$.

3.7.3 Kirchoff's Diffraction Theory

Kirchoff's diffraction theorem expresses the solution of the wave equation at an arbitrary point P in the field in terms of the value of the solution and its first derivations at all points on an arbitrary closed surface surrounding the point P.

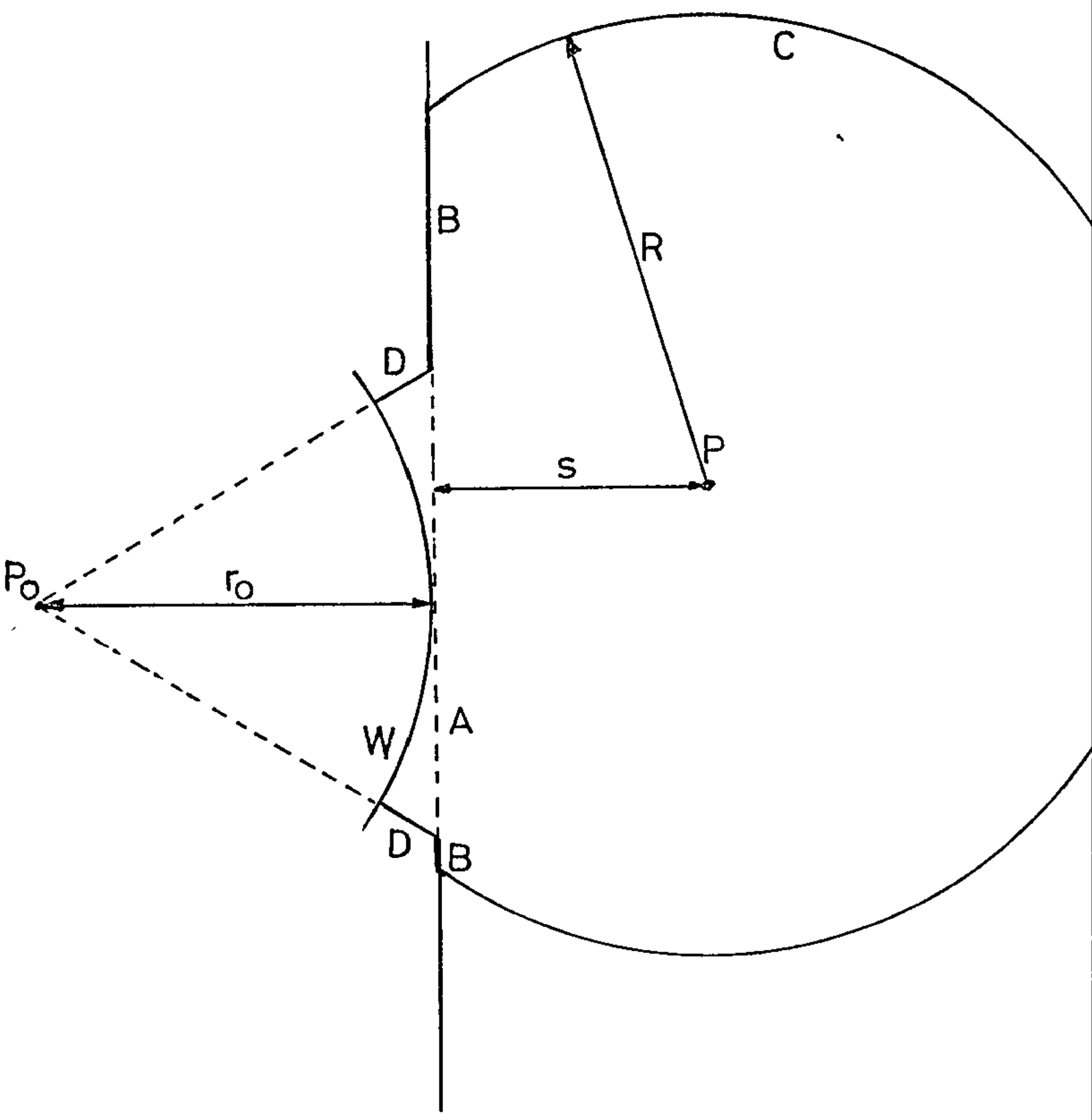


Figure 32. The boundarys of the Kirchoff diffraction theorem.

$$U(P) = \frac{1}{4\pi} \iint_S \left\{ \frac{Ud}{dn} \left(\frac{e^{iks}}{s} \right) - \frac{e^{iks}}{s} \frac{dU}{dn} \right\} dS. \qquad 92.$$

Equation (92) is called the Kirchoff diffraction integral, where S is the closed surface surrounding point P and the $\frac{e^{iks}}{s}$ term has the same interpretation as in Equation (83). To solve Equation (92) requires the values of U and $\frac{dU}{dn}$ at all points on the surface S where $\frac{d}{dn}$ denotes differentiation along the inward normal to S. The Kirchoff integral equation[31] is not derived here and is only used to justify the Huygens-Fresnel theory which is now shown to be a special case of the Kirchoff integral.

In Figure 32, a monochromatic point source P_o illuminates an aperture in an opaque screen a distance r_o from P_o and the point of observation P is a distance s from the screen. The surface of integration S consists of the aperture A and the opaque screen surface B which is bounded by a portion C of the surface of a sphere radius R centred at P.

The Kirchoff boundary conditions give the values of U and $\frac{dU}{dn}$ on these surfaces as,

$$\text{On A} \quad U = \frac{Ae^{ikr_o}}{r_o} \quad \frac{dU}{dn} = \frac{Ae^{ikr_o}}{r_o} \left(ik - \frac{1}{r_o}\right) \cos(n, r_o) \qquad 93a.$$

$$\text{B} \quad U = \frac{dU}{dn} = 0 \qquad 93b.$$

$$\text{C} \quad U = \frac{dU}{dn} = 0 \qquad 93c.$$

where $\frac{Ae^{ikr_o}}{r_o}$ is the value of the incident field in the aperture. The zero contribution from B is due to the opaqueness of the screen and the zero contribution from the surface C is assumed by considering a finite propagation time for the wave before which the field at radius R will be zero.

Substitution of Equation (93) into (92) yields the Fresnel-Kirchoff diffraction formula.

$$U(P) = \frac{-iA}{2\lambda} \iint_A \frac{e^{ik(r_o+s)}}{r_o s} \left(\cos(n,r_o) - \cos(n,s)\right) dS. \tag{94}$$

A portion W of the incident spherical wavefront within the aperture may be selected instead of the surface A and for r_o large, the portion D of the wavefront may be neglected. Then $\cos(n,r_o)=1$ and setting $\chi = \pi - (r_o s)$ in Equation (94)

$$U(P) = \frac{-i}{2\lambda} \frac{Ae^{ikr_o}}{r_o} \iint_W \frac{e^{iks}}{s} (1 + \cos\chi) \; dS. \tag{95}$$

Comparison of Equation (95) with (84) gives for the inclination factor the expression

$$K(\chi) = \frac{-i}{2\lambda} (1 + \cos\chi) \tag{96}$$

which for K(0) reduces to $K_1 = \frac{-1}{\lambda}$ as given by Equation (87).

The assumption of Fresnel that $K(\frac{\pi}{2}) = 0$ is clearly incorrect but the larger values of χ are not encountered in most diffraction problems.

A simplified form of Equation (95) is used in Section 5.5 to calculate the intensity distribution at a caustic surface.

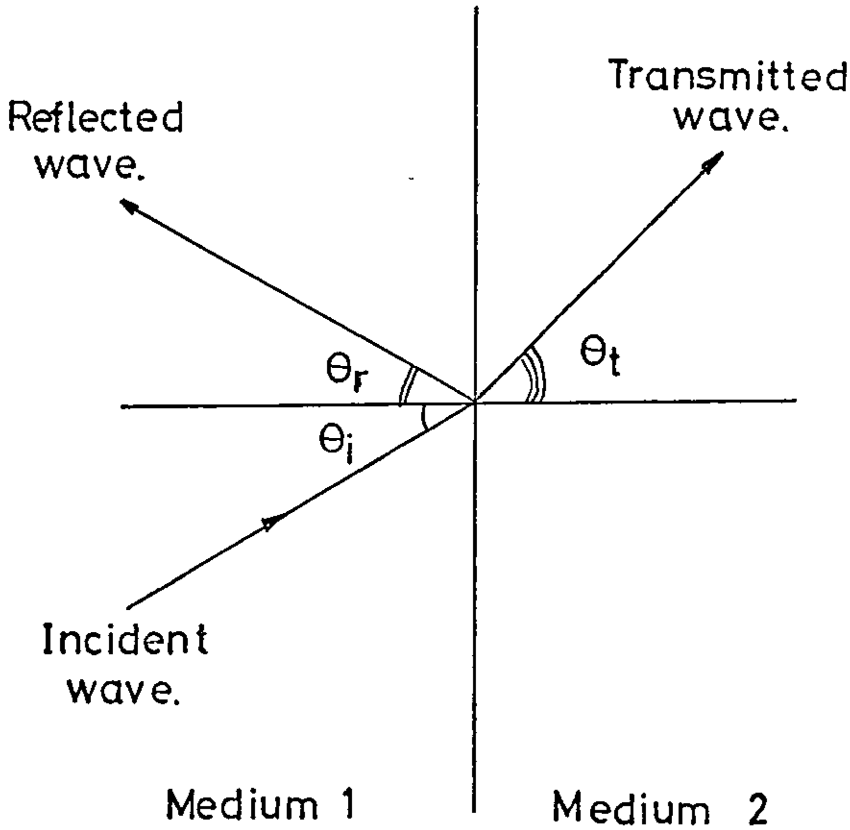


Figure 33. The interface between media of different permittivity showing the reflected and transmitted waves formed by an incident wave.

3.8 The Fresnel Equations

The optical theory considered so far in this chapter has been concerned with wave propagation in homogeneous media where the material constants are independent of position. A simple inhomogenous medium contains a step change of a material constants, for example from ε_1 to ε_2 as shown in Figure 33. The magnetic material constant μ is assumed not to vary significantly from its free space value $\mu o = 1$ for the media considered in this thesis.

The derivation of the Fresnel equations is given in most optical theory textbooks(31-34) and only an outline of the theory and the results are given here. The Fresnel equations describe the complex amplitudes of the wave components which arise in the two media (1, 2) in Figure 33 when a monochromatic plane wave is incident upon the interface between the two media. The derivation of the equations is based upon the boundary conditions which require continuous electric field components tangential to the interface and in the absence of currents, a similar condition for the tangential magnetic field components.

Consider a plane wave of complex amplitude A incident upon the interface from medium 1 at angle θ_1 shown in Figure 33. To satisfy the boundary conditions, two resultant plane waves are postulated, a reflected wave of complex amplitude R and a transmitted wave (refracted wave) of complex amplitude T. Both waves lie in the plane of the incident wave and make angles with the normal at the interface of θ_r and θ_t respectively.

By resolving the three waves into their xyz components and using the condition that at every point on the interface the time variation of all three waves will be the same, the following laws of reflection and refraction may be derived.

Law of reflection $$\theta_r = \pi - \theta_i \cdot \tag{97.}$$

Law of refraction $$\frac{\sin \theta_i}{\sin \theta_t} = \frac{v_1}{v_2}. \tag{98.}$$

Where, from Equation (9),

$$v_1 = \frac{1}{\sqrt{\varepsilon_1 \mu_1}} \qquad v_2 = \frac{1}{\sqrt{\varepsilon_2 \mu_2}}$$

Defining the refractive index n as

$$n = \frac{c}{v} \quad \text{then} \quad n_1 = \frac{c}{v_1} \qquad n_2 = \frac{c}{v_2} \tag{99.}$$

where c is the velocity of waves in free space.

The law of refraction may be written using Equations (98), (99) as

$$n_1 \sin \theta_i = n_2 \sin \theta_t \tag{100.}$$

which is the form known as Snell's Law.

In the section on Polarisation it was shown that an arbitrarily polarised wave may be considered in terms of two waves, one linearly polarised parallel to the plane of incidence, denoted by subscript p, and the second linearly polarised normal to this plane, denoted by subscript n. The field components of each of these waves independently satisfies the boundary conditions and the solutions of the boundary condition equations gives the following amplitudes for the reflected and transmitted waves.

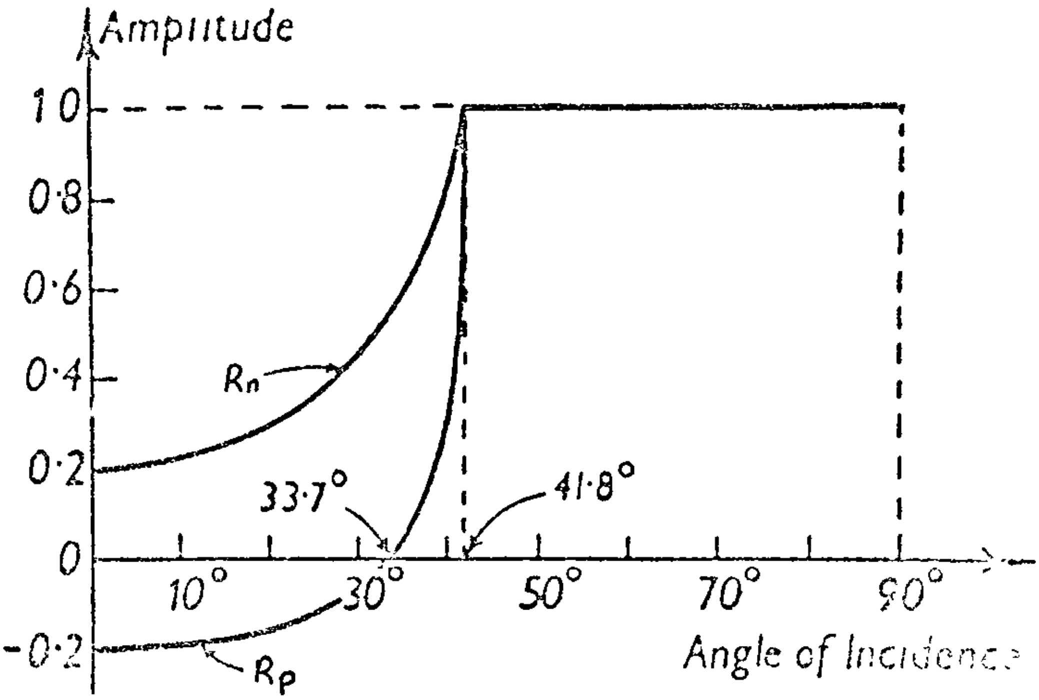


Figure 34b. Graphs showing the amplitudes of the reflected waves at a total internal reflection interface.

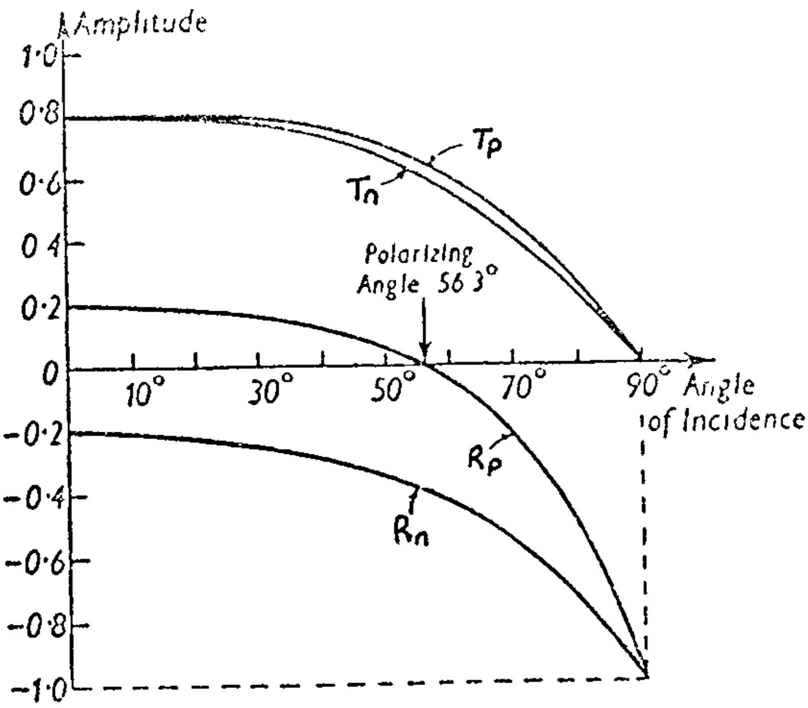


Figure 34a. Graphs showing the amplitude of the refracted and reflected waves at a dielectric interface.

$$T_p = \frac{2n_1 \cos\theta_i}{n_2 \cos\theta_i + n_1 \cos\theta_t} A_p$$

101.

$$T_n = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t} A_n$$

$$R_p = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_2 \cos\theta_i + n_1 \cos\theta_t} A_p$$

102.

$$R_n = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} A_n$$

If $n_2 > n_1$ then from Equation (100), θ_t is real for all θ_i and Figure 34a shows the amplitudes of the reflected and refracted waves according to Equations (101), (102) for $A_p = A_n = 1$ and $\frac{n_2}{n_1} = 1.5$.

When $n_1 > n_2$ then for $(\frac{n_1^2}{n_2^2}) \sin^2\theta_i > 1$, θ_t is imaginary and Equation (102) may be expressed in the form

$$R_p = \frac{n_2^2/n_1^2 \cos\theta_i - i\sqrt{\sin^2\theta_i - n_2^2/n_1^2}}{n_2^2/n_1^2 \cos\theta_i + i\sqrt{\sin^2\theta_i - n_2^2/n_1^2}} A_p$$

103.

$$R_n = \frac{\cos\theta_i - i\sqrt{\sin^2\theta_i - n_2^2/n_1^2}}{\cos\theta_i + i\sqrt{\sin^2\theta_i - n_2^2/n_1^2}} \cdot A_n$$

Either of the Equations (103) may be expressed in the form $\frac{a - ib}{a + ib}$ which in polar form becomes $1\underline{/\delta}$ where $\text{TAN}\ \frac{\delta}{2} = \frac{b}{a}$.

Then

$$|R_p| = |A_p| \quad \text{and} \quad |R_n| = |A_n|. \qquad 104.$$

The condition expressed by Equation (104) is called total internal reflection and indicates that the energy of the incident wave $\alpha\overline{|A^2|}$ is totally reflected at the interface of the two dielectrics. The amplitudes of the reflected fields for Equation (104) are shown in Figure 34b where $A_p = A_n = 1$ and $\frac{n_1}{n_2} = 1.5$. It may be shown that the transmitted field amplitudes T_p, T_n are not zero when total internal reflection occurs, but produce an evanescent field whose amplitude decays rapidly with distance from the interface and which in terms of a time average value propagates no energy away from the interface.

Since in general the parallel and normal components of incident waves have different reflection and refraction amplitudes the state of polarisation of the incident wave will be changed on reflection or refraction. Natural light which is initially unpolarised will become partially polarised after reflection or refraction.

3.9 Geometrical Ray Theory

The analysis of the propagation of a plane wave by consideration of the behaviour of its parallel constant phase loci suggests a simpler representation in which the locus of the tangent to the normal of the constant phase loci will indicate the path taken by the wave. This locus is called a geometrical ray and its orientation may be described using the rules of geometry together with the laws of reflection and refraction.

The plane wave may be considered to be composed of small elements, where the propagation of each element is represented by its own geometrical ray, and the energy carried by each element is assigned to its ray. The intensity of each element after propagation will depend upon the area of the element at the point of observation relative to its initial area.

To determine a limit to the element size, the solution to the 'diffraction by a circular aperture' problem is considered. It is clear from Section 3.7 that confinement of a wave to a small aperture produces rapidly varying amplitudes in the immediate vicinity of the aperture, but at a considerable distance from the aperture the majority of the incident energy is confined within a circular region whose dimensions are wavelength dependent.

If the wavelength of the light is permitted to fall to zero, then it may be shown[31 Ch.3] that the diffraction effects disappear and the propagation of the light through the aperture is due to rectilinear propagation only. This gives rise to sharp boundaries between the illuminated and shadow regions and permits the selection of an arbitrarily small element within the illuminated region as the ray element.

Consider a space dependent component of a wave equation solution of the form

$$U(r) = R\{a(r)e^{iL(r)k_o}\} \qquad 105.$$

where a(r) and L(r) are both real function which vary slowly with respect to the wavelength of the light.

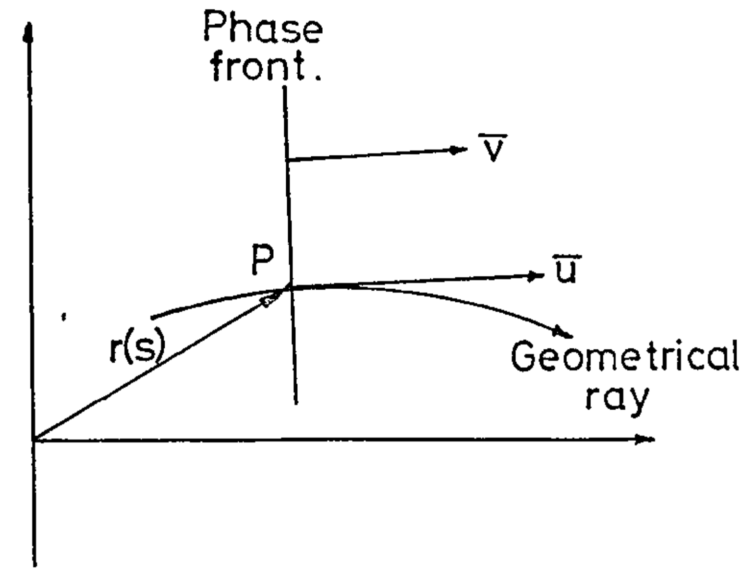


Figure 35a. The relationship between the direction of propagation of the phase front of a wave and a geometrical ray.

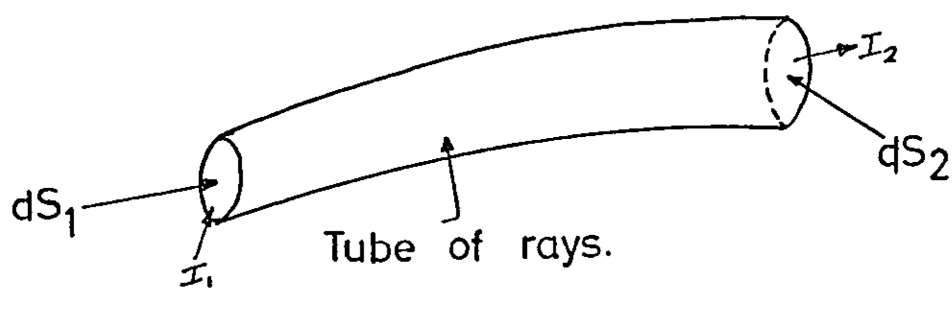


Figure 35b. A tube of geometrical rays, of variable cross section.

The free space wave number

$$k_o = \frac{\omega}{c}. \qquad 106.$$

Substitution of Equation (105) into the time independent wave equation yields

$$k_o^2\left(\frac{k^2}{k_o^2} - \nabla^2 L\right)a - ik_o(2\nabla L\nabla a + a\nabla^2 L) + \nabla^2 a = 0. \qquad 107.$$

The middle term of the LHS of Equation (107) must independently vanish and the last term is negligible compared with the first term since $k \propto \frac{1}{\lambda}$ and λ is assumed to be tending to zero.

When the first term vanishes

$$\nabla^2 L = n^2. \qquad 108.$$

where from Equation (99)

$$n^2 = \frac{c}{\omega} = \frac{k^2}{k_o^2}.$$

Equation (108) is known as the eikonal equation and will permit the calculation of the surfaces of constant phase given by the scalar function

$$L(r) = \text{CONSTANT} \qquad 109.$$

From the opening comments of this section the geometrical rays are the normals to the constant phase surfaces. If the position of a point P on a ray is denoted by a position vector r(s) considered as a function of the distance s along the ray, see Figure 35a then the direction of the ray at point P is given by the unit vector $\overline{u}$

$$\overline{u} = \frac{dr}{ds}. \qquad 110.$$

The vector, v, perpendicular to the phase front L(P) is given by

$$v = \nabla L. \qquad 111.$$

A unit vector $\overline{v} = \frac{v}{|v|}$.

where $$|v| = n. \qquad 112.$$

and the unit vector $\overline{v}$ is parallel to unit vector $\overline{u}$ to give

$$\overline{u} = \frac{\nabla L}{n}. \qquad 113.$$

The equation of a ray is thus

$$\frac{n dr}{ds} = \nabla L. \qquad 114.$$

The optical path length, ΔL, between two points P_1, P_2 on a ray is given by

$$\Delta L = \int_{P_1}^{P_2} n \, ds = L(P_2) - L(P_1). \qquad 115.$$

which for a homogenous medium

$$\Delta L = ns. \qquad 116.$$

and for free space

$$\Delta L = s. \qquad 117.$$

If a tube of rays enclose a surface area dS on a constant phase surface of a wave, intensity I, then the intensity law of geometric optics states that IdS is constant along the tube of rays. From Figure 35b

$$I_1 \, dS_1 = I_2 \, dS_2. \qquad 118.$$

Since each ray represents a local plane wave, the laws of refraction and reflection will apply directly to the propagation of rays across dielectric interfaces. Similarly the divisions of intensity as given by the Fresnel Equations will also apply.

3.10 Conclusions

The characteristics of light waves have been established in this chapter and it has been shown that the path of a plane light wave may be represented by a geometrical ray of light. In the following chapters the propagation of light in optical waveguides is analysed using primarily the geometrical ray theory. However, where interference or diffraction effects are discussed with reference to rays of light it should be understood that the various effects are due to the local plane waves whose propagation paths in the waveguide have been represented by geometrical rays.

CHAPTER 4

4.1 Introduction

The principle features of a geometrical ray analysis of light propagation in core cladded type optical waveguides are well established and contained in at least two textbooks[(3-4)]. The assumptions used in such an analysis are examined in this chapter and a suitable representation of the waveguide is established for use in later chapters. This representation is based upon the behaviour of optical fields at dielectric interfaces according to the Fresnel Equations, and the laws of refraction and reflection.

In addition to consideration of the propagation process, the behaviour of light fields at the entrance aperture of the waveguide is examined. It is shown that the effects of the entrance aperture are similar to those produced by a diffracting aperture of similar dimensions to the waveguide core placed in a medium of core refractive index. The effects of the exit aperture are generally not significant in this study since the fields of interest are within the aperture, (the near field).

The chapter concludes by considering the visible effects of a ramp refractive index profile between the core and cladding materials.

4.2 Geometrical Ray Theory Model of the Core Cladded Waveguide

An idealised core cladded waveguide consists of a dielectric cylindrical core, radius a, refractive index n_1 inside a dielectric tube, inner radius a, wall thickness b, and with refractive index n_2, where $n_2 < n_1$. In general $n_2 > n_o$ where n_o is the refractive index of the medium surrounding the waveguide, for free space $n_o = 1$.

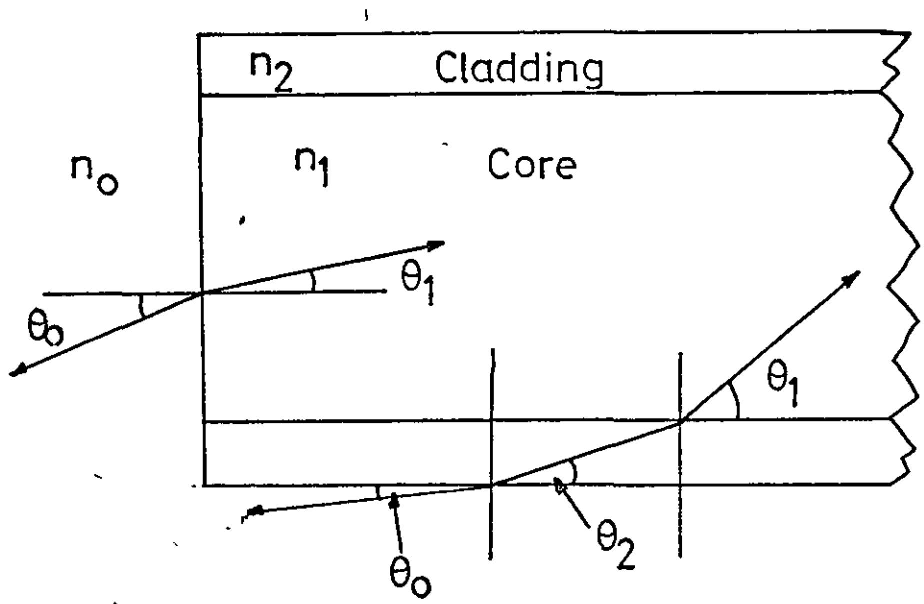


Figure 36. Longitudinal cross section of a cladded dielectric waveguide.

Point of incidence	θ_o	
Core	$\sin\theta_o \leq \frac{1}{n_o}(n_1^2 - n_2^2)^{\frac{1}{2}}$	Reflected at $n_1 - n_2$ interfa
Core	$\sin\theta_o > \frac{1}{n_o}(n_1^2 - n_2^2)^{\frac{1}{2}}$ $\leq \frac{1}{n_o}(n_1^2 - n_o^2)^{\frac{1}{2}}$	Reflected at $n_2 - n_o$ interfa
Core	$\sin\theta_o > \frac{1}{n_o}(n_1^2 - n_o^2)^{\frac{1}{2}}$	Radiates
Cladding, entrance aperture	$\sin\theta_o \leq \frac{1}{n_o}(n_2^2 - n_o^2)^{\frac{1}{2}}$	Reflected at $n_2 - n_o$ interface
Cladding, entrance aperture	$\sin\theta_o > \frac{1}{n_o}(n_2^2 - n_o^2)^{\frac{1}{2}}$	Radiates
Cladding wall		All rays radiate

Table 2.

Three dielectric interfaces are identified, $n_o - n_1$, $n_o - n_2$, $n_1 - n_2$ and since the incident radiation may originate in either medium a total of six refracting conditions exist. These are indicated in Figure 36 in which a meridional longitudinal cross section of the fibre is shown. The $n_o - n_1$ interface accounts for a transmission loss at the entrance and exit apertures of the waveguide and a refraction of light into and out of the waveguide core. The $n_1 - n_2$ interface provides a total internal reflection condition for light within the core and precludes light which may propagate in the core from entering the core except through the end apertures. The $n_o - n_2$ interface is not significant in the geometrical ray theory model of the waveguide although it too provides a total internal reflection condition.

The total internal reflection condition at the $n_1 - n_2$ interface for rays of light entering the core through the entrance aperture at incident angle θ_o is given by

$$\sin \theta_o \leqslant \frac{1}{n_o} (n_1^2 - n_2^2)^{\frac{1}{2}} \qquad 119.$$

Rays of light whose incident angles exceed the meridional critical angle θ_{oc} given by the equality condition of Equation (119), are refracted into the cladding at angle θ_2 where

$$\cos \theta_2 = \frac{1}{n_2} (n_1^2 - n_o^2 \sin^2\theta_o)^{\frac{1}{2}} \qquad 120.$$

The total internal reflection condition at the $n_o - n_2$ interface is given by

$$\cos \theta_2 \geqslant \frac{n_o}{n_2} \qquad 121.$$

Rays which exceed the equality condition of Equation (121) are radiated from the waveguide. Rays with angles within a total internal reflection condition will in theory be reflected unattenuated, but in practice the $n_o - n_2$ interface has a high reflection loss because of its exposure to dirt and damage in the free space environment. The $n_1 - n_2$ interface is protected from free space and is therefore assumed to permit lossless reflection.

Rays of light entering the cladding directly from free space, through the entrance aperture will suffer total internal reflection at the $n_2 - n_o$ interface if

$$\sin\theta_o \leq \frac{1}{n_o}(n_2^2 - n_o^2)^{\frac{1}{2}} \qquad 122.$$

otherwise they will radiate.

Rays incident upon the cladding walls will refract at the $n_o - n_2$ and $n_2 - n_1$ interfaces, pass through the core and refract out of the waveguide through the $n_1 - n_2$ and $n_2 - n_o$ interfaces and thus radiate.

These various ray paths and conditions are summarised in Table 2 and it is clear that only rays within the meridional critical angle θ_{oc}, incident in the core entrance aperture, will propagate unattenuated along the waveguide. Near the entrance aperture other rays may contribute to the intensity within the core depending upon the reflection and transmission coefficients at the various interfaces. This is considered in the next section.

It is proposed to represent this meridional cross section of the core of the waveguide as two plane parallel mirrors, with unity reflection coefficients, a distance 2a apart with the space between

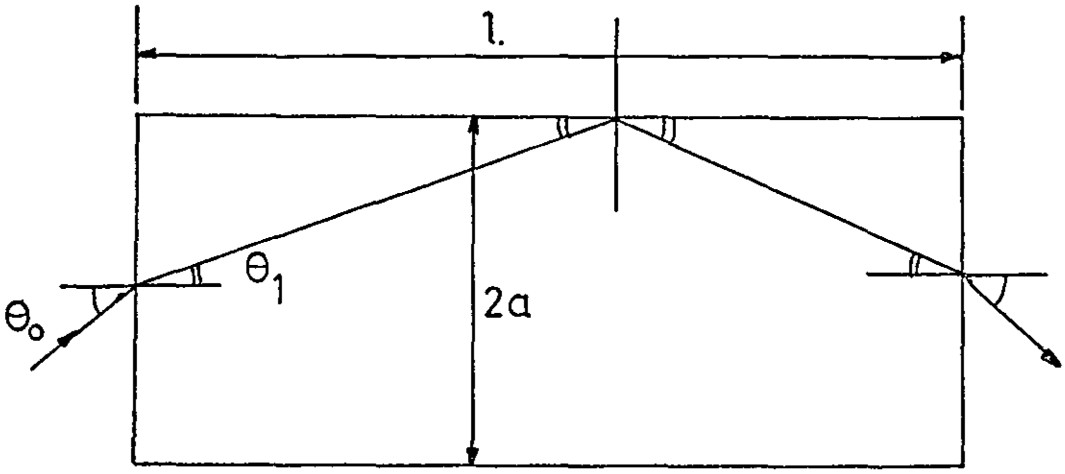


Figure 37. Parallel mirror representation of a cladded waveguide.

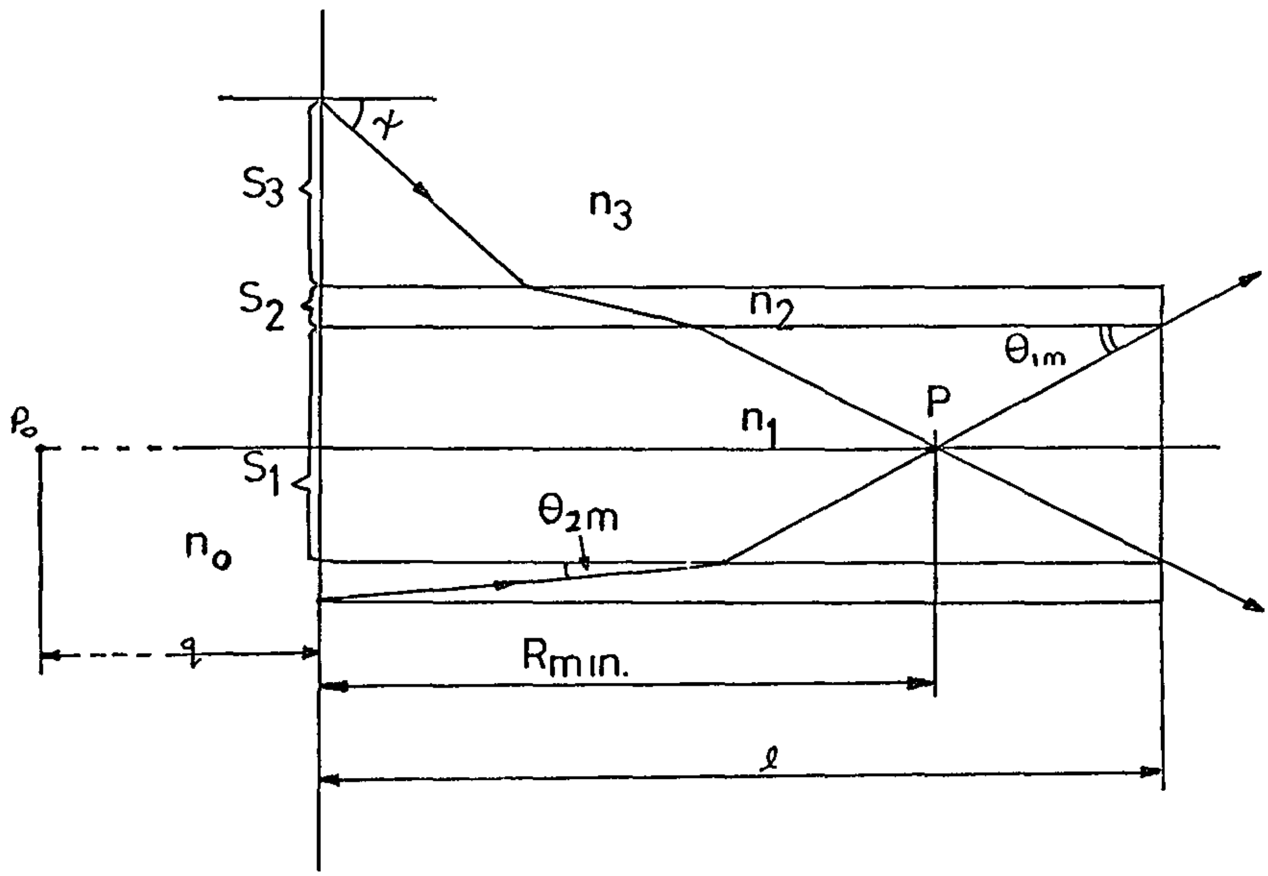


Figure 38. Longitudinal cross section of cladded dielectric waveguide embedded in a medium of refractive index n_3.

them filled with a lossless medium of refractive index n_1.

The characteristics of a typical ray propagating in this model of the core shown in Figure 37 are as follows. The optical path length L of a ray with axial angle θ_1 in the core is given by Equation (116) as

$$L = \frac{n_1 \ell}{\cos\theta_1} . \qquad 123.$$

The number of reflections made by this ray in a length ℓ of core is given by

$$m_\ell = \frac{\ell \tan\theta_1}{2a} . \qquad 124.$$

The direction of the ray is reversed after each reflection but the magnitude of θ_1 is constant. It is further assumed that no phase shift occurs on reflection, although this is rigorously true only for reflections at the critical angle.

4.3.1 Entrance Aperture Diffraction

Consider the experimental arrangement shown diagrammatically in Figure 38 where a short length of waveguide ($\ell \approx 12a$) is embedded in resin, refractive index n_3 (where $n_3 > n_2$) and polished at both ends. A monochromatic point source at P_o is a distance q from the waveguide aperture and it is assumed that q >> 2a so that approximately plane waves are incident over the waveguide aperture and surrounding surface.

The total disturbance at a point P will be the summation of the contributions from secondary sources on the surface S_1 the waveguide core, S_2 the cladding and S_3 the embedding resin. The maximum axial angle θ_{1m} in the core of a contribution to P is limited by the condition that no reflection of the light occurs prior to radiation from the core.

From Figure 38

$$\sin\theta_{1m} = \frac{a}{((\ell-R)^2 + a^2)^{\frac{1}{2}}} \qquad 125.$$

For a ray satisfying Equation (125) to enter the core through the cladding it must first exceed the critical angle θ_{1c} in the core given by

$$\sin\theta_{1c} = \frac{1}{n_1}\,(n_1^2 - n_2^2)^{\frac{1}{2}} \qquad 126.$$

Equating Equation (125) and (126) yields a minimum value for R

$$R_{MIN} = \ell - \frac{a}{\{(\frac{n_1}{n_2})^2 - 1\}^{\frac{1}{2}}} \qquad 127.$$

The maximum axial angle at the core cladding interface of a ray from a secondary source on the surface S_2 contributing to P at R_{MIN} is given by θ_{2m} where

$$\cos\theta_{2m} = \frac{R_{MIN}}{(R_{MIN}^2 + b^2)^{\frac{1}{2}}} \qquad 128.$$

The transmission factors at this $n_2 - n_1$ interface are given by substituting θ_{1c} and θ_{2m} into Equation (101).

$$T_{P2-1}/A_P = \frac{2n_2\sin\theta_{2m}}{n_1\sin\theta_{2m} + n_2\sin\theta_{1c}} \qquad 129.$$

for linearly polarised parallel components of the ray and

$$T_{n2-1}/A_n = \frac{2n_2 \sin \theta_{2m}}{n_2 \sin \theta_{2m} + n_1 \sin \theta_{1c}} \qquad 130.$$

for normal components.

Substitution of typical experimental values into Equation (127), $\ell \simeq 12a$,, $n_1 = 1.62$, $n_2 = 1.52$, yields a value $R_{MIN} \simeq 9a$ and since $b \simeq \frac{a}{5}$ then $\cos \theta_{2m} \simeq 1$ also $\sin \theta_{1c} \simeq 0.35$. Substitution of these last two values into Equation (129), (130) gives a transmission factor of $\simeq 0$ for both components of the ray.

It is therefore assumed that no contribution to the disturbance at P originates from the surface S_2. The radiation from the secondary sources on S_3 will have to cross the $n_2 - n_3$ interface in addition to the $n_2 - n_1$ interface in order to contribute to P. Since $n_3 > n_2$ the $n_3 - n_2$ interface will cause total internal reflection of radiation if it is incident at less than the critical angle. If a ray just exceeds this critical angle it will appear in the cladding with an incidence angle similar to that given by Equation (128) and will suffer the same transmission loss at the $n_1 - n_2$ interface as rays from surface S_2. Rays from surface S_3 which arrive in the core with incident angles greater than the core critical angle will originate from the surface S_3 a considerable distance from the $P - P_o$ axis and it is assumed that they will have small amplitudes due to the value of the inclination factor $K(\chi)$ of secondary sources.

The disturbance at P will therefore be the result of the summation of the radiation from secondary sources in the entrance aperture of the core only. This situation arises in the problem of diffraction by a circular aperture considered in Section 3.7 except that, in this case,

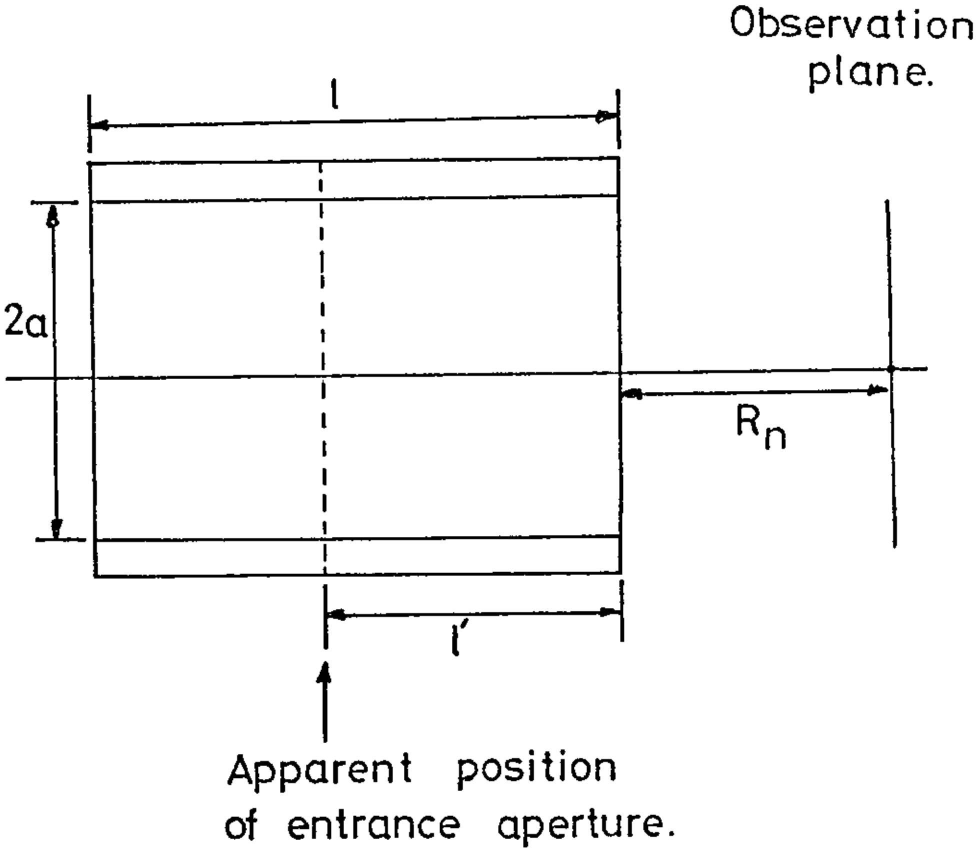


Figure 39. The displacement of the entrance aperture of cladded waveguide due to refraction.

the circular aperture is followed by a block of dielectric material, length ℓ and refractive index n_1.

From the simple theory of imaging by refraction at single surfaces[33 ch.282] this block will have the effect of changing the position of the entrance aperture with respect to the radiation end of the waveguide, as shown in Figure 39. The apparent position of the aperture ℓ' is given by

$$\ell' = \frac{\ell}{n_1} \tag{131.}$$

A simple experiment to test the equivalence of the entrance aperture of the waveguide to a circular aperture followed by a dielectric block is to observe the positions of the maxima and minima of intensity on the axis of the diffraction pattern of a circular aperture as predicted by the Fresnel Zone construction. If in Figure 39, R_n is the distance from the radiation end of the guide, then from Equation (91) a maximum of intensity will be observed when R_n satisfies

$$\ell' + R_n = \frac{a^2}{n\lambda} \quad n = 1, 3, 5 \tag{132.}$$

and an adjacent minima will be observed at R_{n+1} when

$$\ell' + R_{n+1} = \frac{a^2}{(n+1)\lambda} \tag{133.}$$

The radius, a, of the aperture may be determined from these two measurements of R_n, independently of the value of ℓ'.

Subtracting Equation (133) from (132)

$$a = \{(R_n - R_{n+1})\ n(n + 1)\lambda\}^{\frac{1}{2}} \tag{134.}$$

The experimental arrangement of Figure 21, using the laser as a source, was used to test Equations (91) and (134) for various combinations of pinholes and dielectric blocks. The diameters of the pinholes were measured using the TV measuring system, and R_n values obtained using the dial gauge shown in Figure 17. The dielectric blocks were built up to the required thickness using microscope slides and measured using a micrometer. The slides were held together with a thin layer of index matching liquid and the pinhole attached to one face of the block.

4.3.2 Experimental Results

Three pinholes of nominal diameters 25μ, 50μ, and 100μ were illuminated with plane waves from the He-Ne laser source. It was found that the nominally circular pinholes were quasi-elliptical with diameters along the major and minor axis as given in Table 3.

Nominal Diameter (μ)	Major Axis Diameter (μ)	Minor Axis Diameter (μ)
100	96	88
50	48	47
25	26	25

TABLE 3

Each measurement made with the T.V. aided system is corrected to the nearest micron. The values of R for n = 2 to n = 5 in Equation (91) were measured for each of these pinholes to within ±2.5μ and the corresponding values of the diameters of the apertures are given in Table 4.

Nominal Diameter	n = 2	n = 3	n = 4	n = 5	Average Diameter
100	95.1	95.4	94.6	94.7	94.9
50	48.0*	48.5*	48.0*	49.0	48.4
25	27.5	26.1	26.1	26.2	26.5

TABLE 4

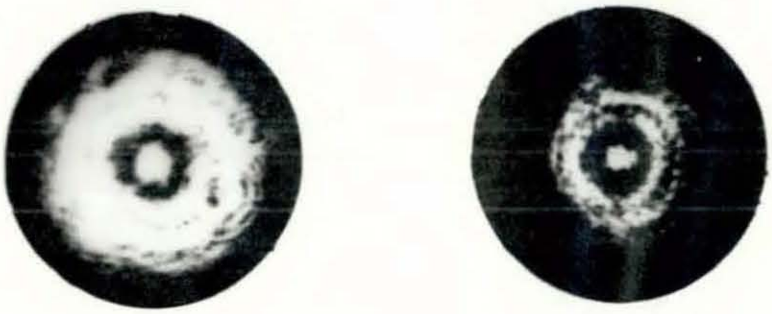
The 50μ diameter pinhole was attached to a block of glass made up from one to four microscope slides each of thickness 160μ and of refractive index n_1 = 1.524, measured using an Abbé refractometer. The values of $(\ell' + R_n)$ were measured for n = 2 to n = 5 as before and the corresponding aperture diameters were calculated using Equations (132) and (133) to give the results shown in Table 5.

Number of Slides	l	Calculated diameter for n = 2	n = 3	n = 4	n = 5	Average
1	160	46.6	47.7	48.2	48.3	47.7
2	320	47.7	47.7	47.2	47.0	47.4
3	480	49.6	48.5	48.0	47.7	48.5
4	640	48.7	48.9	49.2	47.7	48.7

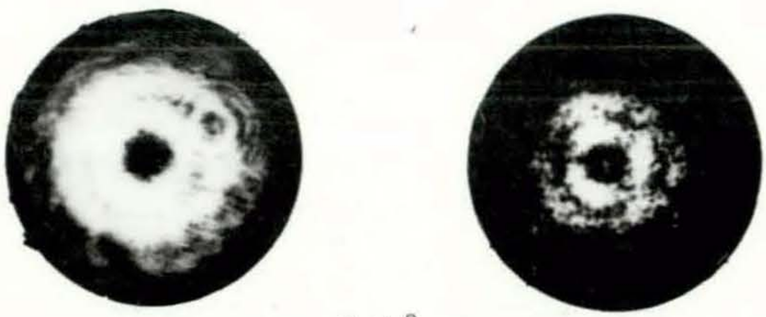
TABLE 5

Two short lengths of embedded fibres (ℓ = 310,330μ) were prepared as described in Chapter 2. The core diameters of three fibres from each sample were measured using the T.V. aided measuring system, whilst illuminating the fibres with white light incident at $\theta_o = 30^o$, in order to highlight the core cladding interface. The values of R_n (see Figure

n = 4



n = 3



n = 2

≃ 50μ diameter pinhole.

≃ 50μ diameter fibre.

Figure 40. Microphotographs of the diffraction patterns produced by a pinhole and a single fibre corresponding with the results marked with * in Table 6.

39) were measured for n = 2 to n = 5 as before and the value of ℓ' calculated from Equation (131) using n_1 = 1.62. The values of apparent core diameters were calculated using Equations (132), (133) for each of the six fibres and the results are given in Table 6.

l	l'	d Measured	n = 2	n = 3	n = 4	n = 5
310	192	50.0	43.8	46.9	48.2	50.2
310	192	48.0	42.0	46.0	46.5	47.7
310	192	49.0	42.7	45.7	46.1	47.7
330	203	45.0	40.7	42.5	43.0	44.0
330	203	46.0	41.6	43.1	43.6	44.6
330	203	50.0	45.1*	46.5*	47.3*	48.2

TABLE 6

Photomicrographs and intensity graphs for the experimental results marked with an asterisk are shown in Figure 40.

Discussion

The results shown in Table 4 suggest that Equation (91) is applicable for the calculation of the positions along the axis of the null points of the diffraction patterns of the nominally circular apertures. The average diameters of Table 4 are within $\pm$ 1.1μ of the major axis diameters of the apertures given in Table 3.

The results in Table 5 suggest that Equations (132), (133) may be used to find the null points of the aperture diffraction patterns when the aperture is followed by a block of glass. The average values of the diameters in Table 5 are all within $\pm$0.7μ of the average effective diameter of the aperture given in Table 4 as 48.4 micron.

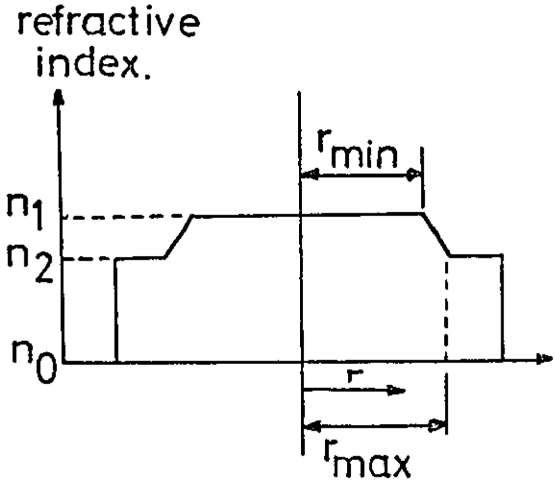


Figure 41. A Ramp refractive index

profile.

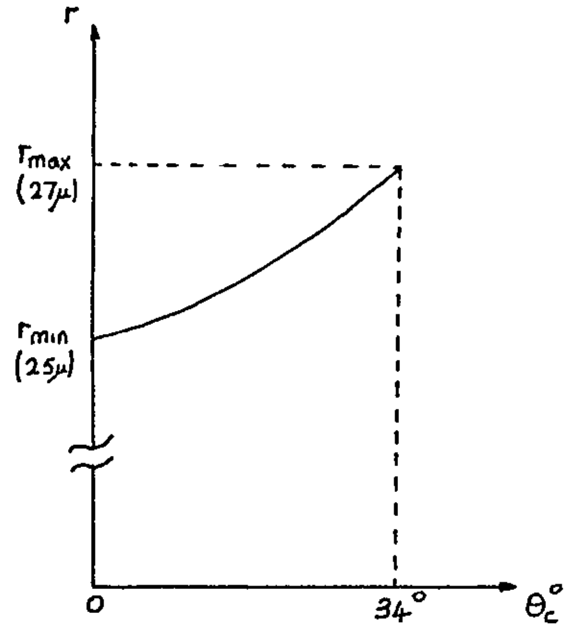


Figure 43. A Graph of r against

θ_c.

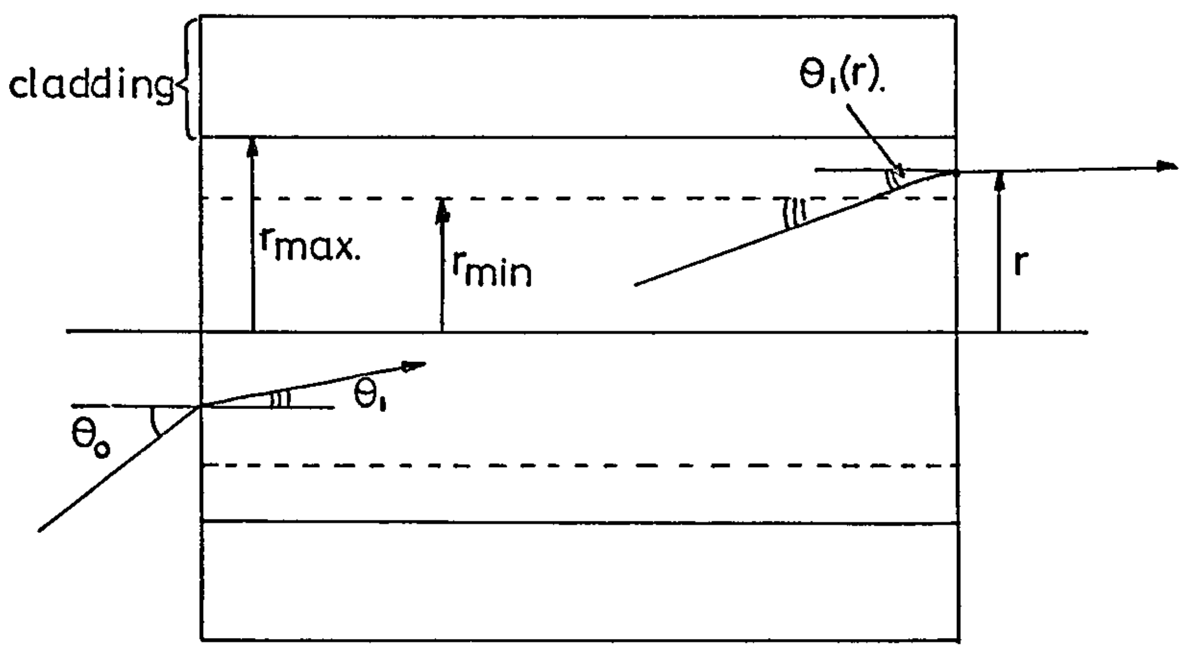


Figure 42. Rays in a longitudinal cross section of a ramp refractive

index profile waveguide.

The calculated diameters of the fibres for $n = 5$ in Table 6 are all within 1.8μ of the measured values, but at $n = 2$ the calculated values fall by between 2.9 to 6.3μ from their values at $n = 5$. This suggests that as the point of observation is moved away from the aperture the effective diameter of the aperture is reduced. One possible mechanism which will display this characteristic is a ramp refractive index profile between the core and cladding glasses, and this is discussed in the next section. The conclusions drawn there from the experimental results suggest that a variation in the effective diameter of the core of $(3 \pm 1)\mu$ between the $n = 2$ and $n = 5$ points of observation is possible in the fibres used for these experiments. Taking this result into account when considering the results in Table 6 suggests that Equations (132), (133) may be applied to the fibre diffraction pattern to find the axis null points.

4.4.1 Ramp Refractive Index Profile

The refractive index profile to be considered is shown in Figure 41. The refractive index $n(r)$ may be expressed as

$$n(r) = n_1 \qquad r \leqslant r_{min} \tag{135.}$$

$$n(r) = n_2 \qquad r \geqslant r_{max} \tag{136.}$$

$$n(r) = n_1 - \frac{\Delta n}{\Delta r}(r - r_{min}), \quad r_{min} < r < r_{max} \tag{137.}$$

where $\quad \Delta n = n_1 - n_2$ and $\Delta r = r_{max} - r_{min}$

A stepped refractive index profile has $\Delta r = 0$ to give a core radius a, where $a = r_{max} = r_{min}$ and Equations (135) and (136) correspond to the core and cladding regions respectively.

Consider a ray of light incident on the entrance aperture of a waveguide with a refractive index profile as shown in Figure 42 and where the ray has an axial incident angle θ_o and initial position $r \leqslant r_{min}$. This ray will be totally internally reflected when it reaches a radial position r, where n(r) satisfies the following equation

$$\sin^2 \theta_o = \frac{1}{n_o^2} (n_1^2 - n^2(r)) \qquad 138.$$

Denoting this value of θ_o as $\theta_c(r)$ and assuming that $n_o = 1$, then on substituting Equation (137) into (138) gives

$$\sin^2 \theta_c(r) = \frac{2n_1(\Delta n)(\Delta r)(r-r_{min}) - (\Delta n)^2(r-r_{min})^2}{(\Delta r)^2} \qquad 139.$$

Rearranging Equation (139) to make 'r' the subject gives a quadratic equation in r whose solution is

$$r = r_{min} + \frac{\Delta r}{\Delta n} \left(n_1 - \sqrt{n_1^2 - \sin^2 \theta_c(r)}\right) \qquad 140.$$

Figure 43 shows a graph of Equation (140) and illustrates the variation of the core radius r as θ_c is varied for $n_1 = 1.62$, $\Delta n = 0.1$, $r = 25\mu$, $\Delta r = 2\mu$. If a plane wave is incident on the waveguide entrance aperture within a radius r where $r \leqslant r_{min}$ with an axial angle of incidence θ_o, then neglecting any divergence of the plane wave due to diffraction at the entrance aperture, the radius of the illuminated core region will be given by Equation (140) with $\theta_o(r) = \theta_o$.

For example when $\sin \theta_o = 0$, $r = r_{min}$ and $\sin \theta_o = \sqrt{n_1^2 - n_2^2}$ (the critical angle of a stepped refractive index profile waveguide) $r = r_{max}$.

The waves which radiate from the core region within a radius $< r_{min}$ will have axial angles of radiation equal to the angle of incidence of the plane wave. However, waves which radiate from the core region $r_{min} < r < r_{max}$ will have axial angles of radiation $\theta_r(r)$, which are dependent upon the radial position of the point of radiation as well as the angle of incidence.

In Figure 42 the axial angle θ_1 of an arbitrary ray of light within the core region $r \lesssim r_{min}$ is given by

$$\sin \theta_1 = \frac{n_o}{n_1} \sin \theta_o \qquad 141.$$

and at a radius r, where $r_{min} < r \quad r_{max}$, the axial angle of the ray becomes $\theta_1(r)$ where

$$\cos \theta_1(r) = \frac{n_1}{n(r)} \cos \theta_1, \qquad 142.$$

The radiation angle $\theta_r(r)$ of this ray is given by

$$\sin \theta_r(r) = \frac{n(r)}{n_o} \sin \theta_1(r), \qquad 143.$$

where it is assumed that the radius, r, of the point of refraction lies in the range $r_{min} < r < r_{max}$. Substituting Equations (141) and (142) into Equation (143) and assuming that $n_o = 1$ gives

$$\sin \theta_r(r) = \sqrt{n(r)^2 - n_1^2 + \sin^2 \theta_o} \qquad 144.$$

Substituting for $n(r)$ using Equation (137) and neglecting the $(\frac{\Delta n}{\Delta r})^2$ term in the expansion of $n^2(r)$, Equation (144) becomes

$$\sin \theta r(r) = \sqrt{\sin^2\theta_o - 2 \frac{\Delta n}{\Delta r} n_1 (r-r_{min})} \qquad 145$$

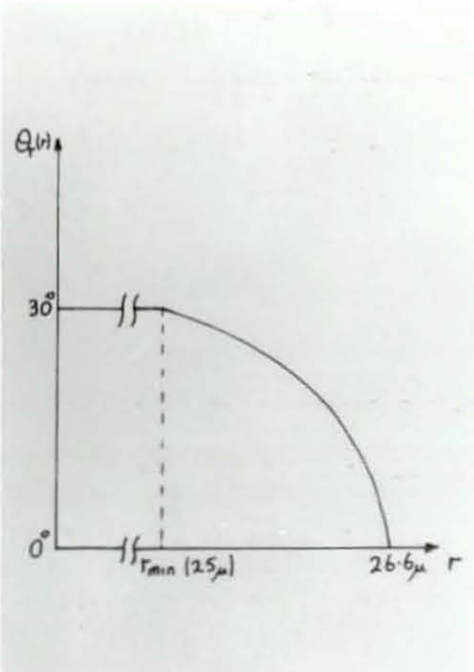


Figure 44. Graph of $\theta_r(r)$ against r.

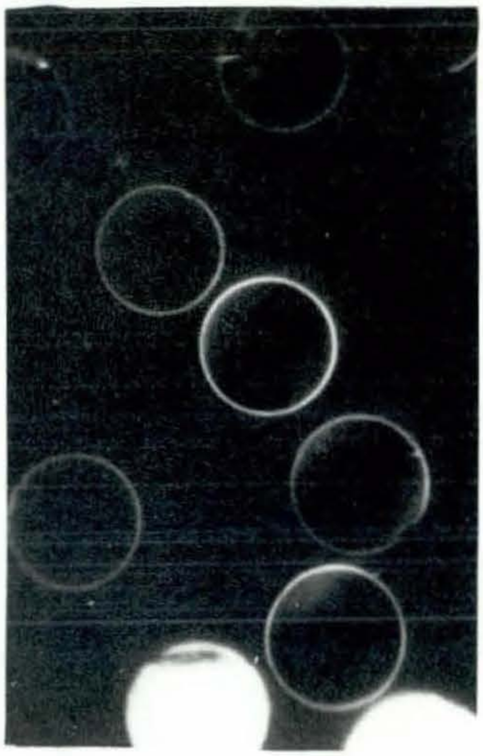


Figure 46. The ring effect.

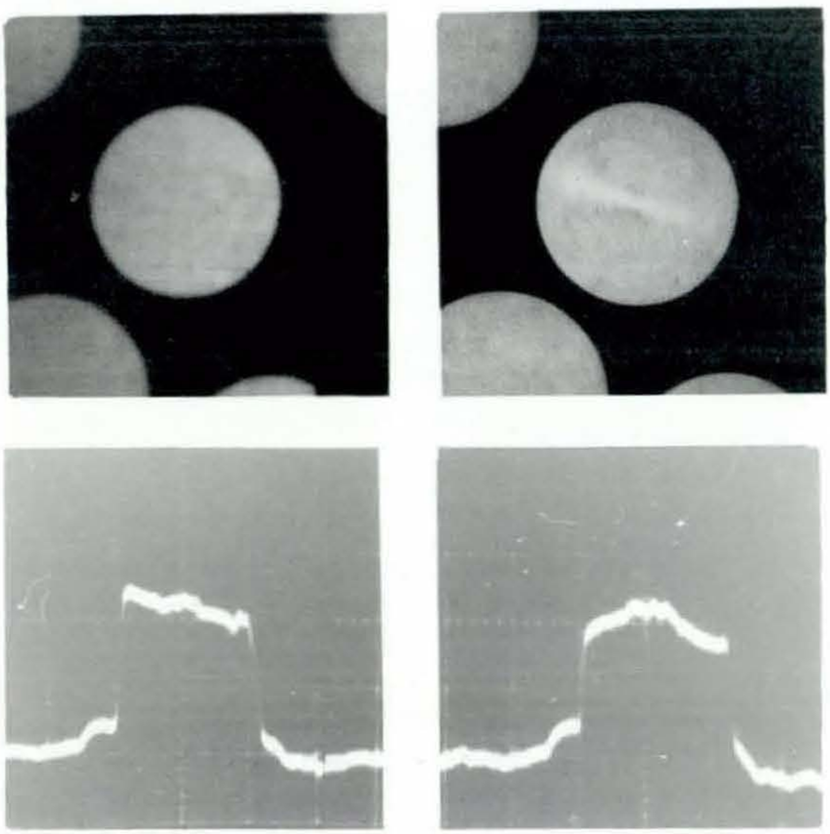


Figure 45. $\theta_o = 0^o$ $\theta_o = 30^o$

A graph of Equation (145) is shown in Figure 44 where $\theta_o = 30^o$, $n_1 = 1.62$, $\Delta r = 2\mu$, $\Delta n = 0.1$ and which illustrates that the angle $\theta_r(r)$ varies from θ_o to zero as the radius of the point of the ray increases from r_{min} to the value given by Equation (140) when $\theta_c(r) = 30^o$.

4.4.2 Experimental Results

A 30 cm length of fibre bundle was used to investigate the variations of the radius of the illuminated core region using the experimental arrangement of Figure 20 with a white light source. The diameter of the illuminated core region of selected fibres was measured using the TV-aided measuring system whilst illuminating the opposite end of the fibres at various angles. At each incident angle the oscilloscope trace of the radiated intensity was normalised to a selected height and the width of the trace at half this height taken as the width of the core region.

The resolution of the measuring system is limited to $\pm.5\mu$ and this prohibits attempts to measure the refractive index profiles which appear to have Δr of the order of 1.5μ or less. However, a measurement of the diameter of the illuminated core at incident angles $\theta_o = 0^o$ and $\theta_o = 30^o$ indicated that Δr was non zero confirming the existence of a finite width refractive index profile, and tests on numerous fibres yielded the average value for Δr of $(1 \pm 0.5\mu)$.

Photomicrographs of the radiation end of a fibre and photographs of the corresponding intensity traces are shown in Figure 45. The value of Δr for this fibre is $\simeq 1.2\mu$.

Equation (145) suggests that if such a fibre is illuminated at incident angle θ_o and the radiation field examined using a microscope objective with an acceptance angle $\theta_a < \theta_o$, a ring of light will be observed at a radius r for which $\theta_r(r) < \theta_a$ in Equation (145). This is demonstrated in Figure 46 where $\theta_o = 44^o$ and $\theta_a \approx 33^o$.

4.5 Conclusions

It is thought that the experiments to measure the entrance aperture diffraction effects are an original contribution to the experimental study of optical waveguides. The theoretical model of a normally incident plane wave of significant amplitude within the core region only, has been considered elsewhere[42]. That solution is given in terms of the efficiency of excitation of surface waveguide modes.

The attempts to measure the refractive index profile are also original, although another worker* has privately confirmed observation of the ring effect. He suggested that this effect was due to diffraction at the radiation end of the waveguide, since such a ring would be expected at sharp dielectric interfaces due to the formation of a cylindrical boundary wave[31, Sec.8-9, and Chap.11].

However, the variation in area of the illuminated core, as the angle of incidence of the illumination is varied, supports the refractive index profile theory. The observation of the intensity of a cylindrical boundary wave produced by a glass-air interface suggests that, if such a wave were present, it would be of considerably less intensity than the ring of light actually observed.

Since it was not possible to measure the refractive index profile, the ramp approximation has been used for simplicity. More appropriate methods of measuring the index profiles of optical waveguides have been reported[29] during the period of this research.

* W.J. Stewart, Plessey Co.Ltd., Private communication.

CHAPTER 5

5.1 Introduction.

This chapter contains an extensive study of primarily geometric ray theory phenomena which may be observed in a core cladded dielectric waveguide which is assumed to be perfectly straight and circular in cross section.

The meridional section representation of the waveguide is used to generate an equivalent cylinder in three dimensions. The concept of skew rays and skew ray paths is introduced and the meridional ray is shown to be a limiting skew ray. The behaviour of the skew rays at the various dielectric interfaces is examined and the behaviour of the waveguide as a thick lens is experimentally verified.

The propagation of skew rays in the waveguide core produces reflection number dependent caustics which are examined experimentally and by computer simulation. The analysis of the propagation of the skew rays leads to the derivation of the uniform radiation cone property of core cladded waveguides used for the alignment procedure described in Chapter 2. The 'black hole' effect is also derived and the effects of sloping end terminations on the radiation cone are examined.

Although the geometric ray theory predicts many of the observed phenomena the microscopic details are explained by reference to the rigorous electromagnetic theory and this has been included where necessary.

5.2 Generalized Geometric Representation of Core Cladded Dielectric Waveguides.

The parallel mirror representation of a meridional cross section of a core cladded dielectric waveguide developed in Chapter 4 will form a cylindrical mirror surface if it is considered in three dimensions.

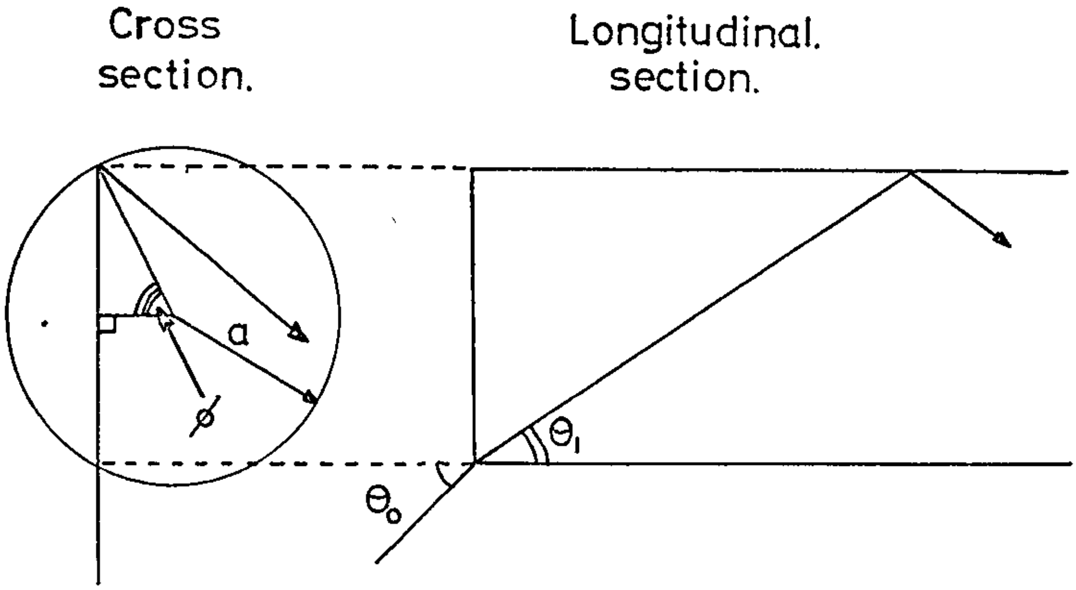


Figure 47. A ray in the longitudinal and cross sections of an internally reflecting cylinder.

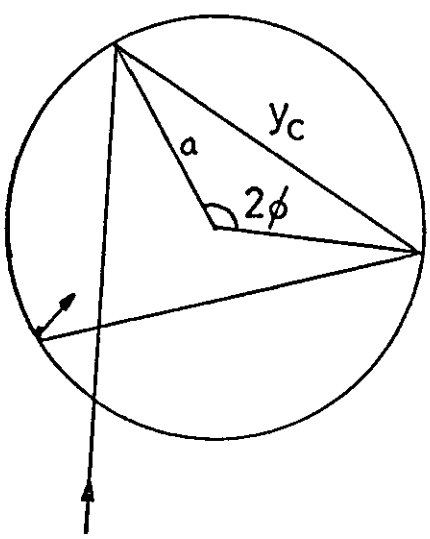


Figure 48. The path of a ray in the cross section of a cylinder.

However, the condition for total internal reflection of rays within the core given by Equation (119) only applies to meridional rays in the cylinder.

A general ray within the core is characterised by two angles θ, ϕ, as shown in Figure 47. The longitudinal section of the core containing the ray defines angle θ and the projection of the ray in the cross section defines angle ϕ. A general ray which enters the core through the core entrance aperture at incident angle θ_o has to satisfy the condition

$$\sin \theta_o \sin \phi \leqslant \frac{1}{n_o} (n_1^2 - n_2^2)^{\frac{1}{2}} \qquad 146.$$

in order to be totally reflected from the $n_1 - n_2$ interface. Clearly Equation (146) reduces to the meridional Equation (119) when $\phi = \frac{\pi}{2}$. Rays which satisfy Equation (146) but which have $\theta_o > \theta_{oc}$ where θ_{oc} is the meridional critical angle, are called leaky rays because they are slightly attenuated on reflection[43]. In the short lengths of waveguide used in this study the effects of this attenuation are negligible.

The behaviour of a general ray at the cylindrical dielectric interfaces is similar to that described in Chapter 4 except that the angle of incidence is formed from a combination of θ and ϕ as in the left hand side of Equation (146). Certain rays are investigated in the next section and it is sufficient here to note that only rays which enter the waveguide through the core entrance aperture and satisfy Equation (146) are propagated unattenuated in the core by the total reflection process.

The characteristics of a general ray propagating in the core are now considered. Since the angle of incidence is preserved in magnitude at each reflection, the values of θ and ϕ also remain constant in magnitude.

The optical path length of a general ray in a length ℓ of waveguide depends only upon θ and is given by Equation (123). The angle ϕ defines the path taken by a ray in the cross section of the core as shown in Figure 48.

The projection of the path of a ray in the cross section between adjacent points of reflection forms a chord of a circle of length y_c where

$$y_c = 2a \sin \phi \qquad 147.$$

and which subtends an angle at the centre of the circle of 2ϕ. If the number of reflections $m(\phi_1 \ell)$ of ray $(\theta_1 \phi_1)$ in a length, ℓ, of the core is given by

$$m(\phi_1 \ell) = \frac{\ell \tan \theta_1}{2a \sin \phi_1} \qquad 148.$$

then the angle ϕ_T subtended at the centre of the circle after $\bar{m}$ reflections where $\bar{m}$ is the integer value of $m(\phi_1 \ell)$ is given by

$$\phi_T = 2\bar{m}\ \phi_1^{o} \qquad 149.$$

If

$$\phi_T = p360^{o} \qquad 150.$$

where p is an integer, then after $\bar{m}$ reflections and p revolutions, the ray path will form a closed figure within the cross section. The closed figures for $p = 1$, $\bar{m} = 2, 3, 4$ and $p = 2$ $\bar{m} = 5$ are shown in

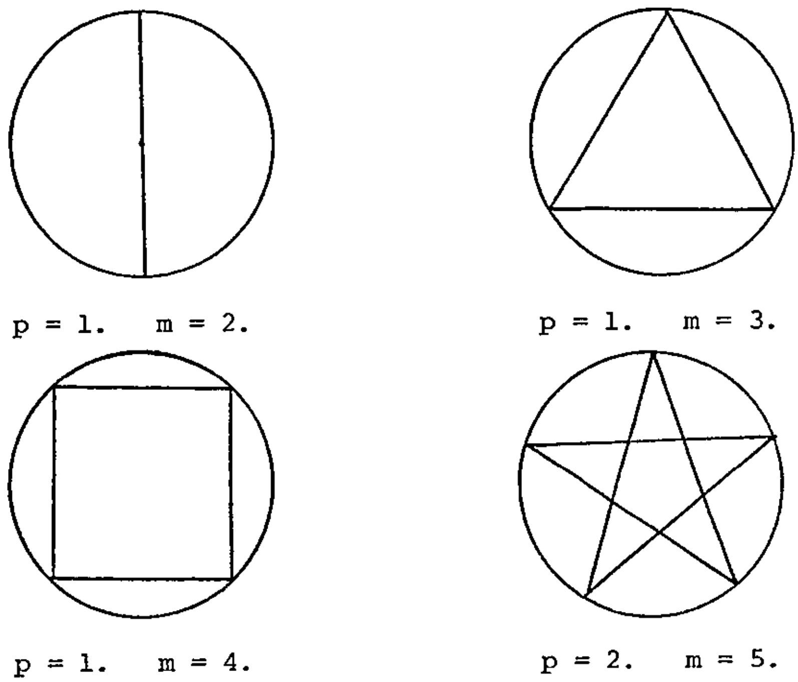


Figure 49. Examples of closed skew ray figures.

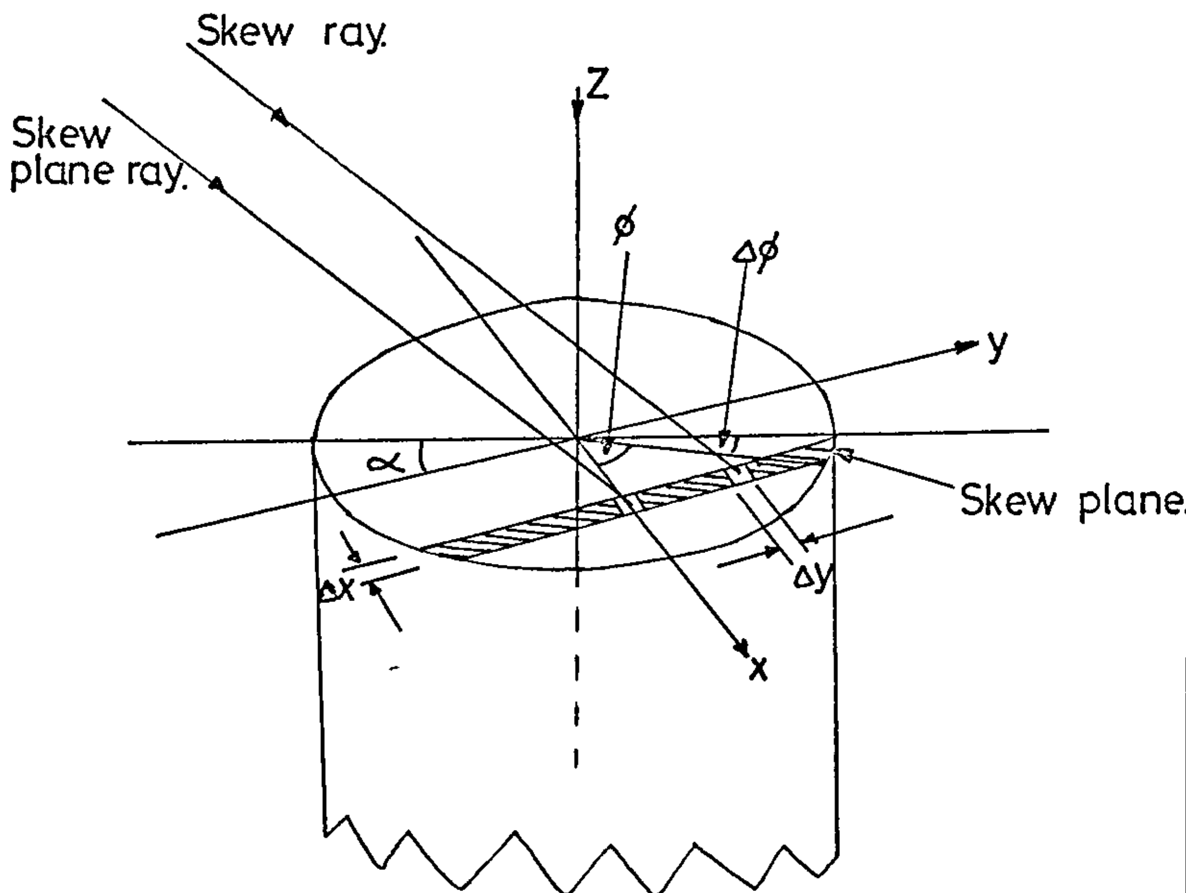


Figure 50. The excitation of a skew plane ray.

Figure 49 and may be compared with the closed figures of Figure 6 which were obtained from the initial experiments described in Chapter 1.

The general experimental arrangement for exciting the patterns described in Chapter 1 is shown in Figure 50. It is assumed that the source illuminates the core aperture with plane waves at incident angle θ_o and a single azimuth angle, α, which for a straight circular waveguide is quite arbitrary. The core aperture is divided into elements of area $\Delta x\ \Delta y$ and the propagation of the light incident upon each element will be represented by a ray whose starting position is the centre of the element and has incident angles θ_o, ϕ where

$$\cos\phi = \frac{x}{a} \qquad 151.$$

If each ray is called a skew ray, then all skew rays with the same value of ϕ will follow a similar path in the core and may be considered collectively as a skew plane. A skew plane has an area in the core aperture of $\Delta x.\ 2a \sin\phi$ and, for uniform incident radiation and flux density D, will represent an intensity $I(\phi)$ where

$$I(\phi) = D.\ \Delta x.2a \sin\phi \qquad 152.$$

The skew ray at the centre (y = 0) of each skew plane will be called a skew plane ray, and the propagation of the skew plane in the waveguide may be represented by the propagation of its skew plane ray, since the skew plane will maintain a fixed spatial relationship with its associated skew plane ray. The two skew plane rays due to $\pm x$ in Equation (151) are similar but they will propagate down the core in opposite directions. The anticlockwise propagation direction ($\cos\phi = \frac{+x}{a}$) is assumed to be positive.

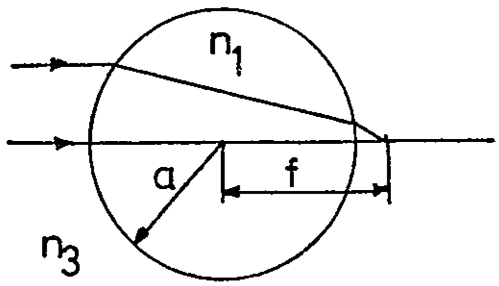


Figure 51a. A cross section of a cylindrical lens.

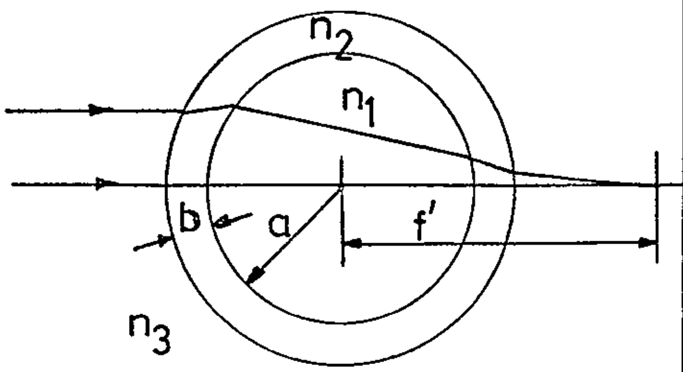


Figure 51b. A cross section of cladded dielectric cylindrical lens.

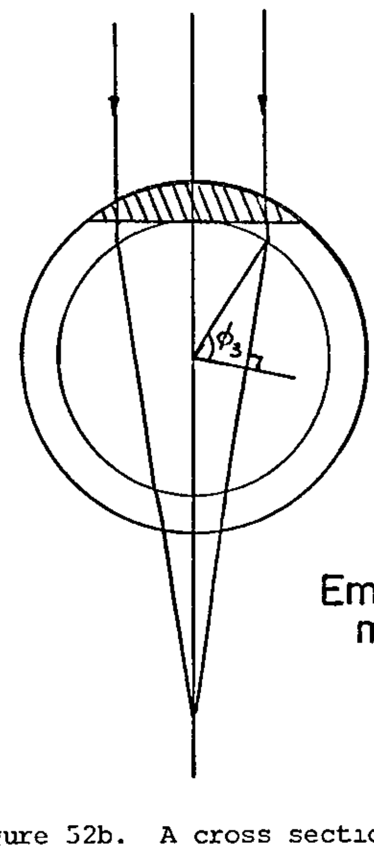


Figure 52b. A cross section of an embedded dielectric cylindrical lens.

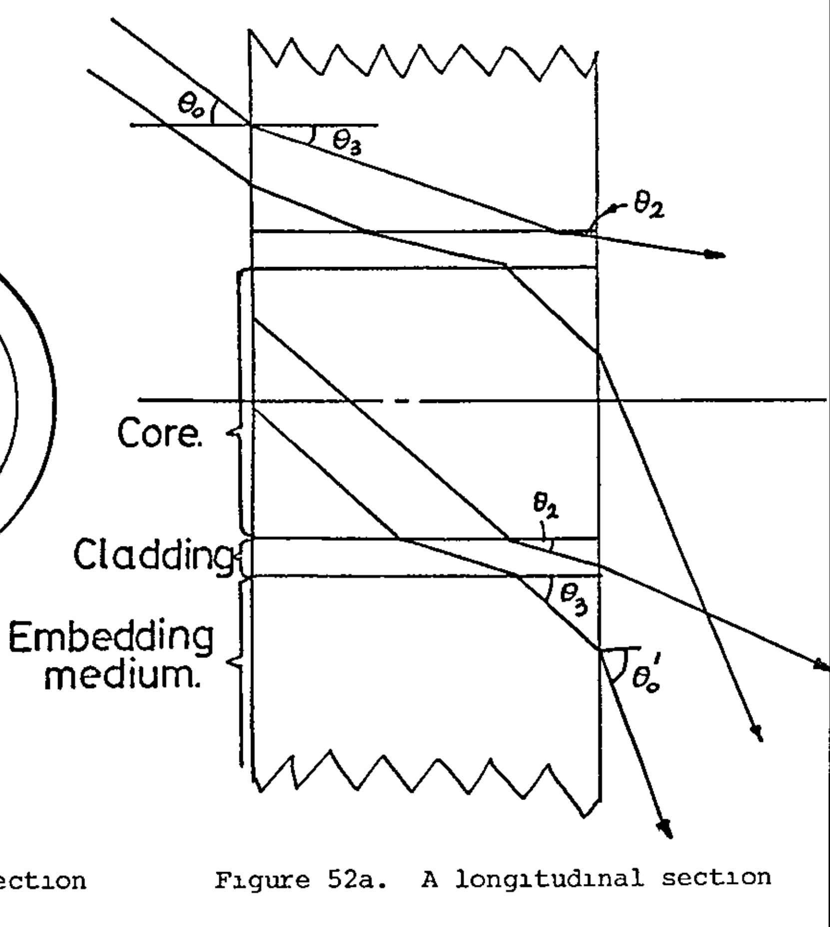


Figure 52a. A longitudinal section of an embedded, dielectric cylindrical lens.

The skew planes which form the closed figures described previously will be called stationary skew planes and the associated skew plane ray a stationary skew plane ray. Before proceeding with the analysis of the propagation of skew plane rays in the waveguide core, the behaviour of light rays in a short length of embedded waveguide is considered. The waveguide is illuminated as in Figure 49, but attention is confined to the light which is incident upon the cladding wall.

5.3 The Dielectric Waveguide "Thick Lens"

It is well known that a cylinder of glass behaves as a thick lens [44,Ch.2] and it is easy to show that within the paraxial approximation a dielectric cylinder refractive index n_1, radius a, in a medium of refractive index n_3 has a focal length f, as shown in Figure 51a where

$$f = \frac{n_1 a}{2(n_1 - n_3)} \tag{153.}$$

The dielectric waveguide has a cladding dielectric, refractive index n_2, thickness b, and it may be shown that the focal length of this device, shown in cross section in Figure 51b, is given by

$$f' = \frac{n_1 \, n_2 \, a}{2n_3(n_1-n_2) - n_1(\frac{n_3}{n_2} - 1)} \tag{154.}$$

where n_3 is the refractive index of the surrounding media, and $n_1 > n_3 > n_2$.

Figure 52a shows a longitudinal section and Figure 52b a cross section of a short length of waveguide embedded in a material of refractive index n_3 and illuminated by plane waves at the incident angle θ_o shown. Rays of light incident on the $n_o - n_3$ interface are refracted into the embedding material at incident angle θ_3.

where

$$\sin \theta_3 = \frac{n_o}{n_3} \sin \theta_o \qquad 155.$$

Rays of light incident upon the cladding wall will refract into the cladding if

$$\sin \theta_3 \sin \phi_3 > \frac{1}{n_3} (n_3^2 - n_2^2)^{\frac{1}{2}} \qquad 156.$$

since they will exceed the total internal reflection condition at this interface.

If θ_3 is such that Equation (156) is just satisfied, the permissible range of ϕ_3 will be small. The refracted rays in the cladding will have $\theta_2 \simeq 0$ and will radiate from the portion of the cladding end face shown shaded in Figure 52b. This effect is demonstrated experimentally in Figure 53a which is a photomicrograph of the end of a 570 micron length of core cladded waveguide illuminated with white light at $\theta_o = 26.5^o$.

Increasing θ_o increases the range of ϕ_3 but more important, θ_2 will now be greater than zero and certain rays will be refracted into the core at angles θ_1, ϕ_1. If it is assumed that rays within the paraxial region of the equivalent cylindrical lens have values of ϕ_1 for which $\sin \phi_1 \simeq 1$ by comparison with the small values of θ_1, then the refraction of the ϕ and θ components of the angle of incidence at each interface may be considered separately. This amounts to applying the thick lens equations to the cross sectional components of the rays and the meridional approximation to the longitudinal components.

Consider the ray components in the longitudinal section, Figure 52a, which have been refracted into the core. Rays near the end of the waveguide will radiate through the $n_1 - n_o$ interface and those which

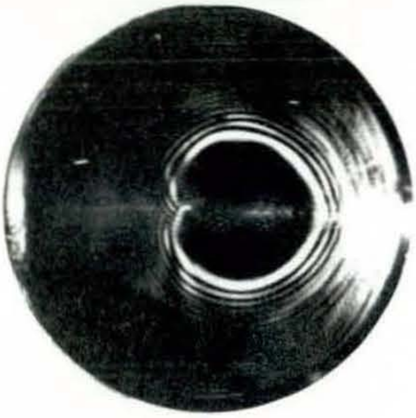


Figure 53a.

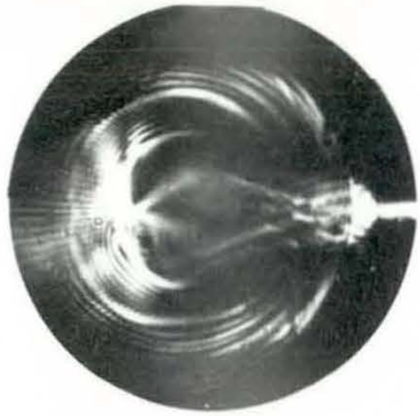


Figure 53b.

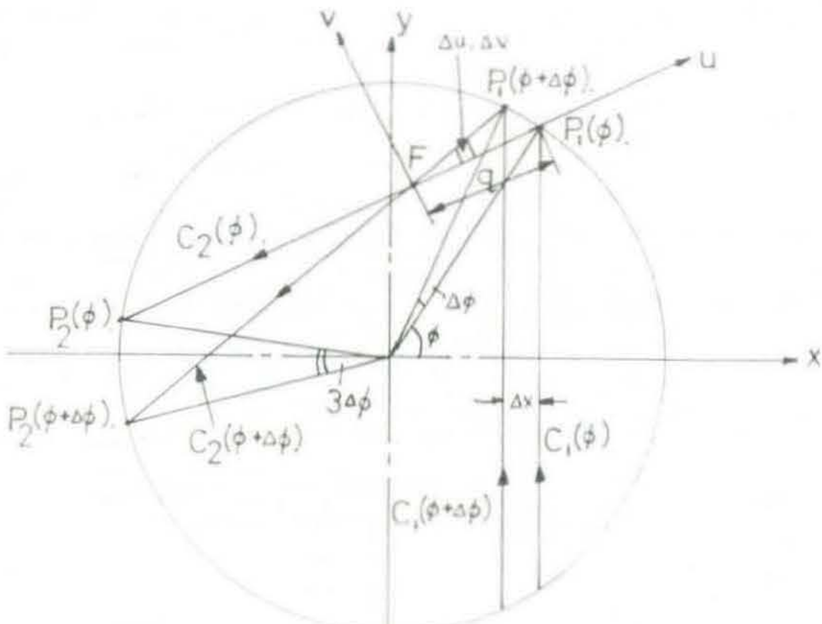


Figure 54. Adjacent skew planes showing the position of their intersection after one reflection.

meet the core cladding boundary prior to radiating will be refracted into the cladding at angle θ_2.

Rays in the cladding near the end of the waveguide will radiate through the $n_2 - n_o$ interface and rays meeting the outside cladding wall will refract into the embedding medium at angle θ_3. Finally the rays in the embedding medium will radiate through the $n_3 - n_o$ interface at an angle θ_o' where within the paraxial approximation $\theta_o' = \theta_o$.

The cross sectional components of all these rays will have been subjected to the focussing properties of the waveguide. The focussing of the incident plane wave will be evident in the radiation field, although the differing angles of radiation from the various regions of the end of the waveguide will present different aspects of the focussed beam. This is demonstrated experimentally in Figure 53b, which is a photomicrograph of the same sample as in Figure 53a but with $\theta_o = 30.5^o$. The different aspects of the focussed beam are evident in the cladding and embedding medium regions. The beam is viewed normally in the cladding region ($\theta_2 \simeq 0$) and obliquely ($\theta_3 \simeq 26^o$) in the embedding region. The radiation from the core contains components of the focussed beam, which may be identified with the help of the ray paths shown in Figures 52a-b. The radiation field from the core due to light incident in the core entrance aperture is also very much in evidence, in the form of caustics, and these are derived in the next section.

5.4 Propagation of Skew Planes and Skew Plane Rays

Consider the cross section of a skew plane bounded by chords $c_1(\phi)$, $c_1(\phi + \Delta\phi)$ shown in Figure 54 where the angle subtended at the centre of the circle, radius a, by the portion of the circumference between the points of reflection $p_1(\phi)$, $p_1(\phi + \Delta\phi)$ is $\Delta\phi$ where

$$\Delta\phi \simeq \frac{\Delta x}{a} \qquad 157.$$

The intensity of a single element of area $\Delta x\Delta y$ of this skew plane is denoted by $I(\Delta x\Delta y)$. After a single reflection the chords $c_1(\phi)$, $c_1(\phi + \Delta\phi)$ form chords $c_2(\phi)$, $c_2(\phi + \Delta\phi)$ and the angles subtended at the centre of the circle by the chords $c_2(\phi)$, $c_2(\phi + \Delta\phi)$ are given by Equation (149) as 2ϕ and $2(\phi + \Delta\phi)$ respectively. Thus the angle subtended at the centre of the circle by the portion of circumference between the new points of reflection $P_2(\phi)$, $P_2(\phi + \Delta\phi)$ is $3\Delta\phi$. As shown in Figure 54 the chords $c_2(\phi)$, $c_2(\phi + \Delta\phi)$ intersect at point F which is a distance, q, along chord $c_2(\phi)$ from point $P_1(\phi)$. Assuming that the triangles $(P_1(\phi), FP_1(\phi + \Delta\phi))$, $(P_2(\phi), FP_2(\phi + \Delta\phi))$ are similar then

$$q = \frac{a}{2} \sin \phi \qquad 158.$$

Consider an element of area between chords $c_2(\phi)$, $c_2(\phi + \Delta\phi)$ of dimensions $\Delta u\Delta v$ a distance, u, from F, where u,v are rectangular co-ordinates with their origin at F and u coinciding with $c_2(\phi)$.

The element Δv is given by inspection of Figure 54 as

$$\Delta v = \frac{|u|\ 2\ \Delta x}{a \sin \phi} \qquad 159.$$

Invoking the conservation of intensity theorem expressed by Equation (118), the intensity in the element $\Delta u\Delta v$ is given by

$$I(\Delta u\Delta v) = I(\Delta x\Delta y)\ .\ \frac{\Delta x\Delta y}{\Delta u\Delta v}$$

If $\Delta u = \Delta y$ and using Equation (159)

$$I(\Delta u\Delta v) = I(\Delta x\Delta y)\ .\ \frac{a \sin \phi}{2\ |u|} \qquad 160.$$

Equation (160) suggests a point of infinite intensity at $u = 0$, which is called a focus. This apparent failure of the geometric ray

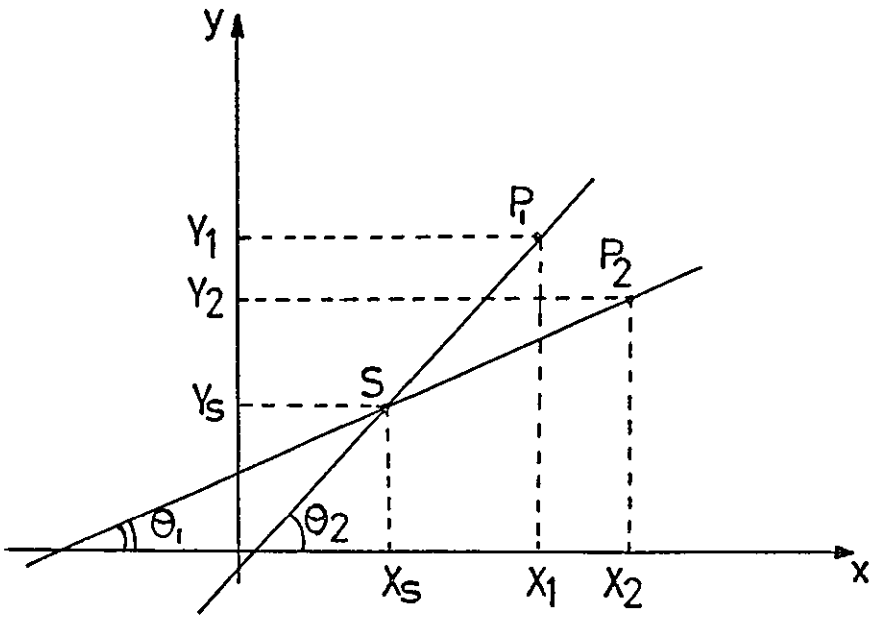


Figure 55. The intersection of two straight lines.

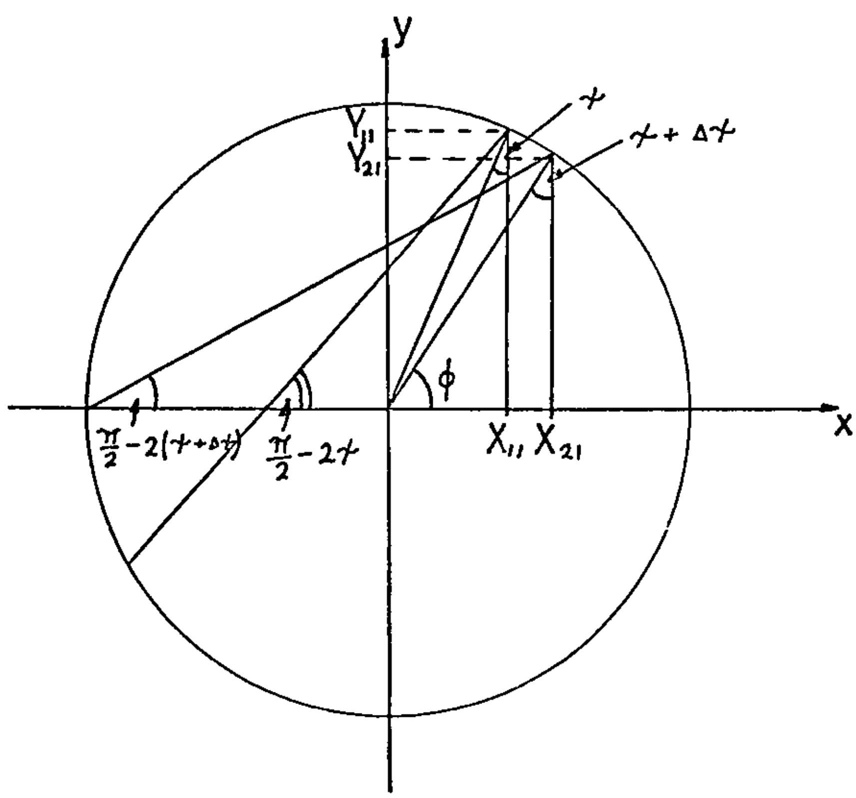


Figure 56. The intersection of skew planes.

theory is caused by attempting to compress the electromagnetic waves into zero volume, where the form of the space dependent component of the wave equation solution (Equation (105)) used for the derivation of the geometrical ray theory is inapplicable. The assumptions used there that a(r) and L(r) vary slowly with respect to the wavelength of the light are incorrect at a focus and the correct intensity at, or near, u = 0 is obtained only by using diffraction theory.

The diffraction pattern at a caustic, which is the locus of foci, is analysed in the next section. The parameters of a caustic required for that analysis are the radius of curvature of the caustic and the position of the caustic with respect to the reflecting surface. The cross sectional components of these two parameters are now obtained for the cylindrical reflecting surface, as a function of the number m of integer reflections made by the skew plane rays, initially illuminated by a plane wave incident at a single azimuth angle.

Consider first the two straight lines shown in Figure 55, where points P_1, P_2 have co-ordinates X_1Y_1, X_2Y_2 respectively and the slopes of the lines are given by T_1, T_2,
where

$$T_1 = \tan\theta_1,\ T_2 = \tan\theta_2$$

The equations defining the two straight lines may be expressed as

$$y = \begin{matrix} X_1 - (Y_1 - y)/T_1 \\ X_2 - (Y_2 - y)\ T_2 \end{matrix} \quad 161a. \qquad y = \begin{matrix} Y_1 - (X_1 - x)\ T_1 \\ Y_2 - (X_2 - x)\ T_2 \end{matrix} \quad 161b.$$

The co-ordinates x_s, y_s of the point S where the two lines intersect may be found by equating the y and x components of the lines.

From Equation (161b)

$$x_s = \frac{(Y_2 - X_2T_2) - (Y_1 - X_1T_1)}{(T_1 - T_2)} \tag{162a.}$$

From Equation (161a)

$$y_s = \frac{T_1(Y_2 - X_2T_2) - T_2(Y_1 - X_1T_1)}{(T_1 - T_2)} \tag{162b.}$$

Let

$$P = (T_1 - T_2) \tag{163a.}$$

$$K_1 = (Y_1 - X_1T_1) \tag{163b.}$$

$$K_2 = (Y_2 - X_2T_2) \tag{163c.}$$

and substitute Equations (163) into Equations (162)

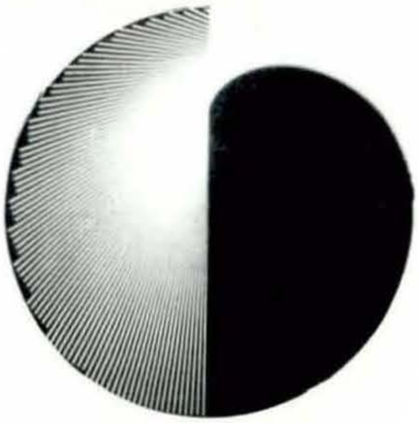
$$x_s = \frac{K_2 - K_1}{P} \tag{164a.}$$

$$y_s = \frac{T_1K_2 - T_2K_1}{P} \tag{164b.}$$

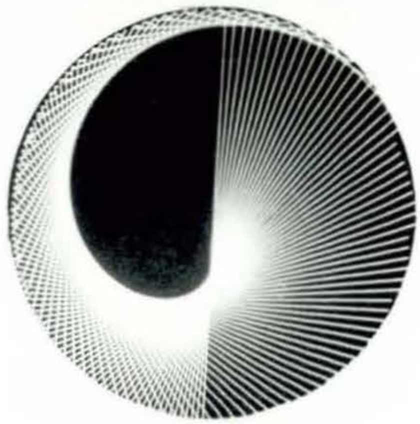
In Figure 56 the initial position of the skew planes is parallel to the y axis. The chords defining the skew plane ϕ, $\phi + \Delta\phi$ satisfy equations of the form of Equations (161), in which the subscripts $_{1,2}$ will denote $\phi + \Delta\phi$, ϕ respectively. It is convenient to define X, Y, T for each chord in terms of the angle, χ, where

$$\chi = \frac{\pi}{2} - (\phi + \Delta\phi) \tag{165a.}$$

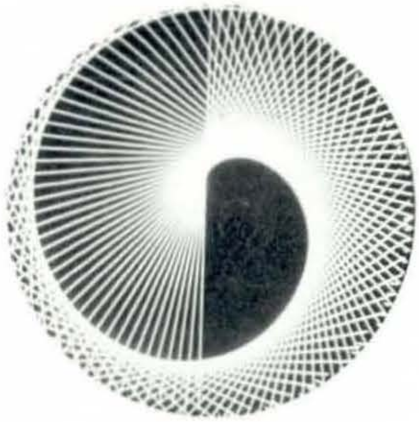
$$(\chi + \Delta\chi) = \frac{\pi}{2} - \phi \tag{165b.}$$



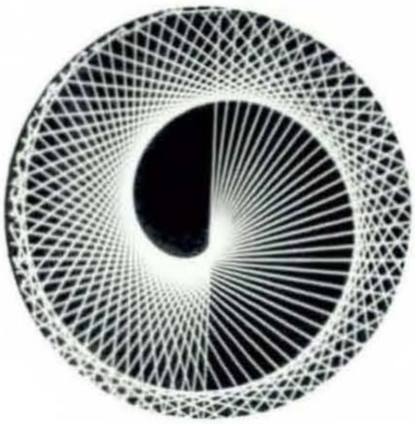
m = 1



m = 2



m = 3



m = 4

Figure 57. Photographs of caustics produced by computer simulation of rays reflected by circular reflectors.

The values of X, Y, T for the two chords after the m^{th} reflection are

$$\left.\begin{aligned} X_{1m} &= a \sin(2m-1)\chi \\ Y_{1m} &= a \cos(2m-1)\chi \end{aligned}\right\} \quad 166a. \qquad \left.\begin{aligned} X_{2m} &= a \sin(2m-1)(\chi+\Delta\chi) \\ Y_{2m} &= a \cos(2m-1)(\chi+\Delta\chi) \end{aligned}\right\} \quad 166b.$$

$$T_{1m} = \frac{\cos 2m\chi}{\sin 2m\chi} \quad 167a \qquad T_{2m} = \frac{\cos 2m(\chi+\Delta\chi)}{\sin 2m(\chi+\Delta\chi)} \quad 167b.$$

The term $\sin(2m-1)\frac{\pi}{2}$ has been dropped from the Equations (166) and is dropped from all subsequent equations except for Equation (169), since it only represents a change of sign for each increment in m.

Substitution of Equations (166), (167) into (163) gives

$$P_m = \frac{\sin 2m.\Delta\chi}{\sin 2m\chi \sin 2m(\chi + \Delta\chi)} \quad 168a.$$

$$K_{1m} = \frac{a \sin \chi}{\sin 2m\chi} \quad 168b.$$

$$K_{2m} = \frac{a \sin(\chi + \Delta\chi)}{\sin 2m(\chi + \Delta\chi)} \quad 168c.$$

Substitution of Equation (168) into (164) and letting $\Delta\chi \rightarrow 0$ gives the co-ordinates x_s, y_s of the function f(s) describing the position of the caustic produced by the m^{th} reflection of the skew planes.

$$x_s = a(\frac{1}{2m}(\sin 2m\chi \cos \chi) - \cos 2m\chi \sin \chi) \sin(2m-1)\frac{\pi}{2} \quad 169a.$$

$$y_s = a(\frac{1}{2m}(\cos 2m\chi \cos \chi) + \sin 2m\chi \sin \chi) \sin(2m-1)\frac{\pi}{2} \quad 169b.$$

Figure 57 shows photographs of computer simulations of rays reflected by a circular reflector for m = 1 to m = 4, and where the locus of the intersections of the rays forms f(s).

Expressing f(s) in $R/\underline{\eta}$ form using Equation (169)

$$R = \sqrt{x_s^2 + y_s^2} = a\sqrt{\sin^2\chi + \frac{1}{4m^2}\cos^2\chi} \qquad 170a.$$

$$\tan\eta = \frac{y_s}{x_s} = \frac{1 + 2m \tan 2m\chi \tan\chi}{\tan 2m\chi - 2m \tan\chi} \qquad 170b.$$

and from Figure 57 it is seen that f(s) forms spirals which originate at

$$R = \frac{a}{2m} \qquad \eta = (2m+1)\frac{\pi}{2}\ ,\ (1-2m)\frac{\pi}{2} \qquad (\chi = 0) \qquad 171.$$

and terminate at

$$R = a \qquad \eta = 0\ ,\ \pi \qquad (\chi = \frac{\pi}{2}) \qquad 172.$$

where η is given for $+\phi$ and $-\phi$ respectively.

The first half revolution of each spiral forms a cardioid type figure which is inverted and reduced in size for each increment in m. The contraction of the spiral origin ($R = \frac{a}{2m}, \chi = 0$) along the y axis is given by

$$D_y = \frac{m}{m+1} \qquad 173.$$

The length R, of the spiral at the first crossing of the x axis is found by setting Equation (169b) equal to zero and substituting the corresponding value of χ in Equation (170a).

Setting Equation (169b) equal to zero results in the Equation

$$\frac{\cos(2m+1)\chi}{\cos(2m-1)\chi} = \frac{(2m+1)}{(2m-1)} \qquad 174.$$

The solution for χ when $m = 2$ in Equation (174) is $\chi \simeq 42^{\circ}$ and substitution of $m = 2$, $\chi \simeq 42^{\circ}$ in Equation (170a) gives a value for R which to a good approximation is given by

$$R = a \sin \chi \qquad 175.$$

Since the term $\frac{1}{4m^2} \cos^2\chi$ in Equation (170a) rapidly diminishes with increasing m, Equation (175) rapidly approaches the exact value for R for all m and χ. The corresponding approximation for Equation (170b) gives

$$\mathrm{TAN}\ \eta = \frac{1}{\mathrm{TAN}\ 2m\chi} \qquad 176.$$

and the values of χ corresponding to the first crossing of the x axis by the spiral are given by

$$\chi = \frac{\pi}{2m} \qquad 177.$$

Substituting Equations (177) into (175) and assuming that

$\sin \frac{\pi}{2m} \simeq \frac{\pi}{2m}$

gives

$$R = \frac{a\pi}{2m}, \qquad 178.$$

and the contraction along the x axis is given by

$$D_x = \frac{m}{m+1} \qquad 179.$$

in agreement with D_y.

Consider now the values of η and R for the two chords $c_1(\phi)$, $c_1(\phi + \Delta\phi)$ of Figure 54 after m reflections.

Using Equations (179), (170) with (165a),

$$\eta(\phi) = 2m\phi \quad (a) \qquad R(\phi) = a \cos \phi \quad (b) \qquad 180.$$

$$\eta(\phi+\Delta\phi) = 2m\phi+2m\Delta\phi \quad (a) \qquad R(\phi+\Delta\phi) = a(\cos\phi \cos\Delta\phi - \sin\phi \sin \Delta\phi) \ (b) \qquad 181.$$

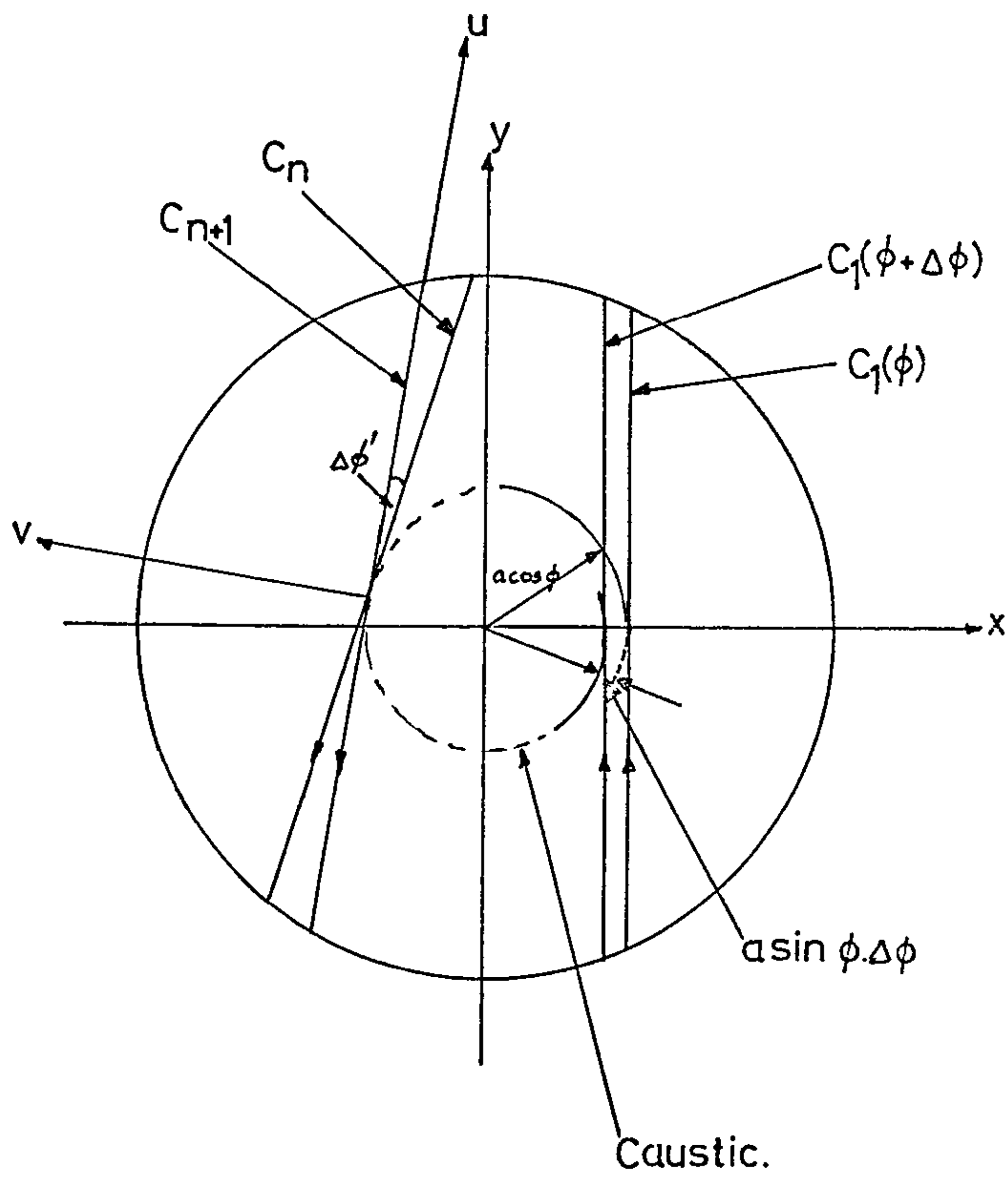


Figure 58. The position of the caustic formed when $m\phi = \pi$.

When $2m\Delta\phi = 2\pi$ the caustic formed by these two chords will be as shown in Figure 58 and the light initially between the chords will appear within the annulus, inner radius a cos ϕ (neglecting the a sin ϕ sin $\Delta\phi$ term of Equation (181b)) outer radius a. This annulus may be considered as being composed of 2m identical sections bounded by chords c_o, c_1, ... c_n (see Figure 58) where adjacent chords are inclined at an angle

$$\Delta'\phi = \frac{\Delta\phi}{\sin\phi} \qquad 182.$$

to each other and intersect at their midpoints where they are tangent to the caustic.

Consider the section bounded by the chords c_n, c_{n+1} and consider an element $\Delta u\Delta v$ of this section a distance u from the midpoint of c_n where u,v are rectangular co-ordinates with their origin at the midpoint of c_n and u coinciding with c_{n+1}.

Assuming $\Delta\phi = \frac{\Delta x}{a}$ as before then

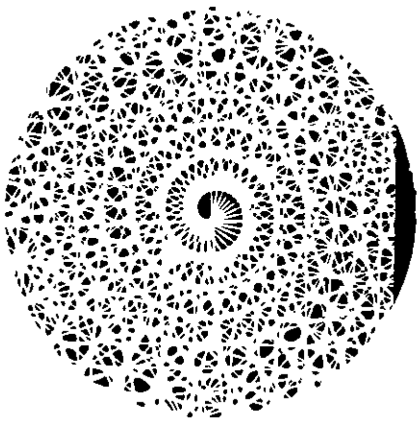
$$\Delta v = \frac{|u|\Delta x}{a \sin\phi} \qquad 183.$$

Since there are 2m such elements then, by the conservation of intensity theorem, the intensity of the element $\Delta u\Delta v$ is given by

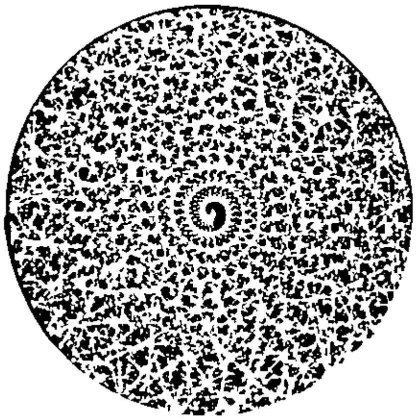
$$I(\Delta u\Delta v) = \frac{I(\Delta x\Delta y)}{2m} \cdot \frac{a \sin\phi}{|u|} \qquad 184.$$

where $I(\Delta x\Delta y)$ is the intensity of an element of the skew plane ray as before (Figure 54).

If the original skew plane ray of width Δx is considered as 2m skew planes each of width $\Delta x/2m$, each of these 2m skew planes



m = 20



m = 40

Figure 59. Photographs of computer simulations of rays reflected by circular reflectors.

corresponds to a single section of the annulus. The light within the original skew plane ray which is in a single azimuthal direction, parallel to the y axis, is now uniformly distributed over $2m\Delta\phi$ azimuth angles. When $2m\Delta\phi = 2\pi$ the light of the original skew plane ray will radiate as a complete hollow cone of light.

Equations (180a,b) represent a spiral which 'winds up' as m is increased as shown in Figure 59 where the spirals for $+\phi$ only are illustrated for $m = 20, 40$. The increment in R for an increment in η of 2π is given by

$$\Delta R = a \sin \phi \, \Delta\phi \tag{185.}$$

and is the width of the skew plane ray whose light is uniformly distributed over 2π azimuth angles. As shown in Figure 59 ΔR reduces as m is increased, and the radiation pattern due to a plane wave incident on the waveguide core entrance aperture rapidly approaches a uniform hollow cone of light.

If a single skew plane ray, bounded by chords $c(\phi')$, $c(\phi'+\Delta\phi)$ is allowed to propagate in the core, then when the width of the skew plane Δx equals the increment ΔR given by Equation (185), the light of the skew plane ray will completely illuminate the annulus, inner radius $a \cos \phi'$, outer radius a. The intensity distribution within the annulus is dominated by diffraction effects due to the caustic as discussed in the next section, but within the geometric approximation, the centre of the core within a circle radius $a \cos \phi'$ will remain in shadow, (i.e. black). Furthermore skew plane rays with $\phi \leqslant \phi'$ will make no contribution to this black circle within the core.

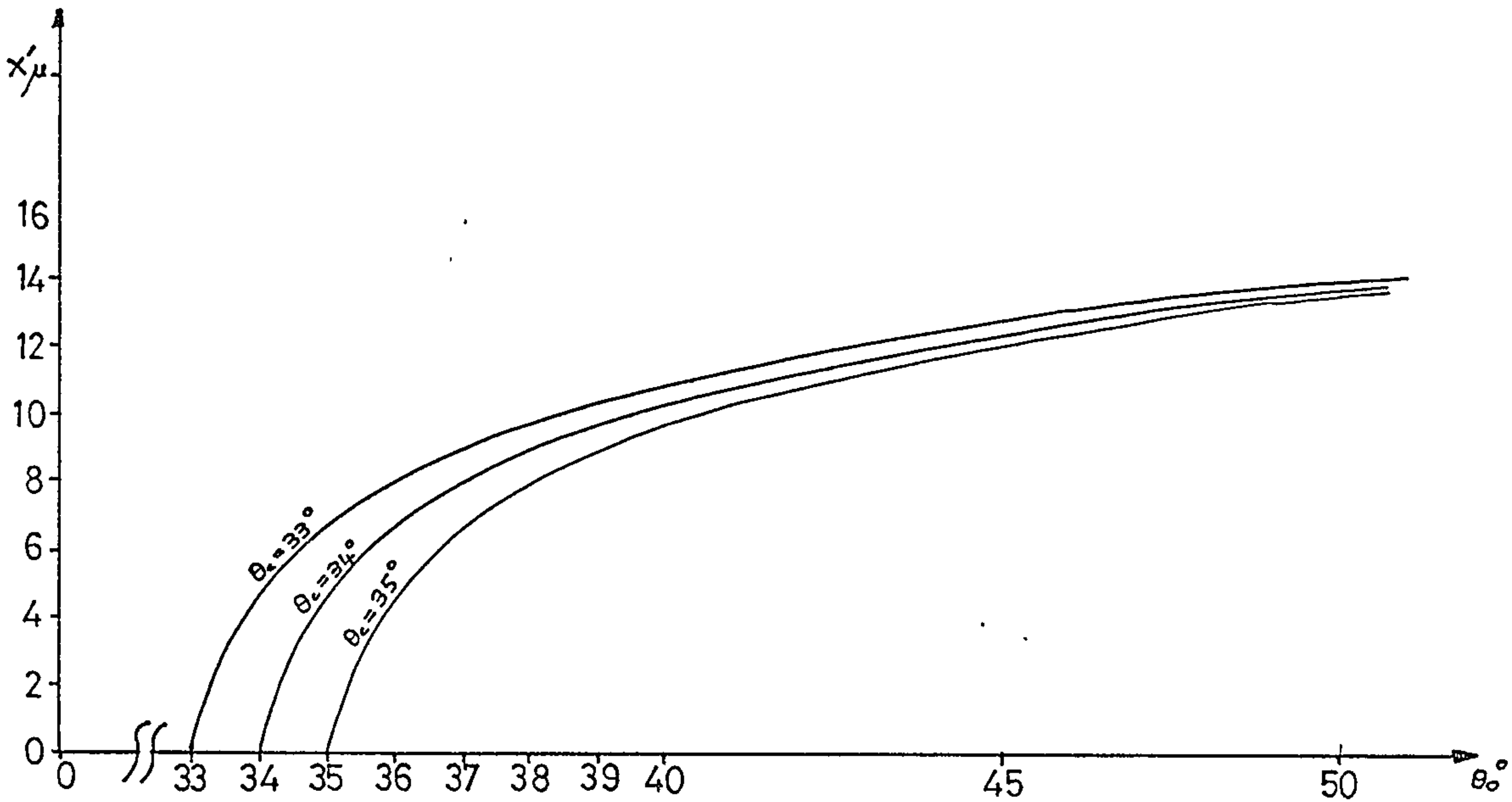


Figure 60. A Graph of Equation 188.

Equation (146) suggests a simple experimental arrangement to allow only skew plane rays satisfying the condition $\phi \leq \phi'$ to propagate. If the angle of incidence θ_o is greater than the meridional critical angle θ_c, the values of ϕ which satisfy Equation (146) are given by

$$\cos\phi \geq \sqrt{1 - \frac{\sin^2\theta_c}{\sin^2\theta_o}} \tag{186.}$$

This corresponds to the condition $\phi \leq \phi'$ if

$$\cos\phi' = \sqrt{1 - \frac{\sin^2\theta_c}{\sin^2\theta_o}} \tag{187.}$$

Denoting the radius of the black circle by x'

$$x' = a\sqrt{1 - \frac{\sin^2\theta_c}{\sin^2\theta_o}} \tag{188.}$$

and rearranging Equation (188) to make θ_o the subject

$$\sin\theta_o = \sin\theta_c / \sqrt{1 - \left(\frac{x'}{a}\right)^2} \tag{189.}$$

which on substitution of $x' = \frac{a}{2}$ gives Equation 1. A graph of Equation (188) is given in Figure 60, for $\theta_c = 33^o$, 34^o, 35^o and from which it may be seen that a 2^o variation in θ_o when $\theta_o \simeq 40^o$, $\theta_c = 34^o$ results in a variation of x' of 1μ when a = 20μ. This is the basis of the confidence limits discussed in Chapter 2 for the determination of θ_o from a measurement of x'.

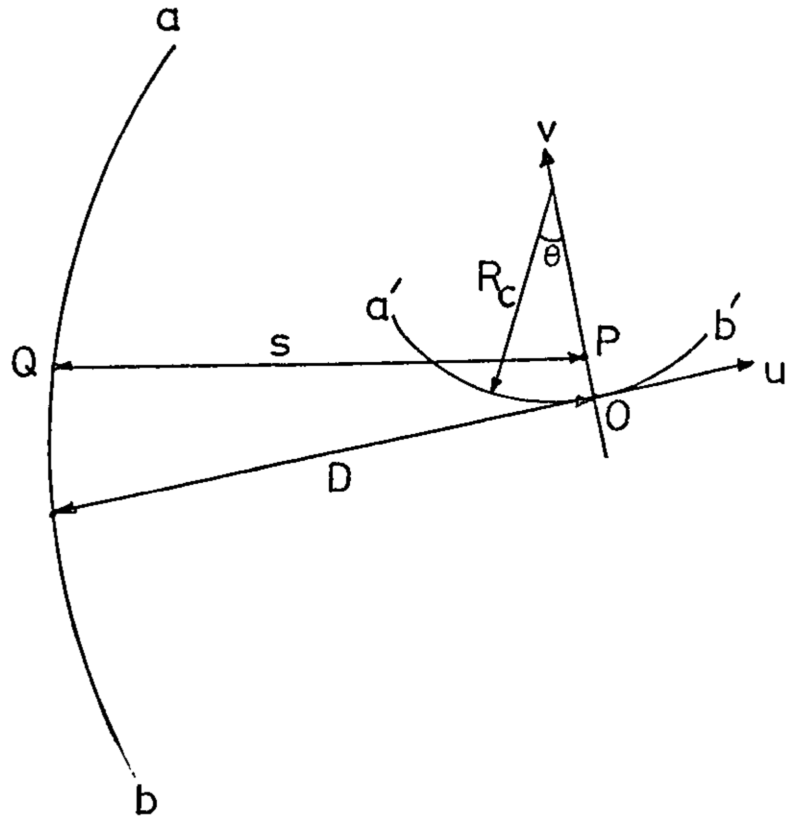


Figure 61. The wavefront (ab) and caustic (a'b') for the derivation of the diffraction at a caustic.

5.5 Diffraction at a Caustic Surface

In Figure 61, ab is a section of a monochromatic wave front which from geometrical theory forms the caustic a'b'. The co-ordinate system u,v has its origin at an arbitrary point O on the caustic, with v coinciding with the radius of curvature and positive v in the direction of the centre of curvature of the caustic at point O.

The disturbance at P a distance v from O due to the contribution from the element Q of the wavefront will be given by Equation (83) as

$$d(U(P)) = A \frac{e^{iks}}{s} \quad dS \tag{190.}$$

where A is the disturbance at Q, s is the distance from P to Q and dS represents the element of the wavefront at Q. Following the derivations of reference (45 pp.148) s may be expressed as

$$s \simeq D - v\theta - \frac{1}{6} R_c \theta^3 \tag{191.}$$

where D and θ are as shown in Figure 61 and R_c is the radius of curvature of the caustic at point O.

Following the diffraction theory of Chapter 3, the total disturbance at P is given by Equation (95), which on substitution of Equation (191) and assuming $K(\chi)$ constant becomes

$$U(P) \sim \int_{-\infty}^{+\infty} e^{-ikv\theta - i \frac{kR_c\theta^3}{6}} \quad d\theta \tag{192.}$$

The term $\frac{1}{s}$ will vary only slowly in comparison with the exponential terms and may be neglected.

Equation (192) may be expressed

$$U(P) \sim 2 \int_0^{\infty} \cos(kv\theta + \frac{kR_c\theta^3}{6}) \; d\theta \tag{193.}$$

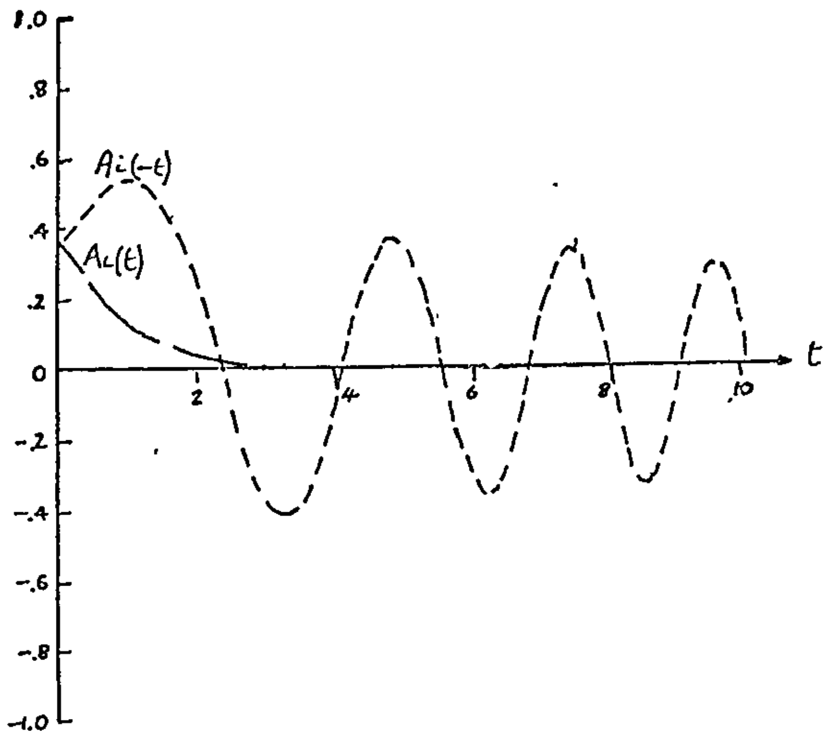


Figure 62. The Airy function.

and introducing a new variable

$$\xi = (\frac{kR_c}{2})^{1/3} \theta,$$

$$U(P) = \Phi \left(v(\frac{2k^2}{R_c})^{1/3}\right) \qquad 194.$$

where $\Phi(t)$ is the Airy function defined as

$$\Phi(t) = \frac{1}{\sqrt{\pi}} \int_0^\infty \cos(\frac{\xi^3}{3} + \xi t)\, d\xi, \qquad 195.$$

The Airy function, shown in Figure 62, decays exponentially for positive argument and oscillates with decreasing amplitude for negative argument. If A represents the electric field of the wave ab then from Equation (27), the intensity at P is given by

$$I(P) \simeq |U(P)|^2 \simeq 2B(\frac{2k^2}{R_c})^{1/6} \Phi^2 \left(v(\frac{2k^2}{R_c})\right)^{1/3} \qquad 196.$$

where from reference 45, B represents the intensity from the caustic calculated from geometrical ray theory and neglecting diffraction effects.

The argument t of the Airy function shown in Figure 62 and the spatial variable v, are related by

$$v = t(\frac{R_c}{2k^2})^{1/3} \qquad 197.$$

The oscillatory behaviour of $\Phi(t)$ for negative argument will produce intensity fringes which are tangent to the caustic with spatial and frequency dependence given by inserting the relevant values of R_c, k and t into Equation (197). For example, the maximum value of $\Phi(t)$, 0.549

is attained when $t = -1.02$ and if $R_c = 25.10^{-6}$, $k = \dfrac{2\pi}{.63.10^{-6}}$

then the corresponding position of the maximum of intensity is $v = -0.5.10^{-6}$. The positive value of v at which the intensity $I = \frac{1}{10} I_{MAX}$ for the same values of R_c and k, is $v \simeq 1.0.10^{-6}$ indicating that the shadow boundary is displaced by 4% from its geometric position.

5.6.1 Experimental Investigation of Caustics

The experimental arrangement of Figure 21 was used to illuminate the entrance apertures of short lengths of core cladded waveguides with plane waves from either a white light or laser source. The short lengths of waveguide were prepared as described in Section 2.5 and the measurements and microphotographs of the caustics presented in this section were made with the microscope focussed on the exit end of the waveguides.

The discussion of Section 4.3 suggests that the light incident upon the entrance aperture of the waveguides will suffer diffraction similar to that produced by a circular aperture corresponding to the core cross sectional dimensions. The study of diffraction by circular apertures contained in reference (31) indicates that after diffraction, 85% of the incident light is contained within an angular width

$$\sim \frac{0.61\lambda}{a}$$

where a is the radius of the aperture.

In the experimental arrangement of Figure 21, the plane waves are incident on the entrance aperture of the waveguides at angles θ_o, α. After diffraction by the entrance aperture 85% of the incident light will be within the cone defined by the azimuthal and axial semi-angles $\Delta\alpha$, $\Delta\theta_o$ respectively where

$$\Delta\alpha \sim \Delta\theta_o \sim \frac{0.61\lambda}{a}, \tag{198.}$$

if the contraction of the aperture at high axial angles of incidence is neglected.

The azimuthal component of the cone $\Delta\alpha$ is of the order of 0.5^o for the values of a and λ used in these experiments and has an insignificant effect upon the experimental results. The axial component $\Delta\theta_o$ represents the divergence of the light of a skew plane along the path of the skew plane ray. After refraction into the core the cone will have an axial semi-angle $\Delta\theta_1$, centred at angle θ_1, where (assuming $n_o = 1$)

$$\sin\theta_1 = \frac{1}{n_1}\sin\theta_o \tag{199.}$$

and

$$\Delta\theta_1 = \frac{\Delta\theta_o}{n_1} \simeq \frac{0.61\lambda}{n_1 a}. \tag{200.}$$

Substitution of $(\theta_1 \pm \Delta\theta_1)$ into Equation (148) and assuming that $\Delta\theta_1 \sin\theta_1 << \cos\theta_1$ gives

$$m(\phi\ell) = m \pm \Delta m \tag{201.}$$

where

$$m = \frac{\ell\tan\theta_1}{2a\sin\phi} \tag{202.}$$

$$\Delta m = \frac{\Delta\theta_1\ell}{2a\sin\phi}. \tag{203.}$$

The term m given by Equation (202) indicates the position of the skew plane ray and hence the position of the skew plane, which will

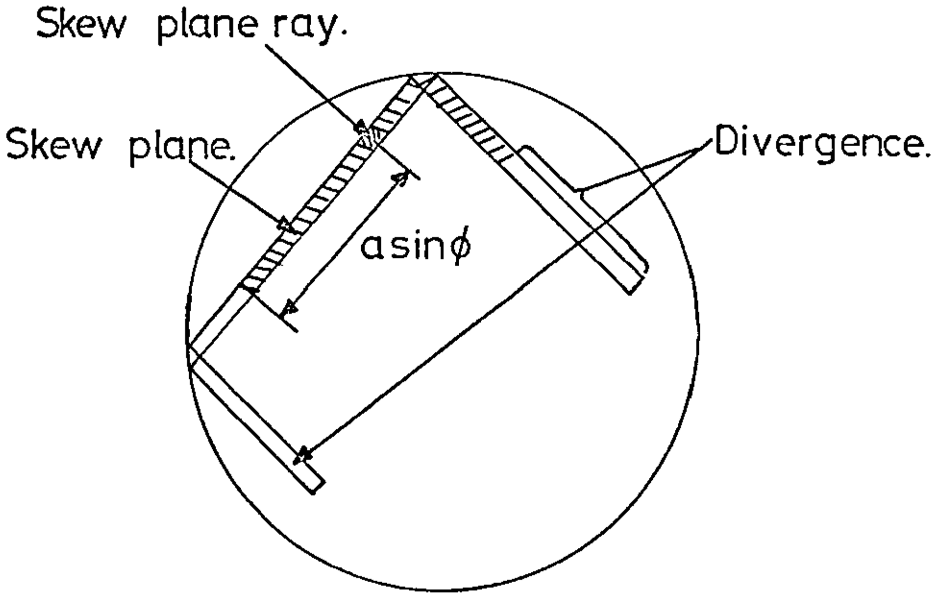


Figure 63a. A diverging skew plane, m non-integer.

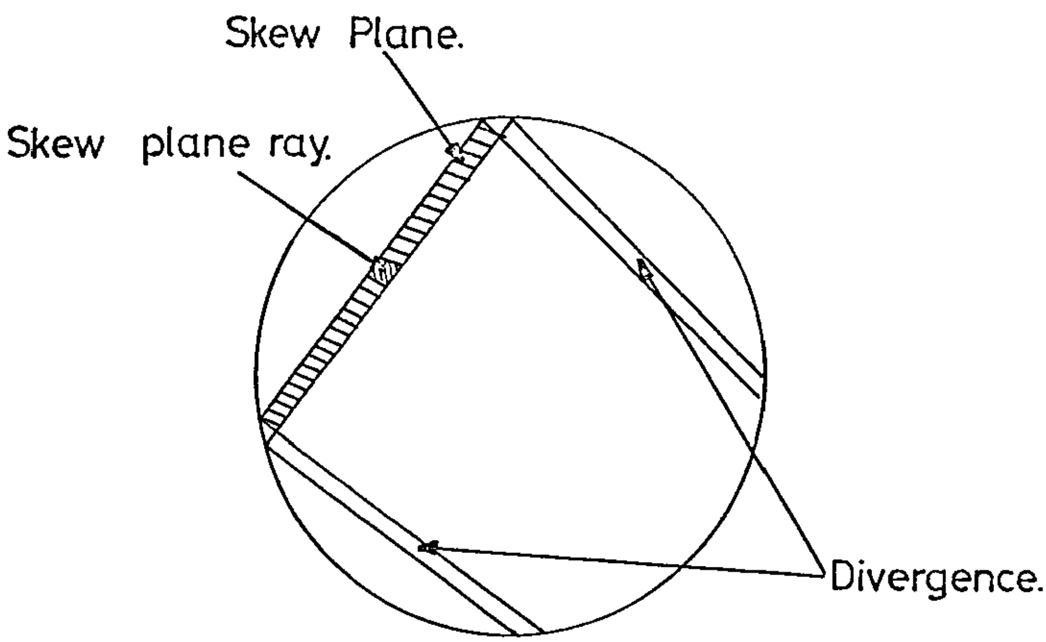


Figure 63b. A diverging skew plane, m an integer.

extend a distance a sin ϕ from the skew plane ray along the skew plane ray path in each direction. The term Δm given by Equation (203), will be called the divergence factor, indicates a divergence of the light originally within the skew plane an additional distance along the skew plane ray path of Δm a sin ϕ. This is illustrated in Figure 63 for m not an integer, $\Delta m = 0.5$, Figure 63a, and m an integer, $\Delta m = 1$, Figure 63b.

It follows from the conservation of intensity theorem and consideration of Figure 63 that the intensity of the skew plane $I(\phi)$ given by Equation (152) will be reduced by the factor $\frac{1}{(1 + 2\Delta m)}$ and the 'lost' intensity appears in the $2\Delta m$ additional skew planes.

The caustics derived in Section 5.4 and described by Equations (169a,b) are based upon integer and single values for m, and it was also assumed that m is independent of ϕ. According to Equation (201), (202) and (203) $m(\phi\ell)$ will, in general, be non-integer, non singular (that is $\Delta m \neq 0$) and dependent upon ϕ. The conditions shown in Figure 63 suggest that a non-integer value for m will result in the partial illumination of adjacent caustics and that when $\Delta m \neq 0$ a total of $(1 + 2\Delta m)$ caustics will be illuminated.

To investigate the effects of the ϕ dependence of $m(\phi\ell)$ on the caustic equations it is assumed that m' is the value of m for $\phi = 90^\circ$, i.e.

$$m' = \frac{\ell \tan \theta_1}{2a} \qquad 204.$$

Substituting Equation (204) into (202), (203)

$$m = \frac{m'}{\sin \phi} \qquad 205.$$

$$\Delta m = \frac{m' \Delta\theta_1}{\sin \phi \tan \theta_1} \qquad 206.$$

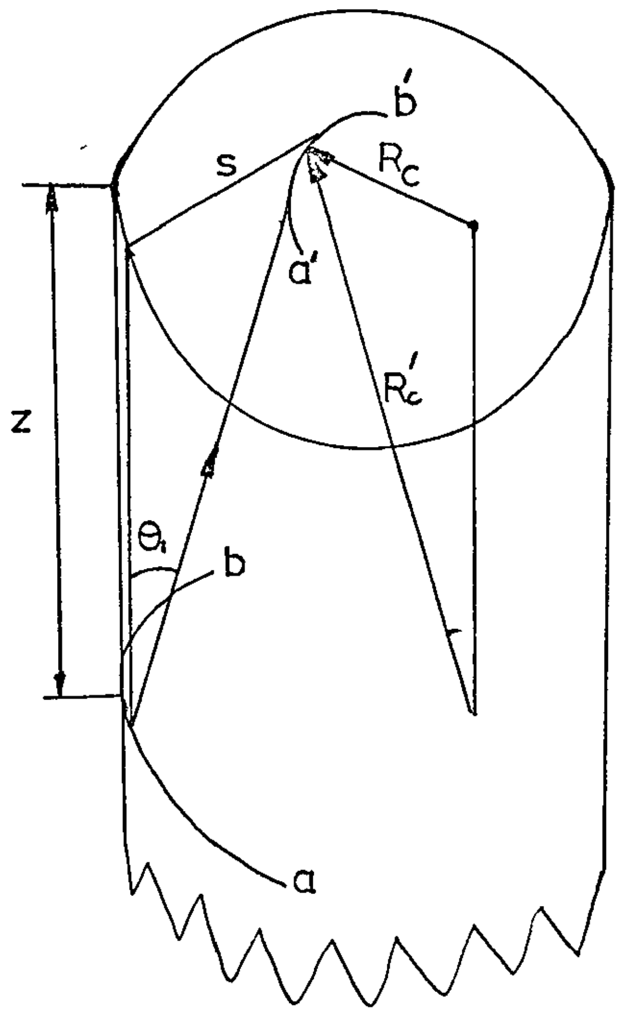


Figure 64. The axial separation of the wavefront and caustic of Figure 61.

Assuming that Equation (205) represents integer values and substituting into Equation (180) gives

$$\eta(\phi) = \frac{2m'\phi}{\sin\phi} \qquad 207.$$

Equation (207) together with Equation (180b) represent the same spiral as Equation (180) for $\phi \sim \frac{\pi}{2}^{(\chi\sim 0)}$, but for $\phi \rightarrow 0 (\chi \rightarrow \frac{\pi}{2})$ the $\phi/\sin\phi$ term of Equation (207) will approach unity[46 Eqn.4.3.74] instead of zero as in Equation (172) and the spirals will now terminate at

$$\eta = 2m' \text{ (radians)} \qquad 208a$$

$$R = a \qquad 208b.$$

Equations (208) represent the helical path of the skew plane, tangent to the core cladding interface, as it propagates down the waveguide as m' increases.

Since the caustics are formed by light which is propagating down the waveguides there will be an axial separation between the wavefront ab and the caustic a'b' of figure 61. In Figure 64 this axial separation is shown as z and the effective radius of curvature of the caustic R_c' is given by*

$$R_c' = \frac{R_c}{\sin\theta_1} \qquad 209.$$

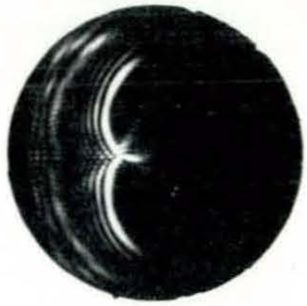
where θ_1 is the axial angle of the rays forming the caustic.

Substitution of Equation (209) and $k = n_1 \frac{2\pi}{\lambda}$ into Equation (197) gives

$$v = tF \qquad 210.$$

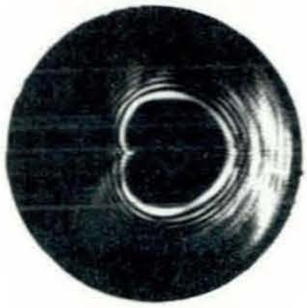
*Assuming R_c and s << z, i.e. θ_1 small

m' = 1



m' = 2

m' = 3



m' = 4

a = White light

m' = 1

m' = 2

m' = 3

m' = 4

b. Laser light

Figure 65. Microphotographs of caustics observed in ≈ 50μ diameter cladded dielectric waveguides.

where

$$F = \left(\frac{R_c \lambda^2}{\sin\theta_1 \; 2(n_1 \; 2\pi)^2}\right)^{1/3} \qquad 211.$$

and F may be called the fringe factor since it scales the Airy function to give the physical fringe spacing.

The other factors which are used to interpret the experimental results are summarised below and expressed in terms of known or measurable quantities. The caustics for $\phi \sim \frac{\pi}{2}$ are of most interest and thus the factor m' is used as the reflection number indicator and will be called the reflection number.

Summary

Reflection number $\quad m' = \dfrac{\ell \tan\theta_1}{2a}$

Skew reflection number $\quad m = \dfrac{m'}{\sin\phi}$

Divergence factor $\quad \Delta m = \dfrac{m' \, 0.61\lambda}{n_1 \, a \sin\phi \tan\theta_1} \qquad 212.$

where Equation (212) is obtained by substituting Equation (200) into (206).

5.6.2 Experimental Results

In Figure 65 the caustics produced by both white light and laser light for m' = 1 to 4 are illustrated. The diffraction patterns associated with the caustics obtained with laser illumination contain additional fringes produced partly by the superposition of the incident and reflected waves and partly as the result of scattered laser light originating from the entrance aperture or from the reflection interfaces. These secondary fringes tend to obscure the diffraction patterns due to the caustics.

Similar secondary fringe systems will be produced by each monochromatic component of the white light source, but since the position of these fringes is frequency dependent, the fringe systems due to different frequency components will overlap and no fringes will be visible. This is demonstrated in Figure 65a where the caustic for m' = 2 produced by white light contains no secondary fringes but only the fringes due to the caustics or those produced by the overlapping of the diffraction patterns of the caustics due to the $\pm\phi$ skew planes.

The absence of the secondary fringes, when illuminating with white light may be interpreted as being the result of the incoherence of the light at the point of observation, as discussed in Section 3.6.5. Such a statement would appear to be inconsistent with the appearance of the caustic diffraction fringes at similar points of observation.

This difficulty is resolved by remembering that the coherence of a source is determined by measuring the visibility of fringes produced by an optical system which arranges to superpose at least two components of light from the same source. In the laboratory experiments (31 Cp.7 associated with interference the optical system generally produces only a single fringe system, which in turn suggests only two possible optical paths through the system, from the source to the point of observation.

If the optical system has a multiplicity of optical paths from the source to the points of observation, as is the case for the optical waveguide, then it is clearly possible for light taking certain of these paths to exhibit coherent properties whilst light following other paths may have differential path lengths exceeding the coherence length of the source and thus exhibit incoherent properties.

Figure 66.

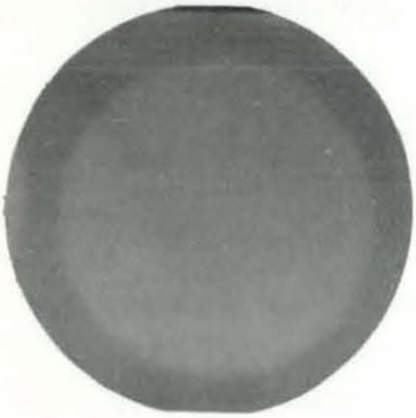


Figure 67a.

Figure 67b.

The caustics are produced by light which follows almost identical paths down the waveguide. If the difference in optical path lengths is very much less than the coherence length of white light, the white light fringes shown in Figure 65a will be produced.

The fringe factor and divergence factor for the caustics shown in Figure 65 are given in Table 7 below.

m'	$F.10^{-6}m.$	Δm
1	1.15	.017
2	.53	.070
3	.37	.075
4	.38	.125

Table 7

The inversion and contraction of the caustics and the spacing of the diffraction fringes show good agreement with the theoretical values.

In Figure 66 the caustic for m' = 2 produced by white light is illustrated but now with a fringe factor $F = 17.10^{-6}$ obtained by increasing the length of waveguide. If ℓ is increased, then θ_1 is reduced to maintain m' = 2 in Equation (204) and the fringe factor which is inversely proportional to θ_1 is increased. As well as increasing the spacing between the fringes the increase in F permits the appearance of different frequency components of the diffraction fringes and Figure 66 has a coloured appearance.

A further increase in the length of the waveguide increases F and θ_1 approaches zero. Ultimately when F is very large, each fringe

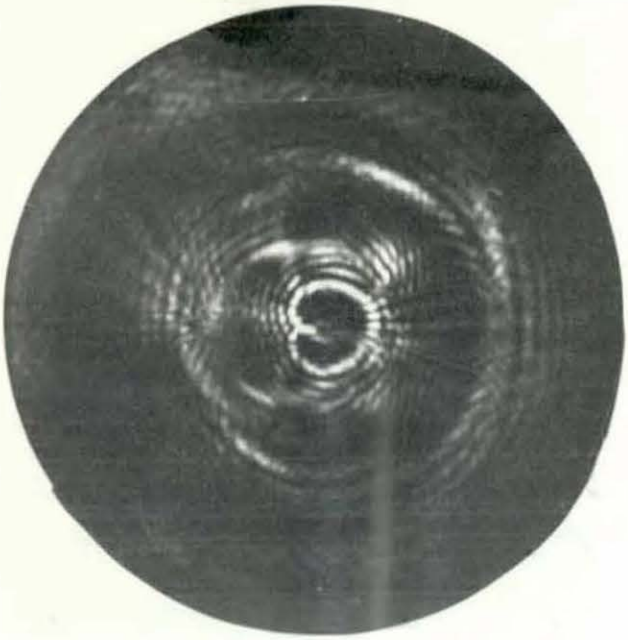


Figure 68

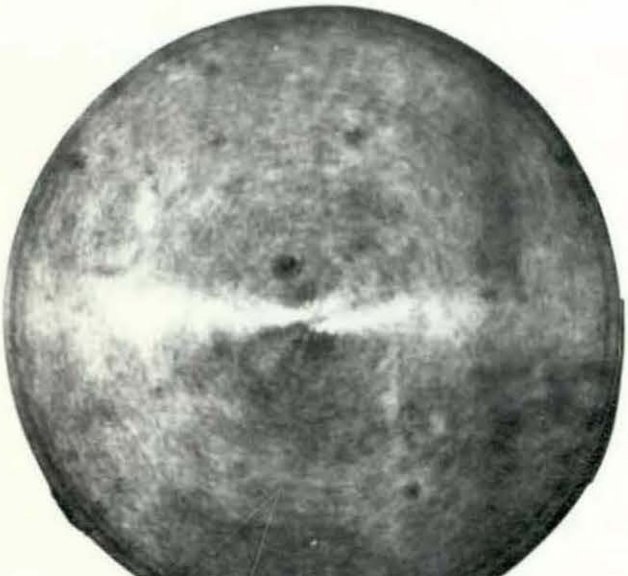


Figure 69.

of the diffraction pattern for low m' will fill the entire aperture of the waveguide, and the different frequency components will overlap. This will produce a uniform white intensity distribution as illustrated in Figure 67a where $\theta_1 = 1^o$ and $\ell = 4.0 cm$ If the same length of waveguide is illuminated using the laser source at the same angle of incidence the resulting intensity distribution is shown in Figure 67b.

Although it is difficult to differentiate between them, there are two types of interference fringes present in Figure 67b. The series of concentric rings correspond to waveguide mode patterns when several modes are excited simultaneously and the larger fringes randomly distributed over the waveguide end face are the interference effects of randomly scattered light.

Concentrating now on white light illumination of the waveguides, Fig ure 68 shows the caustics obtained for m = 12 in a length of waveguide $\ell = 2.0 mm.$ which gives $\theta_1 = 16^o$ The resulting small value of $F = 0.5\ 10^{-6}$ gives the fringe spacing shown in Figure 68 and such a small value for F gives fringe positions which are virtually independent of frequency to give the white fringe pattern illustrated.

Maintaining the value of θ_1 and increasing the length of the waveguide gives increasing values for m' and the well defined caustic shapes obtained for low m disappear as the spiral winds up, as demonstrated in Figure 69. The intensity distributions shown in Figure 69 are white and contain spatial variations of intensity similar to those shown in Figure 5.

If the waveguides are circular in cross section the intensity distributions should be circular symmetric and therefore independent of the azimuth angle of incidence of the white light. It was found that none of the waveguides tested exhibited these characteristics suggesting that all the waveguides tested were non circular.

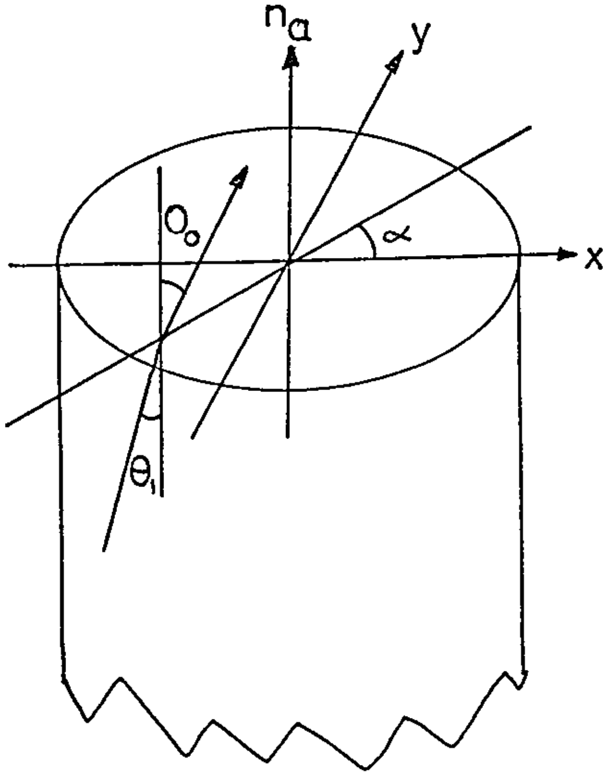


Figure 70. The refraction of rays through the end face of a dielectric waveguide.

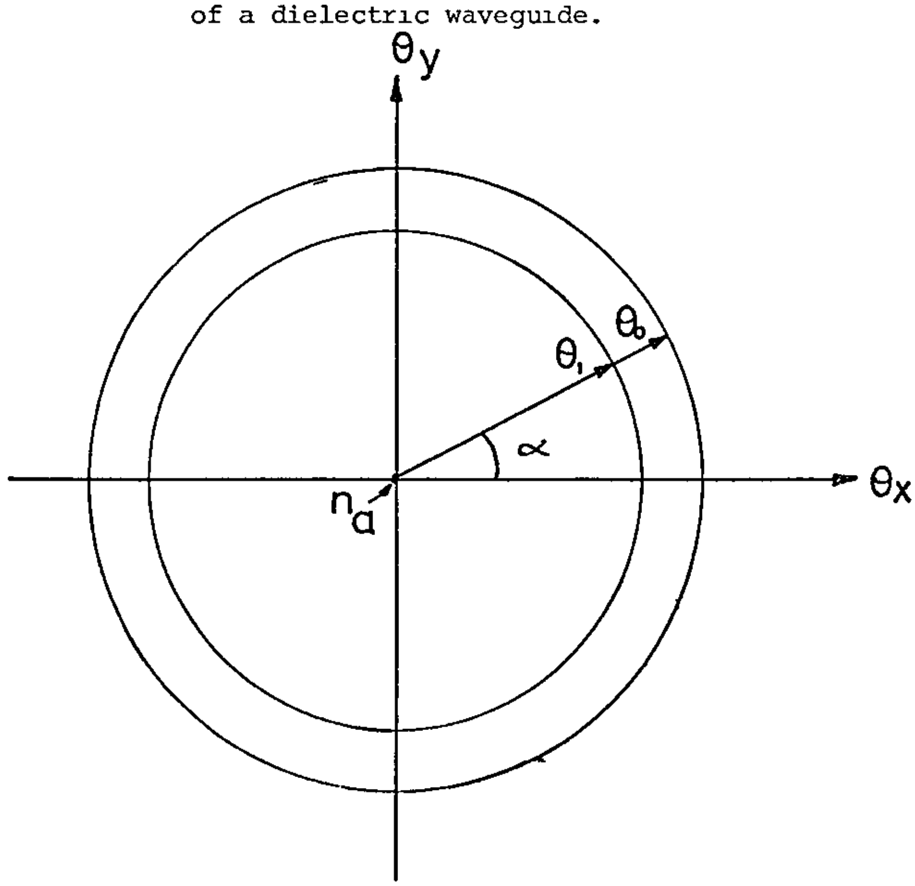


Figure 71. The representation of refraction in cartesian co-ordinates.

5.7 The Effects of Sloping Terminations

The radiation of a hollow cone of light from an optical waveguide when excited by a plane wave was discussed in Section 5.4 The cone of light is formed from a plane wave by divergence in the azimuthal plane as a result of the multiple reflection process, whilst the axial components of the angles of incidence of the light remain constant.

The constancy of the axial angle of incidence is assured if the waveguide is perfectly straight, with end terminations normal to the axis of the waveguide and any divergence due to diffraction neglected. The effect of a bend in the waveguide is considered in Chapter 7, and the result of divergence was discussed in Chapter 2. In this section the effects on the radiated cone of light of a slope at the radiating end face of the waveguide will be considered, assuming a straight waveguide with a normal end face at the entrance aperture.

If the radiating end of the waveguide is normal to its longitudinal axis, then rays of light will refract through the end face as shown in Figure 70 where according to Snell's Law (Equation (100))

$$n_1 \sin \theta_1 = n_o \sin \theta_o \qquad 213.$$

Equation (213) may be expressed in diagrammatic form as shown in Figure 71. The angles of incidence θ_1, and refraction θ_o, are each resolved along orthogonal components θ_x, θ_y which form the axes of a cartesian co-ordinate system. The origin of the co-ordinate system ($\theta_x = \theta_y = 0$) represents the longitudinal axis of the waveguide which also forms the normal n_a to the radiating surface.

Since the angles θ_1, θ_o, are independent of azimuthal position their locii are circles of radius θ_1 and θ_o respectively. All rays of

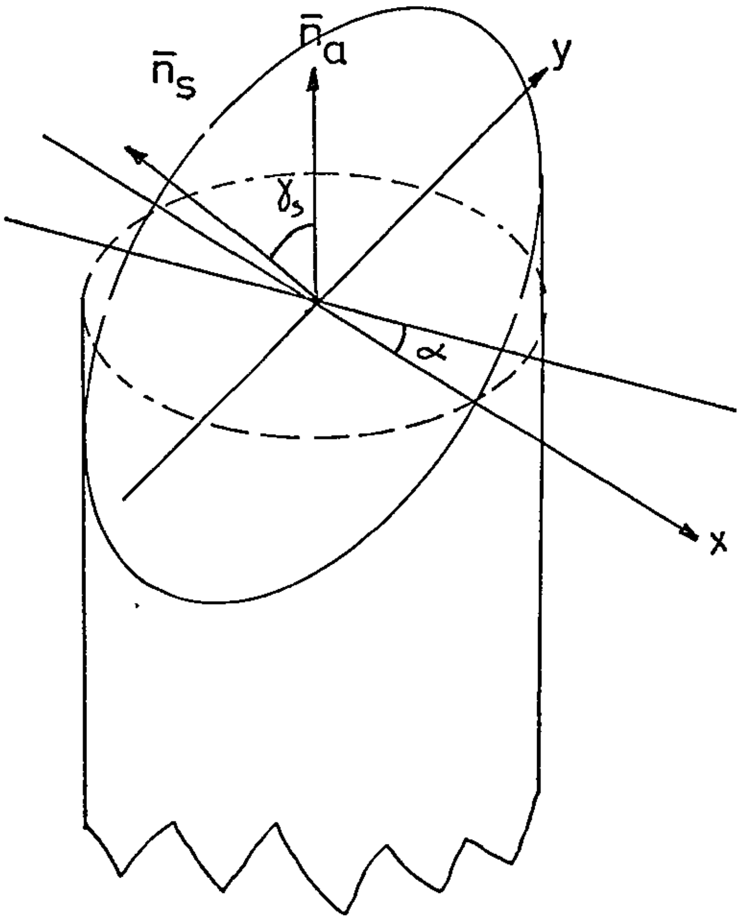


Figure 72. A sloping end face on a waveguide termination.

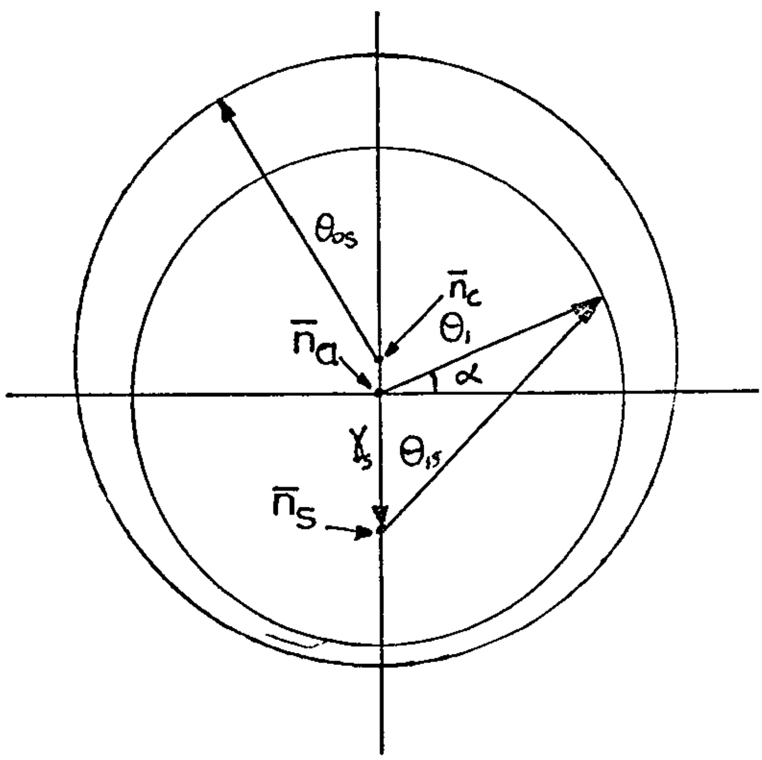


Figure 73. The cartesian co-ordinate representation of refraction at the sloping end face shown in Figure 72.

light which are incident and refracted in a plane parallel to or in a meridional plane at an azimuthal angle α are represented by the vectors in the θ_x, θ_y plane at angle α as shown in Figure 71. Positive angles are measured clockwise from the surface normal.

A flat sloping end face will have a surface normal $\bar{n}_s$ at an angle γ_s to the waveguide axis, $\bar{n}_a$. The co-ordinate system may be rotated until this angle lies solely along either the θ_x or θ_y axis. In Figure 72 the slope is aligned along the y axis and in Figure 73 the slope angle γ_s appears along the θ_y axis to give the position of the normal to the sloping surface $\bar{n}_s$ a distance γ_s along the θ_y axis from the origin $\bar{n}_a$.

Rays of light within the waveguide will still form angles θ_1 with respect to the normal $\bar{n}_a$ and this is represented by the circular locus centre $\bar{n}_a$, radius θ_1 in Figure 73. The angles of incidence θ_{1s} with respect to the sloping end face are given by the length of the vector from the point $\bar{n}_s$ to the circumference of this circular locus.

From Figure 73

$$\theta_{1s} = \sqrt{(\theta_1 \sin\alpha + \gamma_s)^2 + (\theta_1 \cos\alpha)^2}$$

$$= \theta_1 \sqrt{1 + \left(\frac{\gamma_s}{\theta_1}\right)^2 + \frac{2\gamma_s}{\theta_1} \sin\alpha} \qquad 214$$

It will be assumed that $\gamma_s \ll \theta_1$ so that $\left(\frac{\gamma_s}{\theta_1}\right)^2$ may be neglected, and

$$\left(1 + \frac{2\gamma_s}{\theta_1} \sin\alpha\right)^{\frac{1}{2}}$$

may be expanded using the binomial theorem to give

$$\theta_{1s} = \theta_1 + \gamma_s \sin\alpha \qquad 215.$$

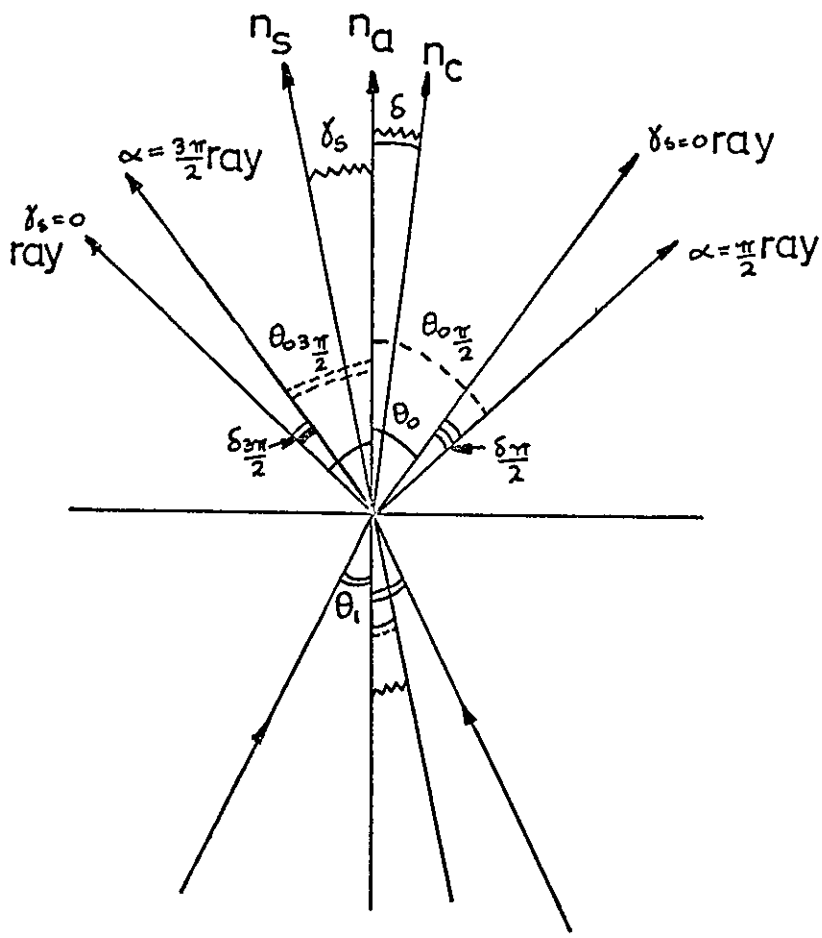


Figure 74. The definition of angles for the refraction of rays at a sloping end face.

The angle of refraction θ_{os} is given by Snell's Law as

$$\sin \theta_{os} = \frac{n_1}{n_o} \sin \theta_{is} \qquad 216.$$

which on substitution of Equation (213) and (215) and assuming $\cos(\gamma_s \sin \alpha) \sim 1$ becomes

$$\sin \theta_{os} = \sin \theta_o + \frac{n_1}{n_o} \gamma_s \cos \theta_1 \sin \alpha \qquad 217.$$

The locus of θ_{os} is shown in Figure 73 where it has been assumed that $\gamma << \theta_o$ to give $\theta_{os} \simeq \theta_o$ when $\alpha = o, \pi$. The centre of symmetry of this locus, which may be called the normal $\overline{n}_c$ of the radiation "cone" of light, will lie on the θ_y axis midway between the $\theta_{os}(\alpha = \frac{\pi}{2})$ and $\theta_{os}(\alpha = \frac{3\pi}{2})$ points.

Since

$$(\theta_{os}(\alpha = \tfrac{\pi}{2}) - \gamma_s) > (\theta_{os}(\alpha = \tfrac{3\pi}{2}) + \gamma_s)$$

the cone normal $\overline{n}_c$ will have an angle δ in the opposite direction along the θ_y axis to the slope normal and where

$$\delta = \frac{\theta_{os}(\alpha = \pi/2) + \theta_{os}(\alpha = 3\pi/2)}{2} - (\theta_{os}(\alpha = \tfrac{3\pi}{2}) + \gamma_s) \qquad 218.$$

A simpler expression for δ is obtained by considering the deviation of the refracted rays at azimuthal angles $\alpha = \frac{\pi}{2}, \frac{3\pi}{2}$ as a slope is introduced on the end face. It is convenient to define the angles as shown in Figure 74 and considering the refraction of the $\alpha = \frac{3\pi}{2}$ ray, Snell's Law gives

$$n_1 \sin(\theta_1 - \gamma_s) = n_o \sin(\theta_{o\,\frac{3\pi}{2}} - \gamma_s) \qquad 219.$$

Expanding the sine functions in Equation (219)

$$\frac{n_1}{n_o}(\sin\theta_1\cos\gamma_s - \cos\theta_1\sin\gamma_s) = (\sin\theta_{o\,\frac{3\pi}{2}}\cos\gamma_s - \cos\theta_{o\,\frac{3\pi}{2}}\sin\gamma_s) \qquad 220.$$

and substituting for θ_1 using Equation (213) and rearranging gives

$$2\cos\frac{(\theta_o + \theta_{o\,3\pi/2})}{2}\sin\frac{(\theta_o - \theta_{o\,3\pi/2})}{2} = \mathrm{TAN}\,\gamma_s\left(\sqrt{\left(\frac{n_1}{n_o}\right)^2 - \sin^2\theta_o} - \cos\theta_{o\,3\pi/2}\right) \qquad 221.$$

From Figure 74 $(\theta_o - \theta_{o\,\frac{3\pi}{2}}) = \delta_{\frac{3\pi}{2}}$ and assuming $\delta_{\frac{3\pi}{2}} \ll \theta_o$

so that $\theta_o + \theta_{o\,\frac{3\pi}{2}} \simeq \theta_o$ equation (221) becomes

$$2\cos\theta_o\sin\frac{\delta_{\frac{3\pi}{2}}}{2} = \mathrm{TAN}\,\gamma_s\left(\sqrt{\left(\frac{n_1}{n_o}\right)^2 - \sin^2\theta_o} - \cos\theta_o\right) \qquad 222.$$

Assuming that $\delta_{\frac{3\pi}{2}}$ and γ_s are small then

$$\delta_{\frac{3\pi}{2}} = \gamma_s\left(\sqrt{\left(\frac{(n_1/n_o)^2 - 1}{\cos^2\theta_o}\right) + 1} - 1\right) \qquad 223.$$

A similar calculation for the $\alpha = \frac{\pi}{2}$ ray gives

$$\delta_{\pi/2} = \delta_{\frac{3\pi}{2}} = \gamma_s R \qquad 224.$$

where

$$R = \sqrt{\left(\left(\frac{(n_1/n_o)^2 - 1}{\cos^2\theta_o}\right) + 1\right)} - 1 \qquad 225.$$

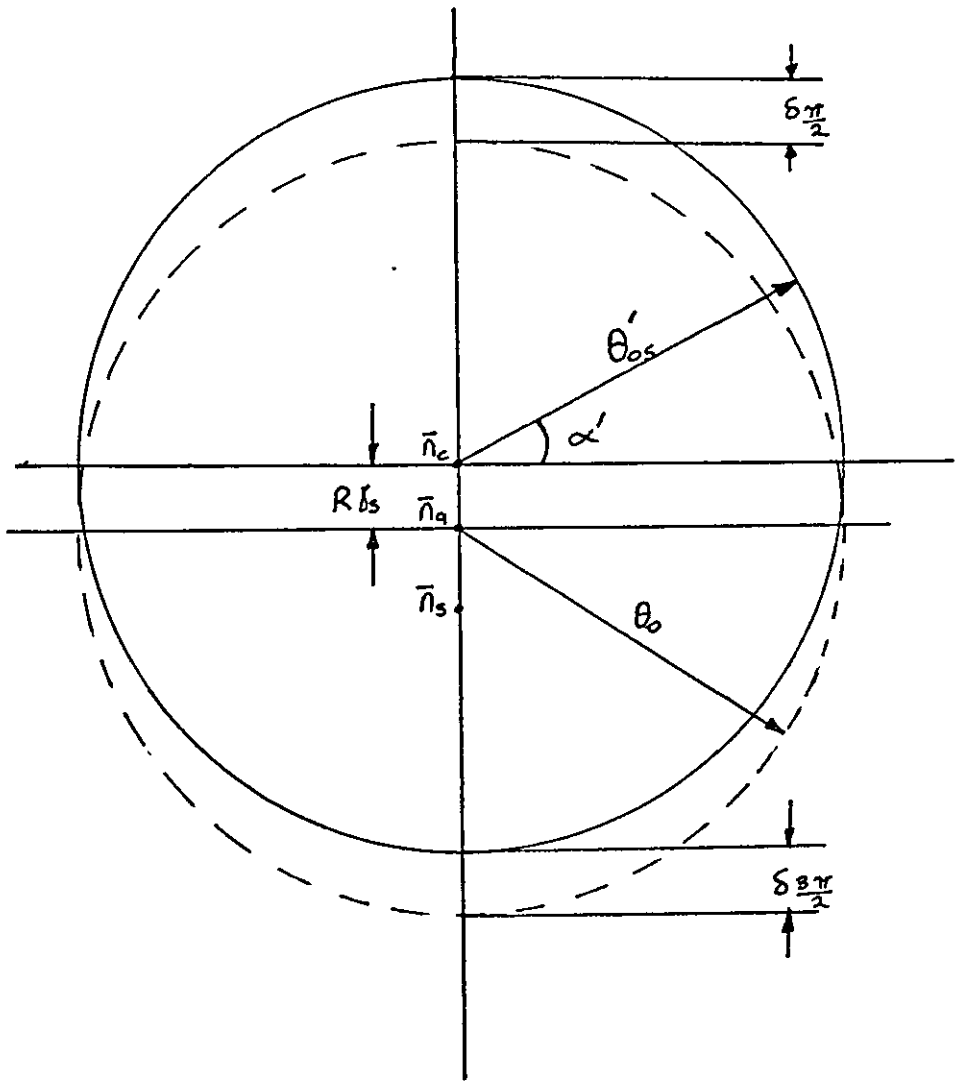


Figure 75. The locus of the angle of refraction θ'_{os} at a sloping end face.

and the cone normal $\overline{n}_c$ is offset from the waveguide axis by an angle δ where

$$\delta = \gamma_s R \qquad 226.$$

The semi-angle of the radiated cone of light, now measured with respect to the cone normal $\overline{n}_c$ may be denoted by $\theta'_{os}(\alpha')$ where α' and the locus of θ'_{os} are shown in Figure 75. Using the definitions of angles shown in Figure 74 and the results expressed by Equation (224)

$$\theta'_{os}(\frac{\pi}{2}) = \theta'_{os}(\frac{3\pi}{2}) \simeq \theta_o \qquad 227.$$

and using the assumption that $\delta << \theta_o$ then from Figure 75

$$\theta'_{os}(0) = \theta'_{os}(\pi) \simeq \theta_o \qquad 228.$$

It would appear from the above analysis that a flat slope on the radiation end of a waveguide will deflect the radiating cone of light by an angle δ given by Equation (226), in a direction opposite to the normal of the slope. The semi angle and circular symmetry of the radiating cone will be preserved as if it were refracted through a normally terminated waveguide, whose longitudinal axis coincides with the cone normal $\overline{n}_c$.

If the experimental procedure described in Section 2.8 is followed so that the end of the wave guide is ground normally to the radiation cone normal $\overline{n}_c$, the sloping end termination produced will have a slope angle γ_{s1} where

$$\gamma_{s1} = \delta \qquad 229.$$

and where the subscript $_1$ denotes the first alignment and grinding of this end of the waveguide. A second alignment and grinding will produce a sloping end termination at an angle γ_{s2} where

$$\gamma_{s2} = R\,\gamma_{s1} \qquad 230.$$

and the n^{th} alignment and grinding will produce a slope with angle γ_{sn} where

$$\gamma_{sn} = R^n\,\gamma_s \qquad 231.$$

Equations (226) and (229) have been used to substitute for δ and γ_{s1} respectively to give this equation.

A typical experimental value for R is given by substituting $n_1 = 1.62$, $n_o = 1$ and $\theta_o = 30^o$ into Equation (225) to give R = 0.779. If the initial slope angle $\gamma_s = 10^o$ this will be reduced by a factor R after each alignment and grinding procedure to give the slope angles shown in Table 8.

n	0	1	2	3	4	5	6	7	8	9	10
γ_{sn}^{o}	10	7.8	6.0	4.7	3.7	2.9	2.2	1.7	1.4	1.0	0.8

TABLE 8

A reduction in θ_o produces a smaller R factor which will reduce the number of grinding stages required to produce a given slope, but the reduction in θ_o is limited by the use of the approximation $\gamma_s \ll \theta_o$ in the analysis.

5.8 Conclusions

The radiation cone, the "black hole" effect and the thick lens phenomena discussed in this chapter are all well known properties of dielectric optical waveguides. However, the utilisation of the radiation cone to correct slopes on the radiation end of waveguides and the measurement of axial angles of incidence of light within the waveguides using the "black hole" effect do not appear to have been suggested previously.

The caustics produced by cylindrical reflectors are not of particular significance but their analysis provides a basis for the investigation of the more interesting caustics produced by non-circular cross section waveguides, for example, the elliptical cross section waveguide which is discussed in the next chapter.

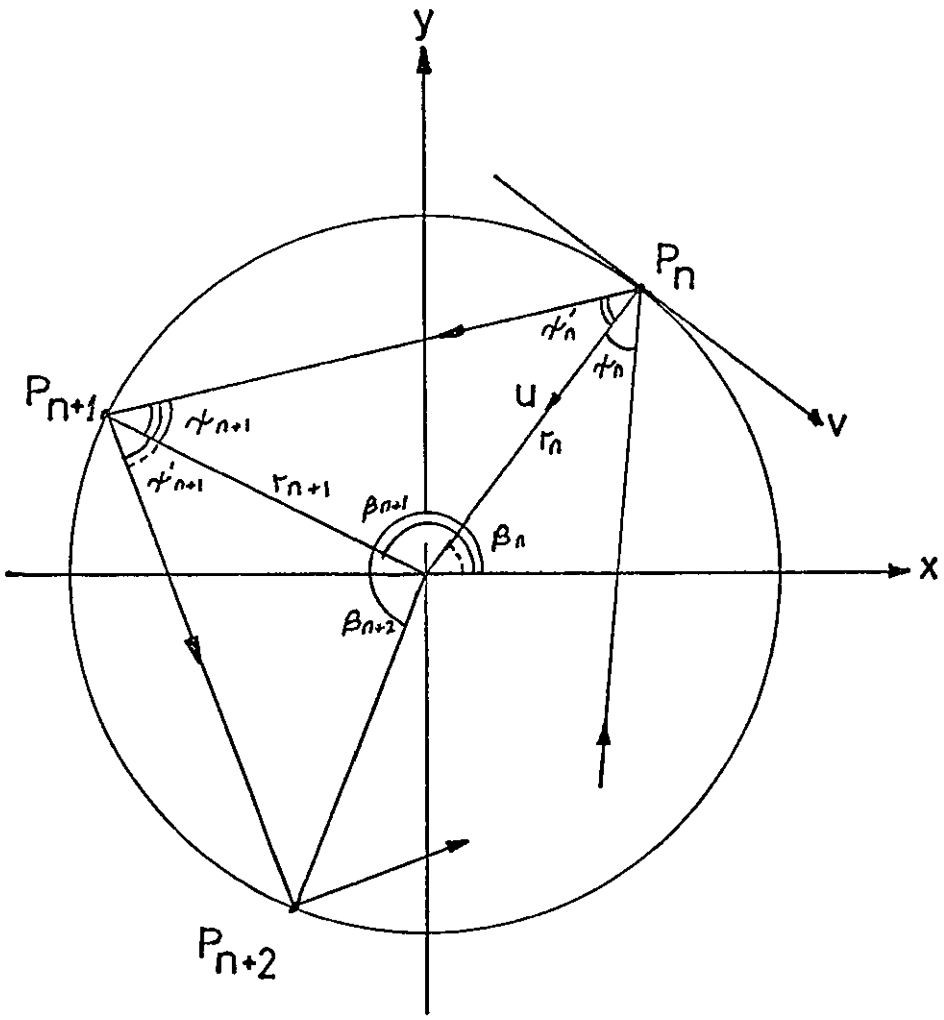


Figure 76. The cross sectional path of a ray in a cylinder.

CHAPTER 6

6.1 Introduction

The experimental results of Chapter 5 suggest that the initial assumptions used there for the analysis of ray propagation in the waveguides are incorrect. It appears that the waveguides do not have circular cross sections but have asymmetries which have a marked effect on the intensity distribution in the waveguide cross section.

In this chapter the waveguide is assumed to have an elliptic cross section and the positions and form of the caustics for this cross section are determined by perturbing the solutions for the circular cross section. To assist this analysis the caustic equations for the circular cross section are derived by considering the skew ray paths in the cross section in terms of angular difference equations.

It is shown that under certain experimental conditions the intensity distribution in the cross section is simply related to the cross sectional dimensions and geometry of the waveguide.

6.2 Difference Equations for Skew Ray Paths in Circular Cross Sections.

In Figure 76 the cross sectional path in a cylindrical waveguide of a skew plane ray ϕ is shown, where P_n denotes the point on the circumference at which the n^{th} reflection occurs. The polar co-ordinates of P_n are $r_n \angle \beta_n$ and the path of the ray is inclined at angle χ_n to r_n prior to reflection and angle χ_n' after reflection.

Consider a co-ordinate system u,v with its origin at P_n and the u axis coinciding with r_n and positive u in the direction of the centre of the circle. If, as shown in Figure 76, a ray is incident in the +u,+v quadrant then adjacent points of reflection P_n, P_{n+1}, P_{n+2} ... have angular separations of $\Delta\beta$ in an anticlockwise direction around the circle. From Figure 76

$$\Delta\beta = \beta_{n+1} - \beta_n = (\frac{\pi}{2} - \chi_n') + (\frac{\pi}{2} - \chi_{n+1}) \qquad 232.$$

Since the radius r_n of the circle is also the normal at the circumference, χ_n, χ_n' become the angles of incidence and reflection respectively and from the law of reflection will be equal in magnitude. i.e.,

$$\chi_n = \chi_n' \qquad 233.$$

From Figure 76

$$r_n \sin \chi_n' = r_{n+1} \sin \chi_{n+1} \qquad 234.$$

Since r_n and r_{n+1} are both radii of the circle then

$$\chi_n' = \chi_{n+1} \qquad 235.$$

Substituting Equations (233), (235) and (165a) (assuming $\Delta\chi \to 0$) into (232) gives

$$\Delta\beta = 2\phi_n$$

After m reflections the angular position of P_m with respect to P_n is given by β where

$$\beta = \int \Delta\beta . dm \qquad 237.$$

and, since $\Delta\beta$ is independent of m, the solution of Equation (237) is

$$\beta = 2\phi . m + \text{constant} \qquad 238.$$

and is in agreement with Equation (180a)

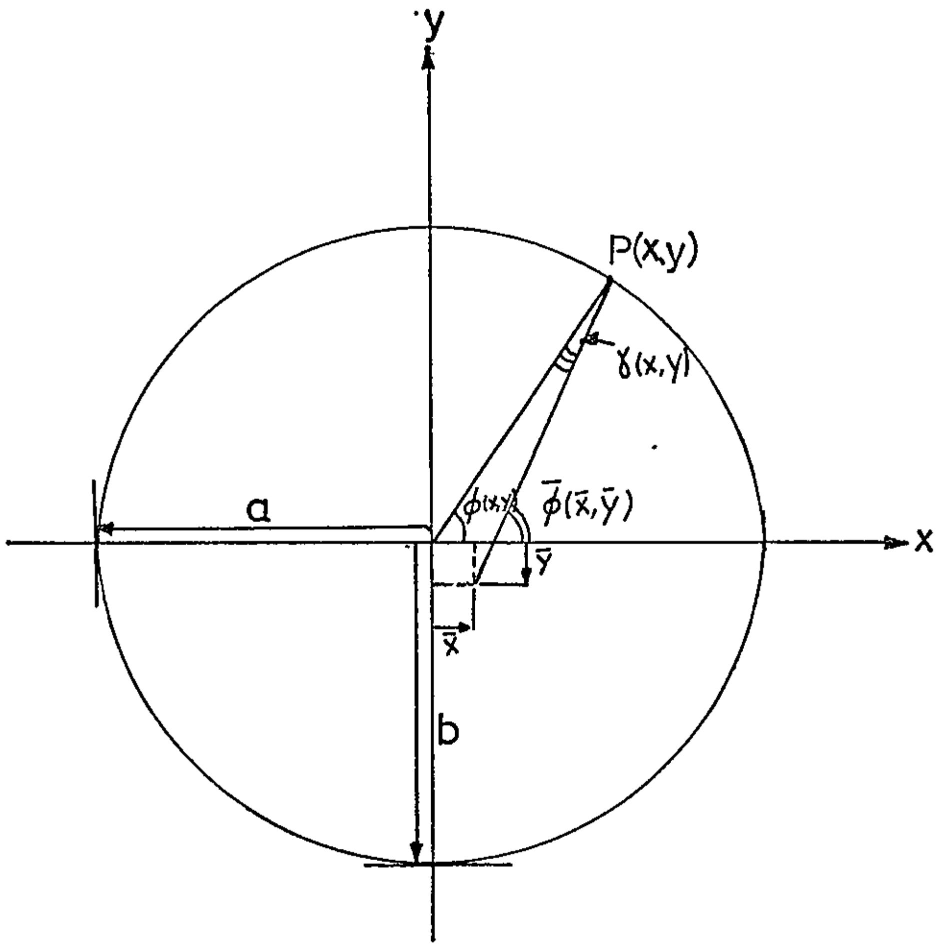


Figure 77. An ellipse.

If a ray is incident in the +u,-v quadrant then adjacent points of reflection are spaced at angular separations of $\Delta\beta$ but now in a clockwise direction around the circumference. Since both ray directions will have paths tangent to the circle radius $r \cos \phi$ this circle forms the caustic for the skew plane ϕ in agreement with Equation (180b).

6.3 Properties of Ellipses

The ellipse shown in Figure 77 is defined by the Equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad 239.$$

From Equation (239)

$$x = \frac{a}{b}\sqrt{b^2 - y^2} \quad 240a. \qquad y = \frac{b}{a}\sqrt{a^2 - x^2} \quad 240b.$$

Consider a point P(xy) on the ellipse whose polar co-ordinates $r\phi$ are given by (47, Pt.1,para.256)

$$r = \frac{ab}{(b^2\cos^2\phi + a^2\sin^2\phi)^{\frac{1}{2}}} \qquad 241.$$

$$\text{TAN } \phi = \frac{y}{x} = \frac{b}{a}\frac{y}{\sqrt{b^2 - y^2}} \qquad 242.$$

The centre of curvature of the ellipse at point P(xy) is given by $\bar{x}\ \bar{y}$ where from reference 48, pp 153

$$\bar{x} = x - \frac{y'(1+y'^2)}{y''} \quad 243a. \qquad \bar{y} = y + \frac{(1 + y'^2)}{y''} \quad 243b.$$

where $\qquad y' = \frac{dy}{dx}, \qquad y'' = \frac{d^2y}{dx^2}$

Substituting Equation (240) into (243) gives (48, pp. 153,Ex.3)

$$\bar{x} = \frac{a^2 - b^2}{a^4} x^3 \qquad 244a. \qquad \bar{y} = \frac{b^2 - a^2}{b^4} y^3 \qquad 244b.$$

The angle $\bar{\phi}$ between the radius of curvature at point P and the x axis as shown in Figure 77 is given by

$$\mathrm{TAN}\ \bar{\phi} = \frac{(y - \bar{y})}{(x - \bar{x})} \qquad 245.$$

Substituting Equations (240a) and (243a,b) into (245) gives

$$\mathrm{TAN}\ \bar{\phi} = \frac{a}{b} \frac{y}{\sqrt{b^2 - y^2}} \qquad 246.$$

Comparing Equations (242) and (246) it is seen that

$$\mathrm{TAN}\ \bar{\phi} = \frac{a^2}{b^2}\ \mathrm{TAN}\ \phi \qquad 247.$$

The radius of curvature $\bar{r}(xy)$ at point P(xy) on the ellipse is given by reference (48 pp. 153) as

$$\bar{r}(xy) = \frac{(1 + y'^2)^{3/2}}{y''}$$

Substituting for y' and y" gives

$$\bar{r}(x) = \frac{(x^2(b^2 - a^2) + a^4)^{3/2}}{a^4 b} \qquad 248.$$

Note that when x = 0

$$\bar{r}(0) = \frac{a^2}{b} \qquad 249.$$

and when x = a

$$\bar{r}(a) = \frac{b^2}{a} \qquad 250.$$

The last two equations give the vertex radii of an ellipse shown in reference (49, Equations F4.35).

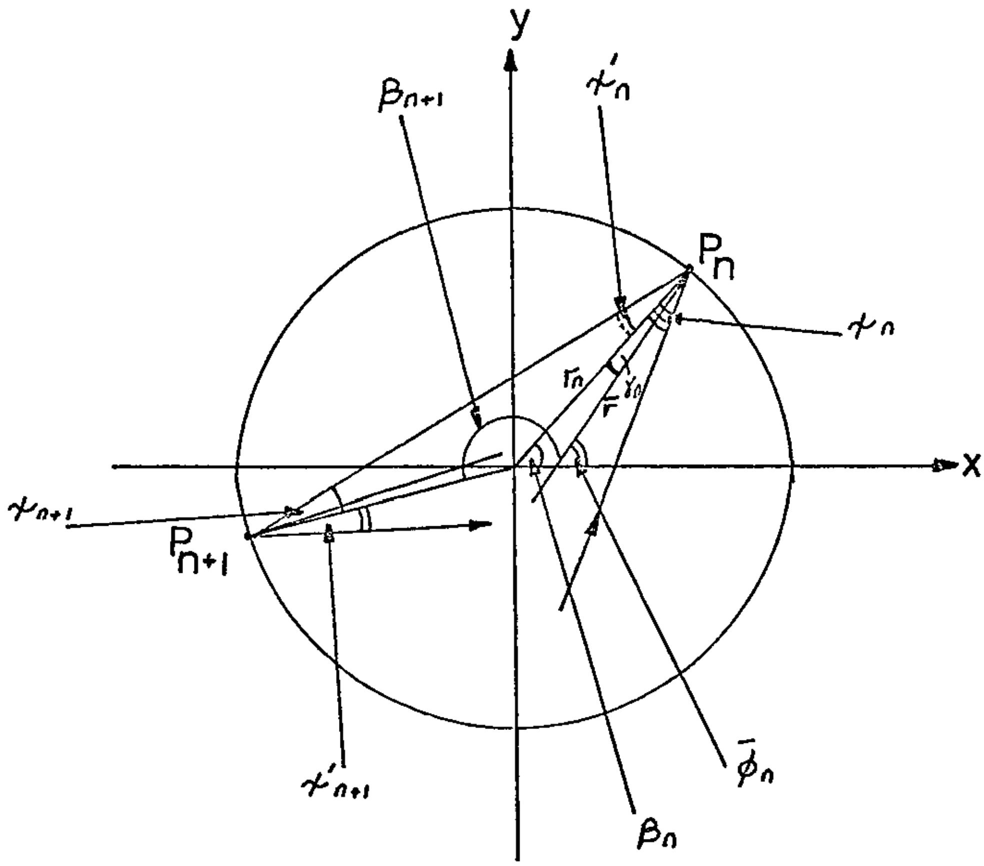


Figure 78. The path of a ray in an ellipse.

6.4.1 Propagation of skew plane rays in elliptical cross section waveguides.

In Figure 78 the cross sectional path in an elliptical cross section waveguide of an arbitrary skew plane ray is shown where P_n, χ_n, χ_n', r_n and β_n are the same as in Figure 76. In addition β_n, $\bar{\phi}_n$ correspond to $\phi(xy)$ and $\bar{\phi}(xy)$ respectively of Figure 77 and $\bar{r}$ is the radius of curvature of the ellipse and hence also the normal to the ellipse at the point P_n.

The angles of incidence and reflection of the skew plane ray at P_n are given by $(\chi_n - \gamma_n)$ and $(\chi_n' + \gamma_n)$ respectively where

$$\gamma_n = \bar{\phi}_n - \beta_n \qquad 251.$$

and by the law of reflection they will be equal in magnitude to give

$$\chi_n = \chi_n' + 2\gamma_n \qquad 252.$$

Equation (251) may be expressed as

$$\mathrm{TAN}\ \gamma_n = \mathrm{TAN}(\bar{\phi}_n - \beta_n) = \frac{\mathrm{TAN}\ \bar{\phi}_n - \mathrm{TAN}\ \beta_n}{1 + \mathrm{TAN}\ \bar{\phi}_n\ \mathrm{TAN}\ \beta_n} \qquad 253.$$

Noting that $\bar{\phi}_n = \bar{\phi}$ and $\beta_n = \phi$ then substituting for $\bar{\phi}_n$ in Equation (253) using Equation (247) gives

$$\mathrm{TAN}\ \gamma_n = \frac{(\frac{a^2}{b^2} - 1)\ \mathrm{TAN}\ \beta_n}{1 + \frac{a^2}{b^2}\ \mathrm{TAN}^2\ \beta_n} \qquad 254.$$

Assuming that γ_n is small, $\frac{a^2}{b^2} \simeq 1$, and denoting $(\frac{a^2}{b^2} - 1) = h$ Equation (254) becomes

$$\gamma_n = \frac{h}{2} \sin 2\beta_n \qquad 255.$$

The angles χ_n' and χ_{n+1} are related by Equation (234) which may be expressed in the form

$$\left(\frac{r_n}{r_{n+1}}\right)^2 = \left(\frac{\sin \chi_{n+1}}{\sin \chi_n'}\right)^2 \qquad 256.$$

Using Equation (241) the left hand side of this Equation becomes

$$\left(\frac{r_n}{r_{n+1}}\right)^2 = \frac{\cos^2 \beta_{n+1} \ (1 + \frac{a^2}{b^2} \mathrm{TAN}^2 \ \beta_{n+1})}{\cos^2 \beta_n \ (1 + \frac{a^2}{b^2} \mathrm{TAN}^2 \ \beta_n)} \qquad 257.$$

Using the assumption that $\frac{a^2}{b^2} \simeq 1$ which is discussed later in this Section, the right hand side of this Equation reduces to 1 thus giving the condition

$$\chi_{n+1} = \chi_n' \qquad 258.$$

Substituting Equations (265) and (258) into (252) and rearranging gives the difference Equation

$$\Delta\chi = \chi_n - \chi_{n+1}$$

$$= h \sin 2\beta_n \qquad 259.$$

A second difference equation is obtained by substituting Equation (258) into (232) to give

$$\Delta\beta = \pi - 2\chi_{n+1} \qquad 260.$$

The values of χ for the skew plane rays of most interest are small, for which $\Delta\beta$ is large. To make $\Delta\beta \to 0$ as $\chi \to 0$, $\Delta\beta$ is modified to $\Delta\beta'$ where

$$\Delta\beta' = -(\Delta\beta + \pi)$$

$$= 2\chi_{n+1} \qquad 261.$$

and β is now measured with respect to the +x and -x axis for alternate reflections.

To solve the difference equations (259), (261) the following ratio is formed which is then assumed to be independent of n. This assumption is justified if $\chi_{n+1} \simeq \chi_n$ since the ratio then applies for a single and arbitrary value of n.

$$\frac{\Delta\chi}{\Delta\beta'} = \frac{h \sin 2\beta_n}{2\chi_{n+1}} \qquad 262.$$

Assuming that $\Delta\beta' \to 0$ then Equation (262) becomes a differential equation which may be solved by separation of the variables and integration.

Separation of variables gives

$$\int 2\chi d\chi = \int h \sin 2\beta d\beta \qquad 263.$$

and integrating both sides gives

$$\chi^2 = -\frac{h}{2}\cos 2\beta + c \qquad 264.$$

where c is a constant of integration which may be evaluated by considering the condition $\beta = \frac{\pi}{2}$.

Substitution of $\beta = \frac{\pi}{2}$ into Equation (264) gives

$$\chi m^2 = \frac{h}{2} + c \qquad 265.$$

where χm is the maximum value of χ which occurs at $\beta = \frac{\pi}{2}$.

Substituting for c from Equation (265) into (264) gives

$$\chi = \sqrt{\chi m^2 - h \cos^2\beta} \qquad 266.$$

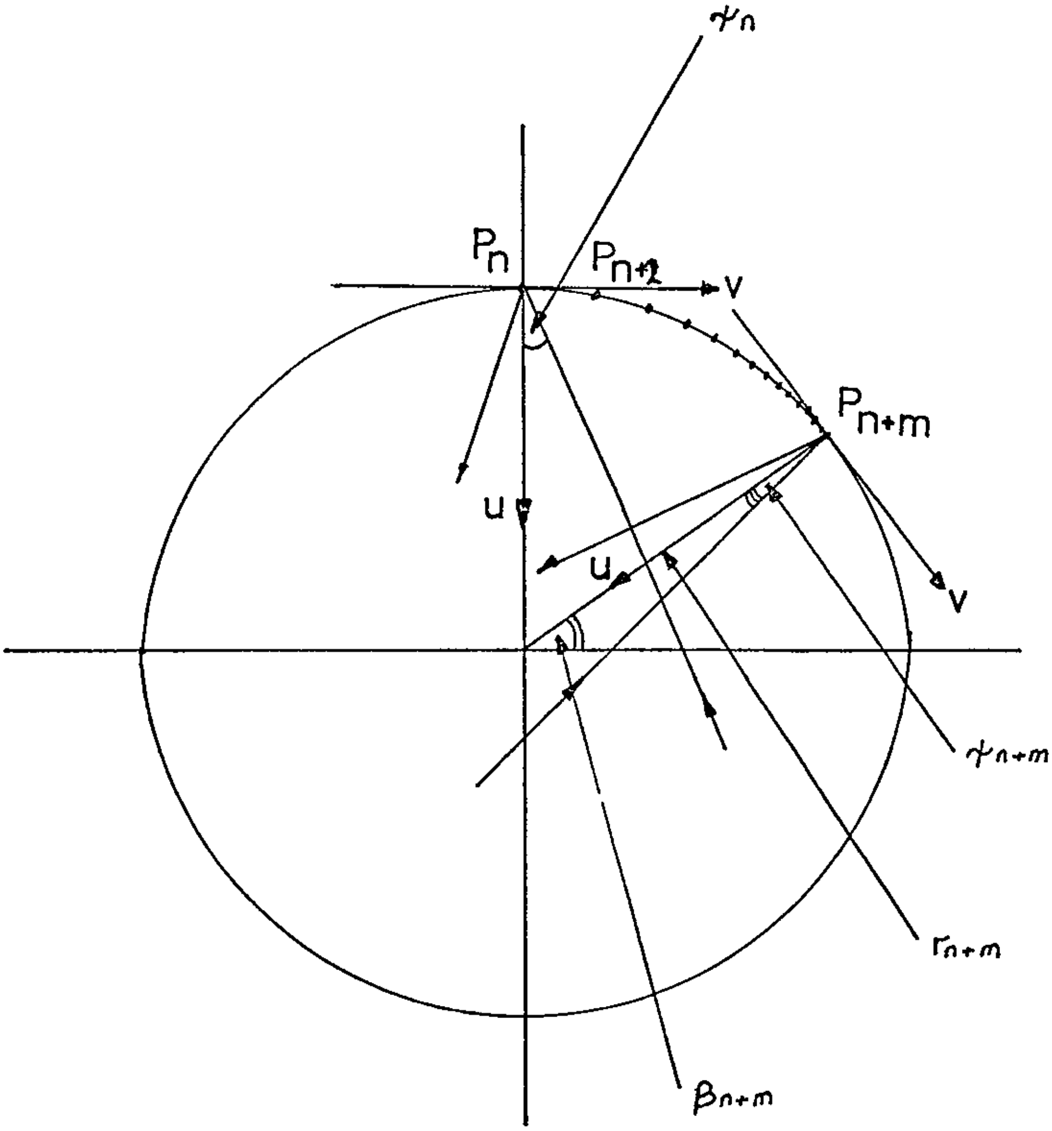


Figure 79. An ellipse, showing rays incident in the positive quadrant of the u,v co-ordinate system.

which, as will be shown, adequately describes the behaviour of the skew plane rays in the elliptic cross section waveguide.

To assist further investigation of Equation (266), χm is expressed in the form

$$\chi m^2 = \sigma^2 h \tag{267.}$$

and Equation (266) becomes

$$\chi = \sqrt{h(\sigma^2 - \cos^2 \beta)} \tag{268.}$$

Consider the point P_n on the ellipse at $\beta = \frac{\pi}{2}$ as shown in Figure 79 with a co-ordinate system u, v with its origin at P_n and with the +ve u axis coinciding with the normal to the reflecting surface at P_n. A skew plane ray in the +u, +v quadrant, incident on the reflecting surface at P_n will have anticlockwise increments in β for each reflection. If χ is small, alternate points of reflection P_n, P_{n+2}, P_{n+4} ... will appear in the same quadrant of the ellipse and have reducing values of β.

If $\sigma < 1$ then according to Equation (268), χ will be zero when $\beta = \cos^{-1}\sigma$ and the angle of incidence will be given by substituting $\chi = 0$ in Equation (252) to give

$$\chi' = -2\gamma_n \tag{269.}$$

The ray will now be incident in the +u, -v quadrant, and successive reflections will result in clockwise increments of β and the alternate points of reflection will have increasing values of β. Since Equation (268) is dependent upon $\cos^2\beta$, similar behaviour will occur for β in the range $\frac{\pi}{2} < \beta < \pi$ and will result in the skew plane ray path oscillating about the y axis of the ellipse between $\beta = \cos^{-1}\sigma$ and $(\pi - \beta) = \cos^{-1}\sigma$.

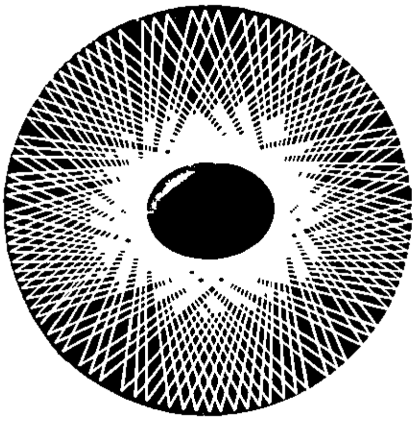


Figure 80b

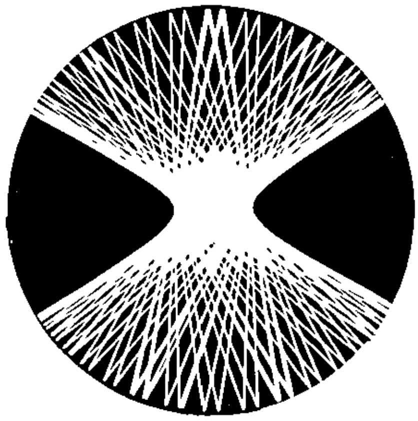


Figure 80a

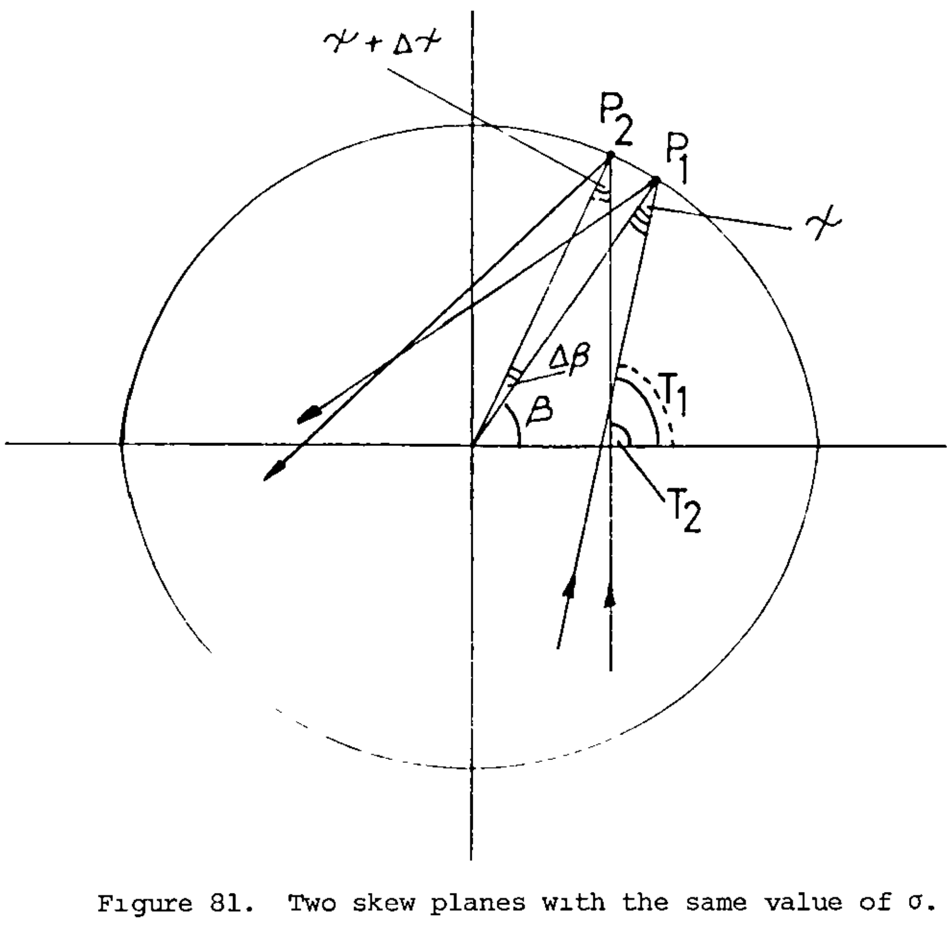


Figure 81. Two skew planes with the same value of σ.

If $\sigma > 1$ the ray will not change quadrants since $\chi > 0$ for all β values and the resultant ray path will be similar to that for a skew plane ray in a circular cross section waveguide. Examples of computer simulated ray paths in elliptic cross section waveguides are shown in Figure 80 where

$\sigma = .8660$ Figure 80a. and $\sigma = 1.5$ Figure 80b.

The caustic shown in Figure 80b is similar to that for a circular cross section except that it has been deformed into a quasi-elliptical figure. The caustic shown in Figure 80a. obviously differs from the circular caustic in that the ray paths cross the centre of the ellipse and the resultant caustic is quasi-hyperbolic in shape.

Before considering the equations for these caustics, the approximation $\frac{a^2}{b^2} \simeq 1$ is considered. With prior knowledge of the experimental results presented later in this chapter, the maximum variation between major and minor axis for the waveguides used in this study is of the order of 2%. This may be expressed by setting

$$a = b + \Delta \qquad 270.$$

then

$$\frac{\Delta}{a} \leqslant 0.02$$

and

$$1 < \frac{a^2}{b^2} < 1.04$$

The corresponding maximum value of h is .04 which gives $\chi m \sim 12^{\circ}$ for $\sigma = 1$ which in turn permits the small angle approximation for χ to be used when $\sigma \lesssim 1$.

In Figure 81, two skew plane rays with the same value for σ are incident at points P_1, P_2 on the ellipse and the points P_1, P_2 have co-ordinates X_1, Y_1, X_2, Y_2 respectively. The tangents of the rays before reflection are T_1 and T_2 and the expressions for X, Y, T are obtained from Figure 81 in terms of the angles χ, $\Delta\chi$ β $\Delta\beta$ and the nominal radius a of the ellipse.

$$X_1 = a \sin\beta \qquad\qquad X_2 = a \sin(\beta + \Delta\beta)$$

$$Y_1 = a \cos\beta \qquad 271a. \qquad Y_2 = a \cos(\beta + \Delta\beta) \qquad 271b.$$

$$T_1 = \tan(\beta + \chi) \qquad\qquad T_2 = \tan(\beta + \Delta\beta + \chi + \Delta\chi)$$

The expression for T_2 contains the term $\Delta\chi$ which may be combined with $\Delta\beta$ using Equation (262) to give

$$T_2 = \tan(\beta + \chi + t\,\Delta\beta) \qquad 272.$$

where

$$t = \left(1 + \frac{h \sin 2\beta}{2\chi}\right) \qquad 273.$$

and the assumption that $\chi_{n+1} \simeq \chi_n$ has been used and the subscript n dropped.

Substituting Equations (271), (272) into (163) gives

$$K_1 = \frac{-a \sin\chi}{\cos(\beta+\chi)}$$

$$K_2 = \frac{-a \sin(\chi+\Delta\beta(t-1))}{\cos(\beta+\chi+ t\Delta\beta)} \qquad 274.$$

$$P = \frac{-\sin t\,\Delta\beta}{\cos(\beta+\chi)\cos(\beta+\chi+t\Delta\beta)}$$

Substituting Equation (274) into (164) and letting $\Delta\beta \to 0$ gives the expressions for the co-ordinates of the caustic

$$x_s = a(\cos\beta - \frac{1}{T}\cos(\beta+\chi)\cos\chi)$$

$$\qquad 275.$$

$$y_s = a(\sin\beta - \frac{1}{T}\sin(\beta+\chi)\cos\chi)$$

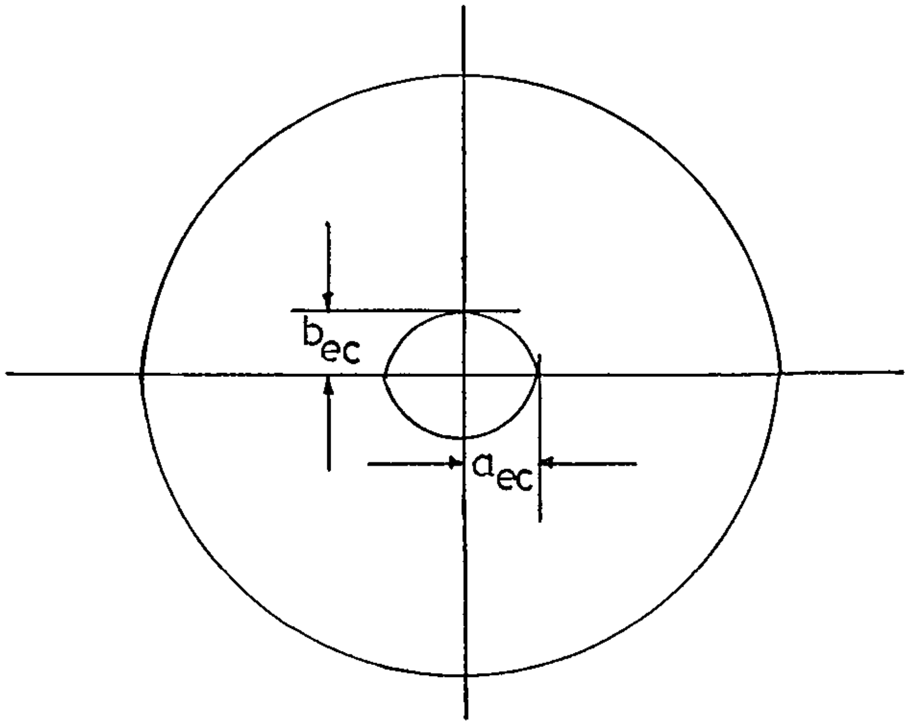


Figure 82a. A quasi-elliptical caustic.

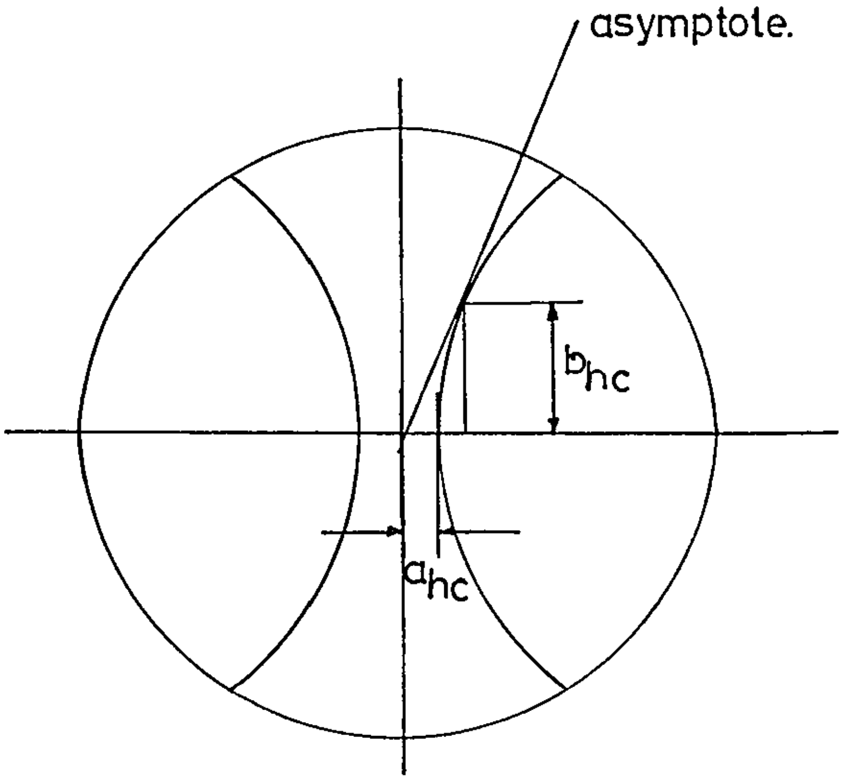


Figure 82b. A quasi-hyperbolic caustic.

and finally expressing Equations (275) in polar co-ordinates using Equation (170a) gives

$$R^2 = x_s^2 + y_s^2$$

$$= a^2(1 + (\frac{1-2t}{t^2}) \cos^2 \chi) \qquad 276.$$

When $t = 1 (a=b)$ Equation (276) reduces to the caustic equation for rays propagating in a circular waveguide as given by Equation (175). Expanding the term $(\frac{1-2t}{t^2})$ using Equation (272) gives

$$(\frac{1-2t}{t^2}) = \frac{-(1 + \frac{h}{\chi} \sin 2\beta]}{(1 + \frac{h}{\chi} \sin 2\beta + \frac{h^2}{4\chi^2} \sin^2 2\beta} \qquad 277.$$

and neglecting the last term of the denominator which will be small compared to the first two terms Equation (277) reduces to -1 which on substitution into Equation (276) again gives Equation (175).

The equation for the caustic shown in Figure 80b is obtained by substituting for χ in Equation (175) using Equation (268) assuming that $\sigma > 1$ which ensures positive χ for all β. Assuming also that χ is small so that $\sin \chi \simeq \chi$ the equation for the caustic for $\sigma > 1$ becomes

$$R^2 = a^2 h(\sigma^2 - \cos^2 \beta) \qquad 278.$$

The maximum and minimum values of R, shown in Figure 82a as a_{ec}, b_{ec} respectively are given by the following expressions

$$a_{ec} = a\sigma \sqrt{h} \qquad (\beta = \frac{\pi}{2}) \qquad 279a.$$

$$b_{ec} = a \sqrt{h(\ ^2 - 1)} \qquad (\beta = 0) \qquad 279b.$$

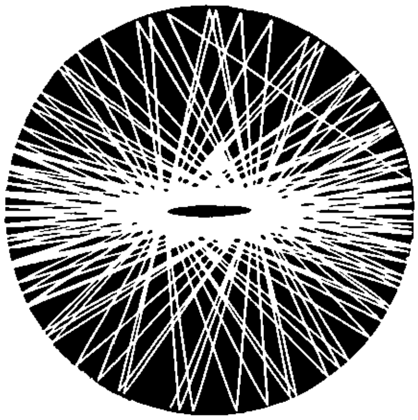


Figure 83a. $\sigma = 1.01$

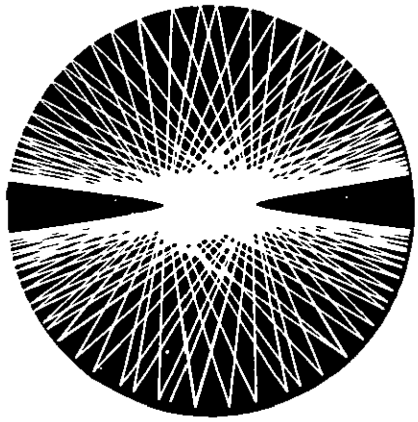


Figure 83b. $\sigma = .99$

Figure 83. Photographs of computer simulations of rays reflected by an elliptical reflector with $a = 1.05$, $b = 1.0$.

If Equation (278) is assumed is assumed to approximately represent an ellipse for $\sigma > 1$ then the quasi-hyperbolic figure given by Equation (278) for $\sigma < 1$ may be assumed to be hyperbolic as shown in Figure 82b. The constants of this hyperbola a_{hc}, b_{hc} are given by

$$a_{hc} = a\sigma\sqrt{h} \qquad (\beta=\tfrac{\pi}{2}) \tag{280a}$$

$$b_{hc} = a\sqrt{h}(1-\sigma^2) \qquad (\beta= \cos^{-1}\sigma) \tag{280b.}$$

The computer simulated caustics for the skew plane rays with σ just above and just below unity are shown in Figures 83a, 83b respectively. Assuming $\sigma \simeq 1$ in Equations (279a) and (280a) then

$$w = 2a_{eh} = 2a_{hc} \tag{281.}$$

where w is the cut-off width of the ellipse and is related to the difference between the vertex radii of the ellipse according to the expression

$$\Delta = \frac{w^2}{8a} \tag{282.}$$

As a numerical example consider the caustic (the edge of the black hole) at the centre of the waveguide cross section shown in Figure 5e to be that due to skew plane rays with $\sigma > 1$, $w \simeq 4.5\mu$ and $a = 23\mu$ to give $\Delta \sim .1\mu$. If the minimum measurable value for w is 1μ in a 50μ core diameter waveguide, the corresponding value for Δ is $.005\mu$ which is .02% of the radius.

This order of sensitivity to variations in the waveguide geometry justifies the comments made in the introduction to this chapter, that the waveguides used for the experiments exhibit asymmetry, since it is unlikely that they are manufactured within a tolerance of $.005\mu$.

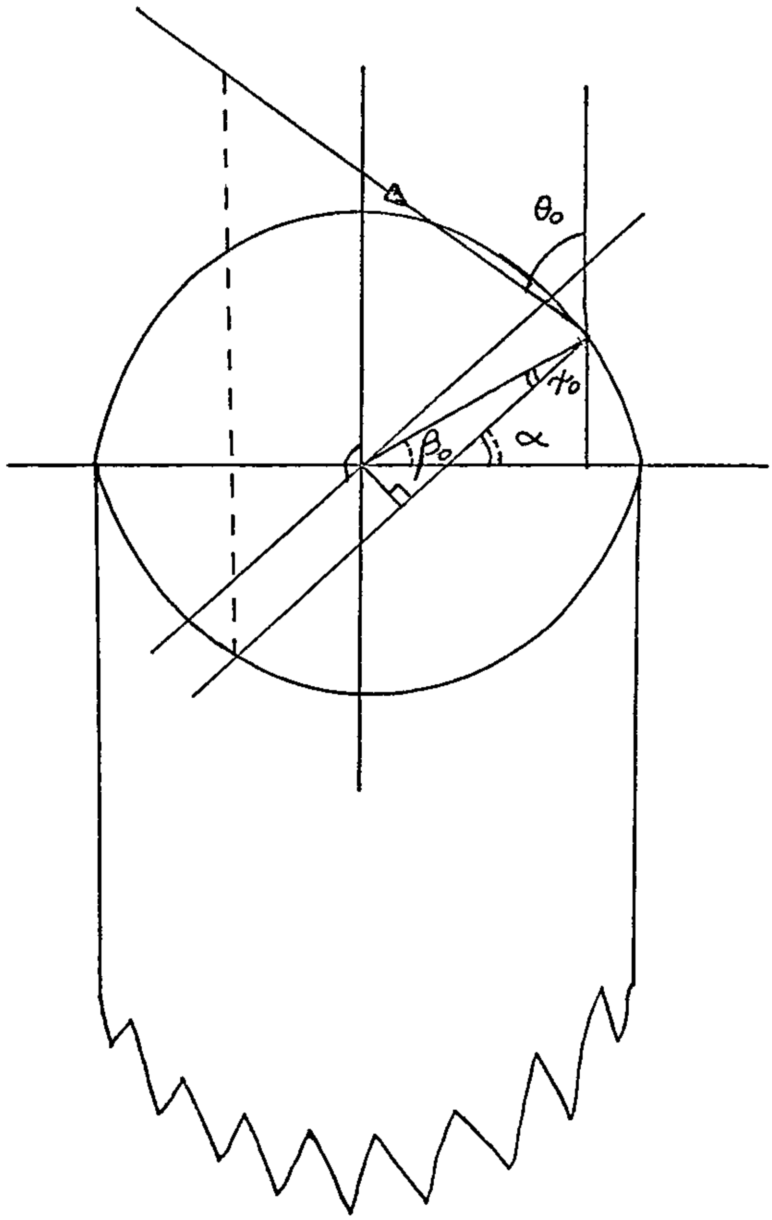


Figure 84. A single ray of a plane wave incident on the entrance aperture of an elliptic cross section waveguide.

Consider a plane wave incident upon the entrance aperture of an elliptic cross section waveguide at angles θ_o, α, as shown in Figure 84. The longitudinal plane of the waveguide selected to define the azimuthal angle α coincides with the major axis.

The skew planes illuminated by this plane wave are identified by their characteristic angles χ_o, β_o in the entrance aperture of the waveguide as shown in Figure 84. Substitution of χ_o, β_o into Equation (268) will give the value of σ for each skew plane, as

$$\sigma = \sqrt{\frac{\chi_o^2}{h} + \cos^2 \beta_o} \qquad 283.$$

Skew planes with $\sigma < 1$ will be called trapped skew planes since they oscillate about the minor axis of the ellipse to produce the quasi-hyperbolic caustics. Skew planes with $\sigma > 1$ will be called non trapping skew planes.

In Figure 84, χ_o, β_o are related by the expression

$$\chi_o = \alpha - \beta_o \qquad 284.$$

and the relation between χ_o and β_o for trapped skew planes is found by substituting $\sigma < 1$ into Equation (283) to give

$$\chi_o < \left| \sqrt{h} \sin \beta_o \right| \qquad 285.$$

Substituting for χ_o in Equation (285) using the right hand side of Equation (284), gives the range of β_o corresponding to trapped skew planes as

$$\beta_o = \alpha \pm \sqrt{h} \sin \beta_o \qquad 286.$$

and the corresponding values for χ_o are obtained from Equation (284) and are

$$\chi_o = 0 \qquad \text{when} \qquad \beta_o = \alpha \qquad 287.$$

$$\chi_o = \mp \sqrt{h}\,\sin\beta_o \qquad \text{when} \qquad \beta_o = \alpha \pm \sqrt{h}\,\sin\beta_o \qquad 288.$$

Substituting for χ_o, β_o in Equation (283) using Equations (287) and (288) gives the range of σ in terms of the azimuth angle of the incident plane wave as

$$\cos\alpha < \sigma < 1 \qquad (0 < \chi_o < \sqrt{h}\,\sin\beta_o) \qquad 289.$$

The gradient of the asymptote to the quasi-hyperbolic caustic, denoted in Figure 82b as β_{co} is given by

$$\tan\beta_{co} = \frac{b_{hc}}{a_{hc}} = \frac{\sqrt{1-\sigma^2}}{\sigma} \qquad 290.$$

which on substitution for σ from Equation (289) gives the range of the cut-off angle β_{co}

$$0 \lesssim \beta_{co} \lesssim \alpha \qquad (\sqrt{h}\,\sin\beta_o > \chi_o > 0) \qquad 291.$$

When $\alpha = 0$, Equation (289) gives $\sigma = 1$ indicating that no trapped skew planes ($\sigma < 1$) are illuminated by a plane wave at this azimuth angle of incidence.

The relation between χ_o and β_o for non trapped skew planes is found by substituting $\sigma > 1$ into Equation (283) to give

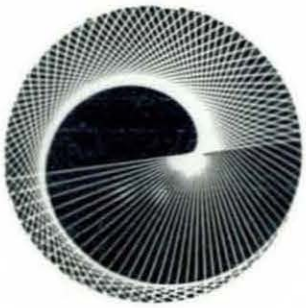
$$\chi_o > \left|\sqrt{h}\,\sin\beta_o\right| \qquad 292.$$

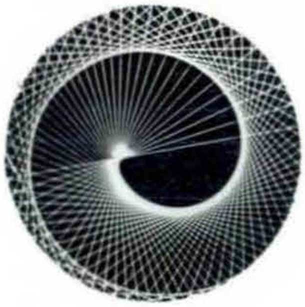
and the initial positions of these skew planes are given by Equation (284) as

$$\beta_o = \alpha \pm \chi_o \qquad 293.$$

The significant difference between the trapped and non trapped skew plane dependence upon the azimuth angle of incidence α, is that

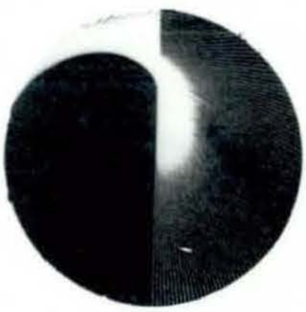
m = 1 m = 2



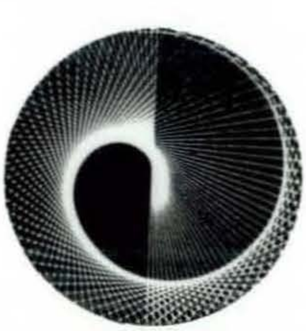


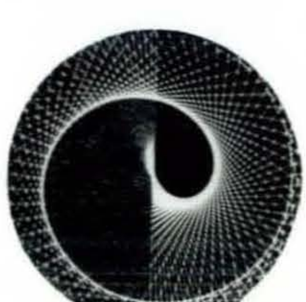
m = 3 m = 4

$\alpha = 0$



m = 1 m = 2





m = 3 m = 4

$\alpha = \pi/2$

Figure 85. Photographs of computer simulation of rays reflected by an elliptical reflector, a = 1.05, b = 1.0.

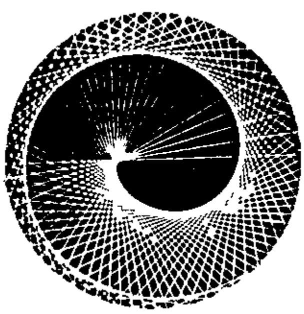
only certain trapped skew planes are illuminated at each value of α (Equation (289)) whereas the nontrapped skew planes are illuminated independently of α. The α dependence of the nontrapped skew planes only appears in the expression for the initial position of the skew planes, Equation (293).

A mathematical description of the caustics produced by the incident plane wave as a function of the number of reflections is complicated by the transition from the quasi hyperbolic caustics for $\sigma < 1$ to the quasi elliptic caustics for $\sigma > 1$. It has not been possible to produce a useful mathematical expression to describe this process and the following qualitative description of the caustics is based upon the results of computer simulation of the multiple reflections of skew planes.

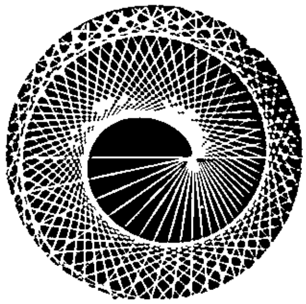
Figure 85 shows computer simulations of the caustics produced by skew planes reflected by an elliptical cross section reflector ($a = 1.05$ $b = 1.0$) for $m = 1$ to 4 and with azimuthal angles of incidence $\alpha = 0$ and $\frac{\pi}{2}$. Comparing these caustics with those for the circular cross section reflector shown in Figure 57, suggests that the low m caustics are insensitive to variations in the cross sectional geometry of the reflector.

However, as m is increased the effects of the ellipticity and the dependence upon the azimuthal angle of incidence become apparent. In Figure 86 the caustics for $m = 6,7, 10,11$ are illustrated for the elliptical cross section reflector with $\alpha = 0, \frac{\pi}{2}$.

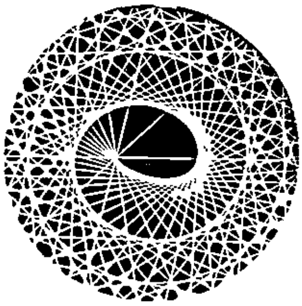
The circular reflector caustics form the characteristic spirals which originate on the y axis as given by Equation (171), and wind up as m is increased. The caustics produced in the elliptical reflector



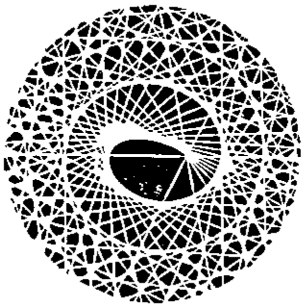
m = 4



m = 5

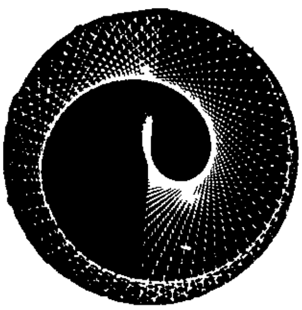


m = 8

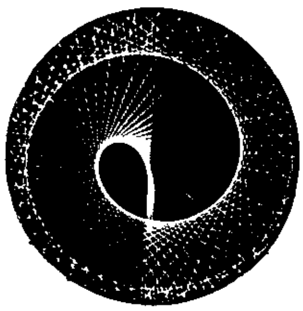


m = 9

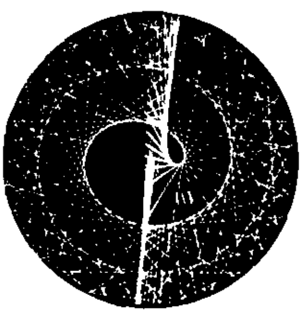
$\alpha = 0$



m = 4

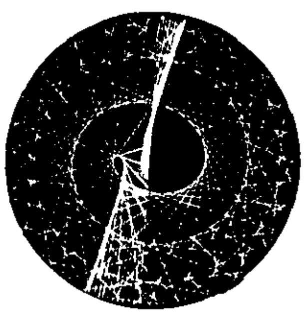


m = 5



m = 8

$\alpha = \ /2$



m = 9

Figure 86. Photographs of computer simulation of rays reflected by an elliptical reflector, a = 1.05, b = 1.0.

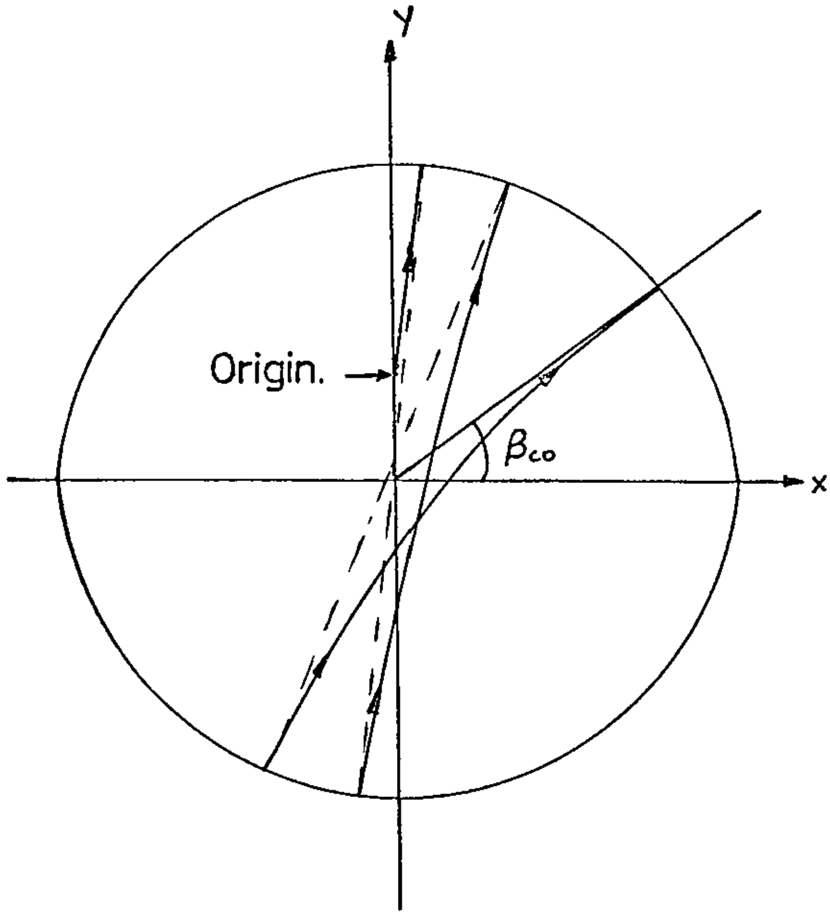
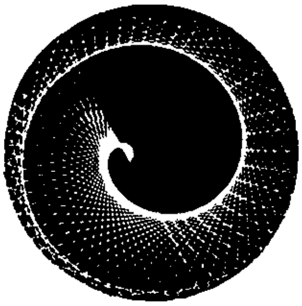
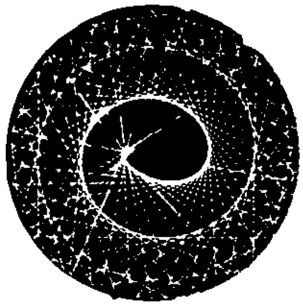


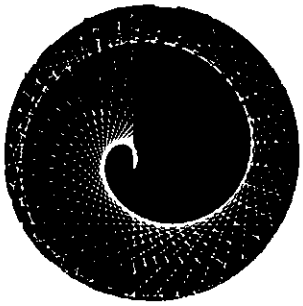
Figure 87. A typical trapped skew plane caustic.

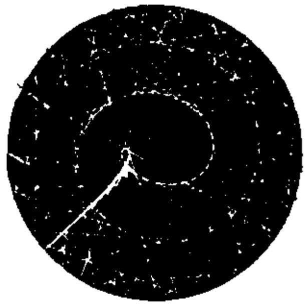




m = 4 m = 8

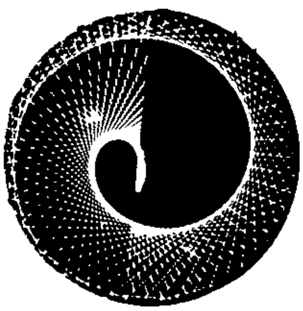
$\alpha = 22.5^{\circ}$

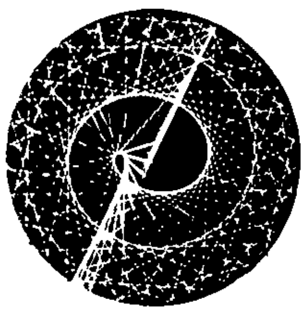




m = 4 m = 8

$\alpha = 45^{\circ}$





m = 4 m = 8

$\alpha = 67.5^{\circ}$

Figure 88. Photographs of computer simulations of rays reflected by an elliptical reflector, a = 1.05, b = 1.0.

when $\alpha = 0$ form what may be called quasi elliptical spirals which originate on the x axis at $x = a\sqrt{h}$, where this value is obtained from Equation (279a) by substituting $\sigma = 1$. These spirals correspond to the non trapped skew planes, and also 'wind up' as m is increased.

Similar spirals are obtained for the elliptical reflector when $\alpha = \frac{\pi}{2}$ since nontrapped skew planes are excited independently of the value of α. The trapped skew planes excited when $\alpha = \frac{\pi}{2}$ form a caustic which originates on the y axis and follows a path of the form shown in Figure 87. This caustic terminates at the point of origin of the non trapped skew planes since both these conditions correspond to $\sigma = 1$.

The cut off angle β_{co} shown in Figure 87 is reduced as m' is increased until in the limit of large m, $\beta_{co} = 0$ corresponding to the value of β_{co} given by Equation (290) when $\sigma = 1$. The origin of the trapped skew plane caustic depends upon the azimuthal angle of incidence and for $\alpha \neq \frac{\pi}{2}$ also depends upon m. This is demonstrated in Figure 88 where the computer simulated caustics for $\alpha = \frac{3\pi}{8}, \frac{\pi}{4}, \frac{\pi}{8}$ and m = 6, 10 are given. When m is large the maximum angle of the caustic origin is given by the upper limit of Equation (291)

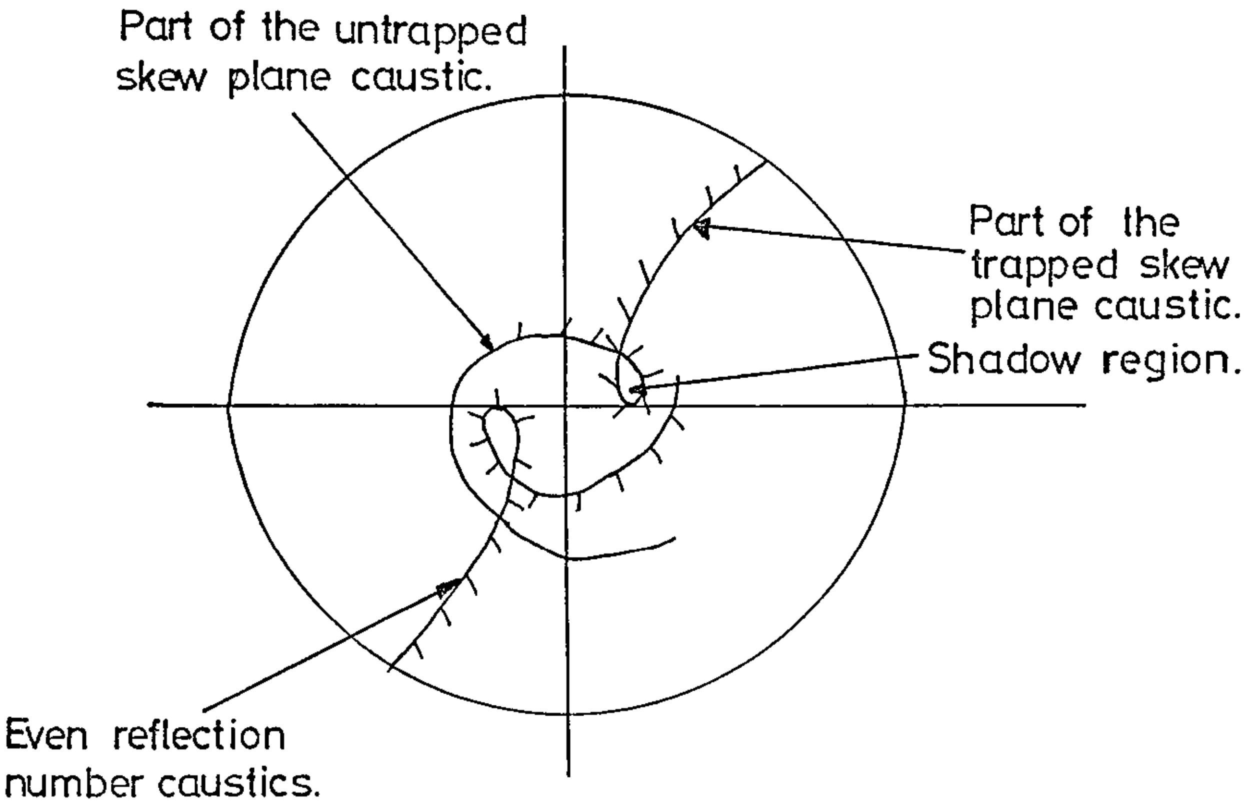


Figure 89. The transition from the quasi hyperbolic caustic to the quasi elliptical caustic after a small number of reflections.

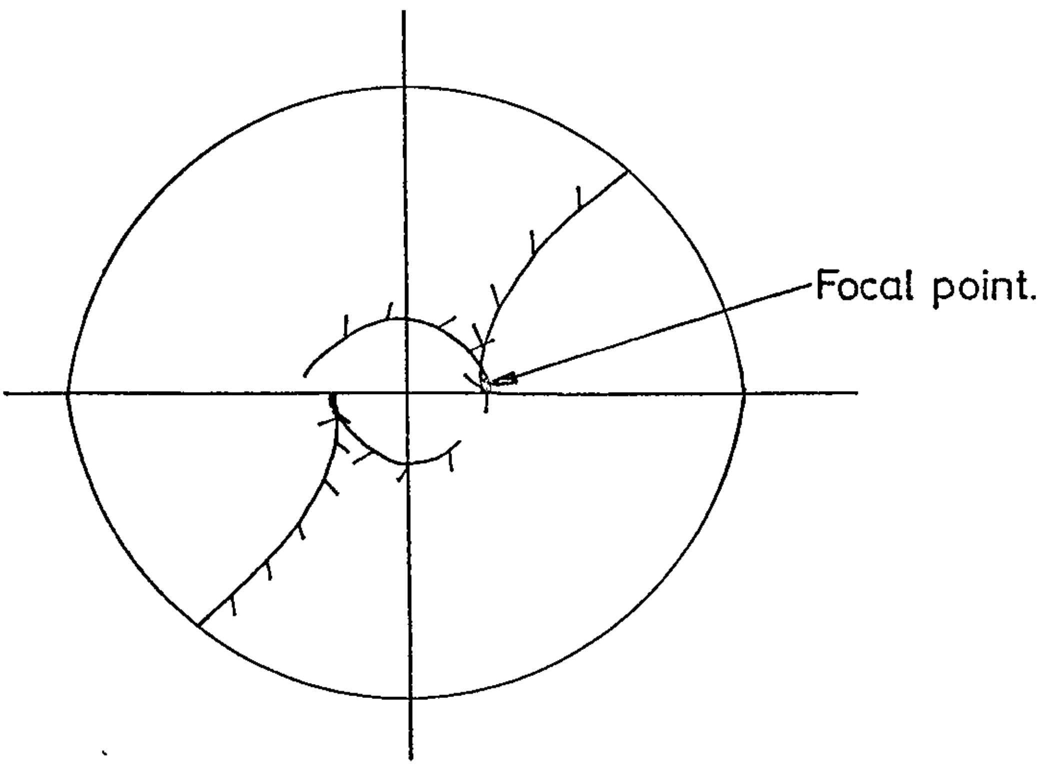


Figure 90. The transition caustic after a high number of reflections.

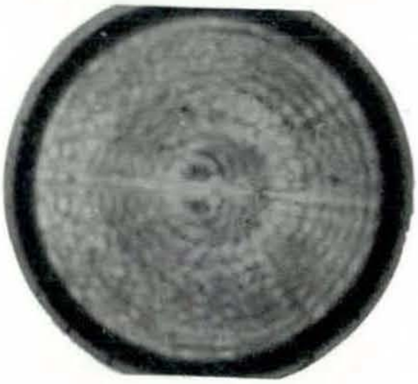


Figure 91.

Figure 92.

Figure 93.

Figure 94.

6.4.2 Experimental Results

A 5.75 mm length of fibre bundle was illuminated with white light at an axial angle of 10^{o} to obtain the caustic shown in Figure 91. The caustic appears to be similar to that obtained by computer simulation shown in Figure 86 where $\alpha = \frac{\pi}{2}$, m = 8. In figure 91, m ≃ 8 and the effects of increasing m are shown in Figures 92 and 93 where the angles of incidence are 20^{o} and 30^{o} respectively. In Figure 92, m ≃ 20 and although the caustic is still visible at the centre of the fibre, its form is less distinct. In Figure 93 where m ≃ 35 the centre caustic is indistinct and this result yields no useful information about the behaviour of rays in elliptic cross section waveguides after many reflections.

The behaviour of the caustic as it changes from a trapped to an untrapped mode has a significant feature which was suggested by the observation of high intensity spots at positions close to the centre of fibres as shown in Figure 94. Figure 94 was obtained from a 40 cm length of fibre illuminated with white light at an axial angle of 20^{o} giving m ≃ 170. As the caustic changes from the trapped quasi-hyperbolic form to the untrapped quasi-elliptical form, the caustic turns to form a tangent to the major axis of the ellipse as shown in Figure 89. In so doing a shadow region is formed which is illustrated experimentally in Figure 91. As the number of reflections is increased the shadow region is reduced in dimensions until the caustic turns through an angle of 270^{o} at a single point adjacent to the major axis of the ellipse. Because of the diffraction which occurs at a caustic this single turning point appears as a bright spot of light and may be considered as the focal point of skew rays with $\sigma \simeq 1$. Its position corresponds to a_{ec} or a_{hc} along the major axis of the ellipse, and for rays with $+\phi$ appears in the 1st and 3rd quadrants for alternate reflections, and in the 2nd and 4th quadrants for rays launched with $-\phi$. In the experimental result shown in Figure 94, spots appear in the 2nd and 4th quadrants simultaneously because the divergence of the rays after 170 reflections results in numerous components with adjacent reflection numbers.

Further experimental examples of the trapped quasi hyperbolic caustics are shown in Figures 5 and 7. In each of these photographs the form of the caustic is dependent upon the azimuthal angle of incidence and this is a demonstration of the range of trapped skew planes excited by plane waves incident at particular azimuthal angles. (Equations 287, 288). In particular when $\alpha \simeq 270.0$ as in Figure 5e, no quasi hyperbolic caustics are in evidence, excepting the bright spots of the focal points of skew planes with $\sigma \simeq 1$.

There is also a shadow region in evidence in Figure 5e, which is not suggested by the theory of the untrapped caustics, since all the untrapped skew planes are excited at all azimuthal angles of incidence. It may be that the cross section of the waveguide is not exactly elliptical and that the skew planes with $\sigma \simeq 1$ suffer a misalignment at points within the fibres length. Alternatively, the bending of the fibre bundle during the experiment may cause the centre shadow effect. These two possibilities are considered in the following sections and Chapter 7.

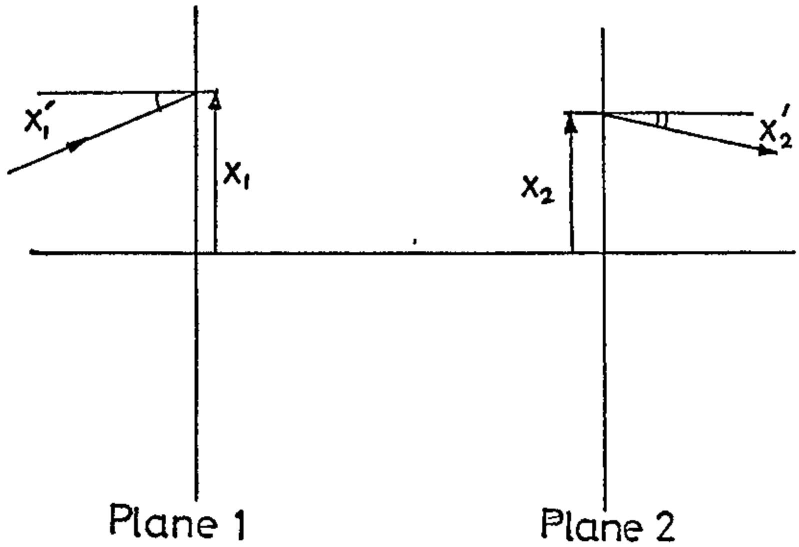


Figure 95. The characterisation of rays in an optical system.

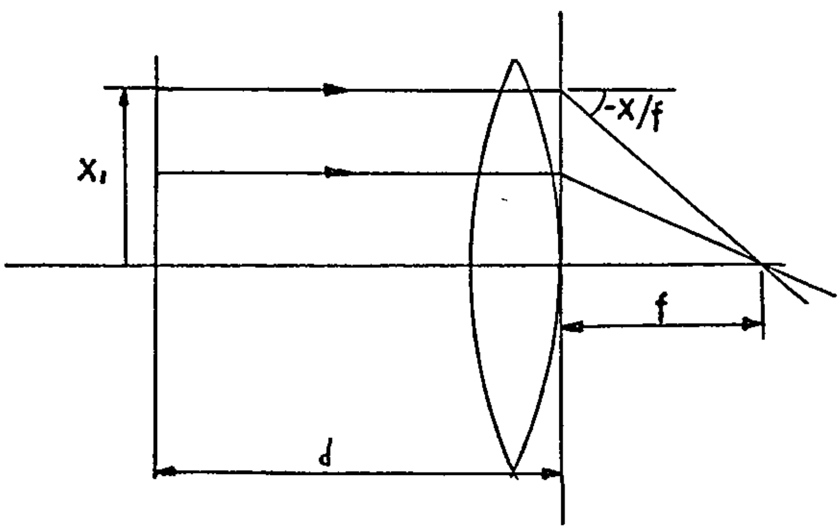


Figure 96. A thin lens system.

6.5 Representation of the Elliptical Waveguide as a Periodic Sequence of Lenses.

Another author[20] has noted the similarity between the theory of resonating cavities and the theory of propagation in elliptical waveguides. He adapted the resonator theory to predict the waveguide modes observed when an elliptical dielectric waveguide is illuminated by a monochromatic source at specific azimuthal angles of incidence.

In this section the trapping condition for skew plane rays with X is derived from the geometrical theory of resonating cavities, where the multiple reflection of a light ray is represented by a ray passing through a periodic lens sequence.

The path of a ray through an optical system is characterised[50] by its distance x from the optical axis and its slope x' both measured at the input and output planes of the system. In Figure 95 x_1, x_1' represents the input quantities measured at plane 1 and x_2, x_2' are the output quantities measured at plane 2. When a ray has x_1, x_2 small and its slopes x_1', x_2', may be equated to their angles ($x_1' = \mathrm{TAN}^{-1} x_1'$) the rays are said to be confined to the paraxial region of the optical system. (31 Sec.4.9.

The output quantities x_2, x_2' of paraxial rays are linearly related to their input quantities and this relationship may be expressed in the following matrix form

$$\begin{vmatrix} x_2 \\ x_2' \end{vmatrix} = \begin{vmatrix} A & B \\ C & D \end{vmatrix} \begin{vmatrix} x_1 \\ x_1' \end{vmatrix} \qquad 294.$$

The values of the matrix $\begin{vmatrix} AB \\ CD \end{vmatrix}$, known as the ray transfer matrix, are given in reference (50) for various simple optical systems. The system of interest here is shown in Figure 96 and consists of a thin lens, focal length f, with the input plane a distance d from the output plane.

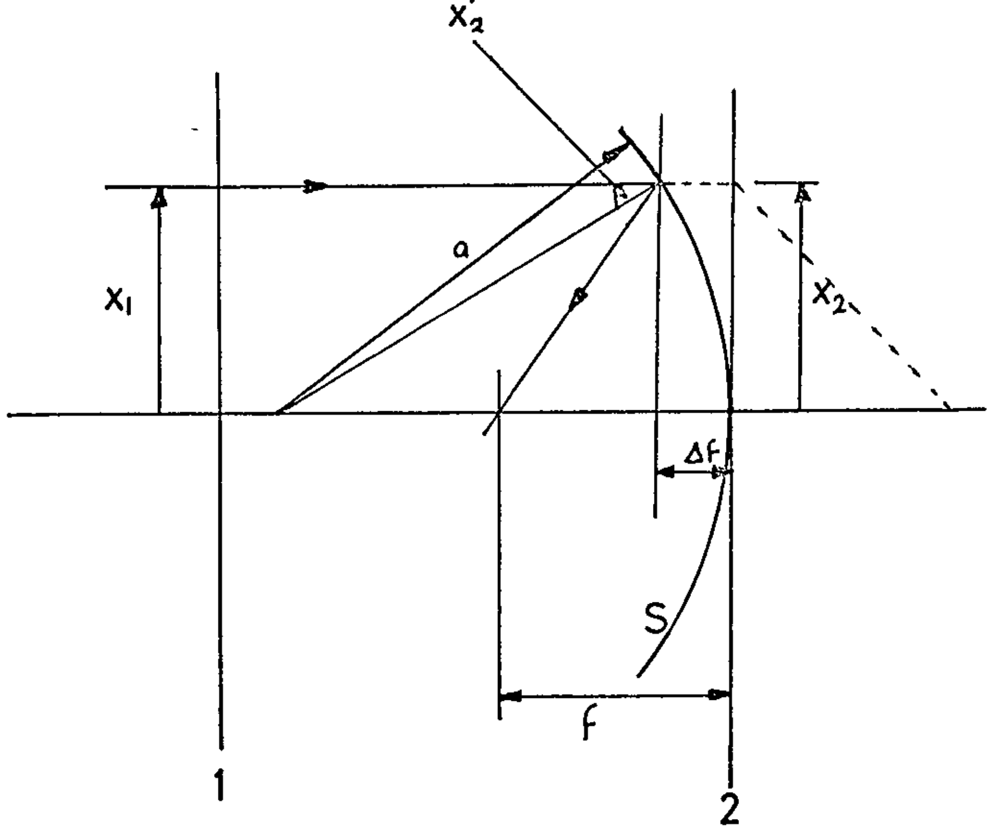


Figure 97.

Its ray transfer matrix is given by reference (50) as

$$\begin{vmatrix} 1 & d \\ -\frac{1}{f} & 1-\frac{d}{f} \end{vmatrix} \qquad 295.$$

which gives the following equations for the output quantities

$$x_2 = x_1 + dx_1' \qquad 296.$$

$$x_2' = x_1' - \frac{x_1}{f} - \frac{dx_1'}{f} \qquad 297.$$

A set of rays parallel ($x_1' = 0$) to the optical axis at plane 1 with various values of x_1 will have

$$x_2 = x_1 \qquad 298.$$

$$x_2' = -\frac{x_1}{f} \qquad 299.$$

at plane 2 to give the ray paths shown in Figure 96. The focal length of the system is clearly the distance from the output plane of the system at which a set of parallel input rays converges to a focus.

Consider a ray parallel to the axis of the optical system shown in Figure 97 where S is a section of a cylindrical reflector, radius a. The ray has an angle of incidence at the reflector of $\frac{\mathrm{TAN}^{-1} x_2'}{2}$ and the reflected ray crosses the optical axis a distance -f from plane 2. From Figure 97.

$$x_2' \sim \frac{x_2}{(f-\Delta f)} \qquad 300.$$

$$\sin\left(\frac{\mathrm{TAN}^{-1} x_2'}{2}\right) = \frac{x_1}{a} \qquad 301.$$

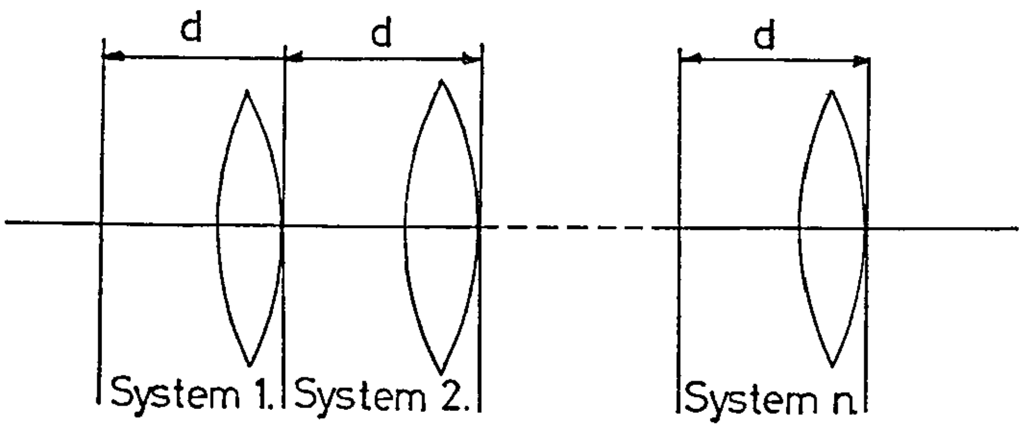


Figure 98. A periodic lens sequence.

$$\Delta f = a(1 - \cos(\mathrm{TAN}^{-1} x_2')) \qquad 302.$$

$$x_1 = x_2 \qquad 303.$$

Using the small angle approximations ($\sin\theta \sim \theta$, $\mathrm{TAN}\,\theta \sim \theta$) and neglecting Δf then from Equations (300) and (303)

$$f = \frac{x_1}{x_2'} \qquad 304.$$

and from Equation (301) and (304)

$$f = \frac{a}{2} \qquad 305.$$

It appears that rays within the paraxial regions of the systems shown in Figures 96 and 97 will follow similar paths if the -f of Figure 97 is replaced by +f.

The path of a paraxial ray making N reflections in an optical system consisting of concentric cylindrical reflectors, each of radius a, may therefore be represented by the path of a paraxial ray passing through N successive systems of the type shown in Figure 96 as shown in Figure 98.

The relationship between the input ray quantities x_o, x_o' and the ray quantities x_n, x_n' at the n^{th} plane is given by

$$\begin{vmatrix} x_n \\ x_n' \end{vmatrix} = \begin{vmatrix} A & B \\ C & D \end{vmatrix}^n \begin{vmatrix} x_o \\ x_o' \end{vmatrix} \qquad 306.$$

The standard textbook treatment of ray propagation in lens sequences suggests that if a beam of rays passirg through such a system is confined to within the paraxial region of the system axis, the periodic sequence

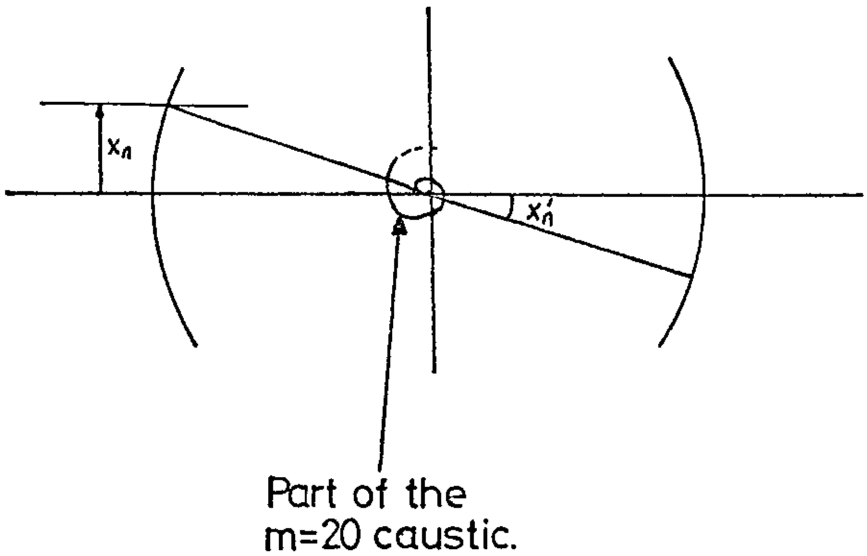


Figure 99. A section of the m = 20 caustic shown in Figure 69, to illustrate the divergence of the rays which form this caustic.

is said to be stable and its matrix components satisfy the condition

$$-1 < \tfrac{1}{2}(A + D) < 1 \qquad 307.$$

Substitution of the relevant values of matrix 295 into Equation (307) gives

$$-1 < (1 - \frac{d}{2f}) < 1 \qquad 308.$$

which may be recast into the form

$$0 < (1 - \frac{d}{4f}) < 1 \qquad 309.$$

The concentric reflector system (d = 2a) gives

$$(1 - \frac{d}{4f}) = 0 \qquad 310.$$

since Equation (305) gives $f = \frac{a}{2}$ for paraxial rays and this suggests that all paraxial rays are stable in a concentric system. This is not so because the exact expression for f given by Equations (300) to (303) gives a value $f = \frac{a}{2}$ for $x_1 = x_1' = 0$ only, all other values of x_1 and x_1' give $f < \frac{a}{2}$.

If $f < \frac{a}{2}$ then

$$(1 - \frac{d}{4f}) < 0 \qquad 311.$$

and the system is unstable and the rays for which $f < \frac{a}{2}$ are not trapped and will rapidly diverge away from the system axis.

The divergence of rays with $\chi > 0$ propagating in a concentric system is also predicted by the reflection number dependent caustic given by Equations (170a-b). In Figure 99 a section of the caustic shown in Figure 59 (number of reflections m = 20) is shown together with the paraxial sections of the reflecting surfaces. When m is large

(>10) the caustic equations may be replaced by the approximations given by Equations (175) and (176). Assuming that a ray is tangent to the caustic at its mid point (which is established by Equation (175)), the limit of paraxial trapped rays shown in Figure 99 as x'_n will be given by

$$x'_n = \text{TAN } 2m\chi \tag{312.}$$

which clearly falls to zero as m is increased, for a fixed value of χ.

Consider now the case when the spacing d between the reflectors is reduced by an amount 2Δ so that

$$d = 2a - 2\Delta \tag{313.}$$

To satisfy the stability condition the focal length of trapped rays must lie in the range

$$\frac{a - \Delta}{2} < f < \frac{a}{2} \tag{314.}$$

The locii of focus for a cylindrical reflector is given by the caustic Equations (169a-b) with m = 1, and by reference to Figure 56 the focal length of the reflector will be given by

$$f = a - y_s \tag{315.}$$

which on substitution for y_s using Equation (169b) becomes

$$f = a(1 - \tfrac{1}{2}(\cos 2\chi \cos \chi) - \sin 2\chi \sin \chi) \tag{316.}$$

After algebraic manipulation and substitution for the double angles, Equation (316) becomes

$$f = \frac{a}{2} \cos \chi - a \cos \chi \sin^2 \chi + 2a \sin^2 \frac{\chi}{2} \tag{317.}$$

The focal length of the outer trapped ray is given by the lower limit of Equation (314) which may be equated to the above expression for f to give an expression for the limit of trapped rays in terms of χ. Taking Equation (317) as $f = \frac{a}{2} \cos \chi$ (accurate to within 3% at $\chi = 10^{o}$) and denoting the cut off value of χ as χ_m then

$$\frac{a}{2} \cos \chi_m = \frac{a-\Delta}{2} \qquad 318.$$

Letting $a-\Delta = b$ where a,b correspond to the constants of the ellipse shown in Figure 77 then from Equation (318)

$$\sin^2 \chi_m = \frac{a^2}{b^2} h \qquad 319.$$

Although this expression shows agreement with Equation (267) (with $\sigma = 1 \frac{a^2}{b^2} \simeq 1$) the paraxial lens sequence theory cannot tolerate the cut-off value of $x_n' = \frac{\pi}{2}$ which is encountered in the analysis of ray propagation vetween elliptical reflectors given in Section 6.4.

This is demonstrated by calculating the cut-off angle x_n' using the rigorous expression for f for the cylindrical reflector given by Equation (317). When $\frac{\Delta}{a} = 0.02$ Equation (319) gives $x_m = 12^{o}$ corresponding to $x_n' = \frac{\pi}{2}$, whereas the rigorous analysis gives $\chi_m \sim 6^{o}$ to give a corresponding value for $x_n' = 25^{o}$ which lies within the paraxial region.

However, both theorys demonstrate that other skew rays of light as well as those with $\chi = 0$ are trapped when the spacing between the cylindrical reflectors is less than twice the radius of curvature. This feature of the lens sequence theory is used in the next section to investigate the possibility of rays becoming trapped along low order stationary skew plane paths other than that considered here.

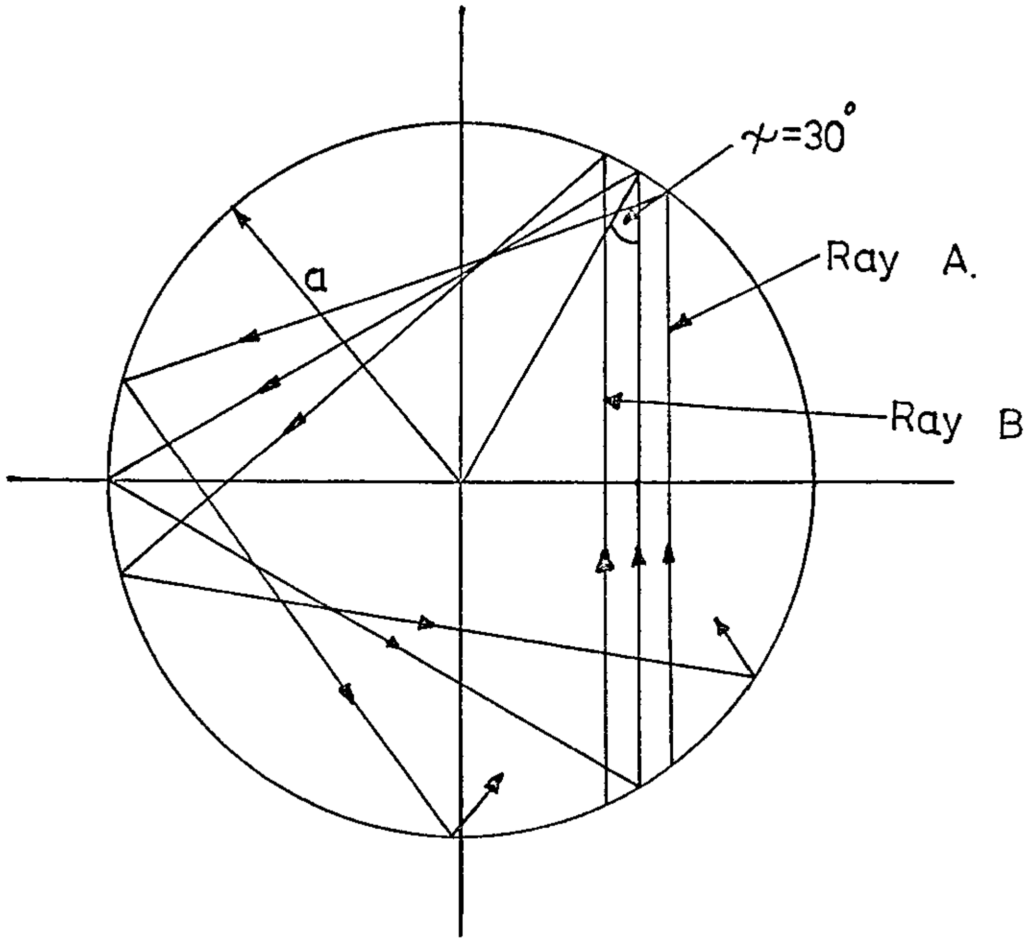


Figure 100. Ray paths in a circular cross section reflector.

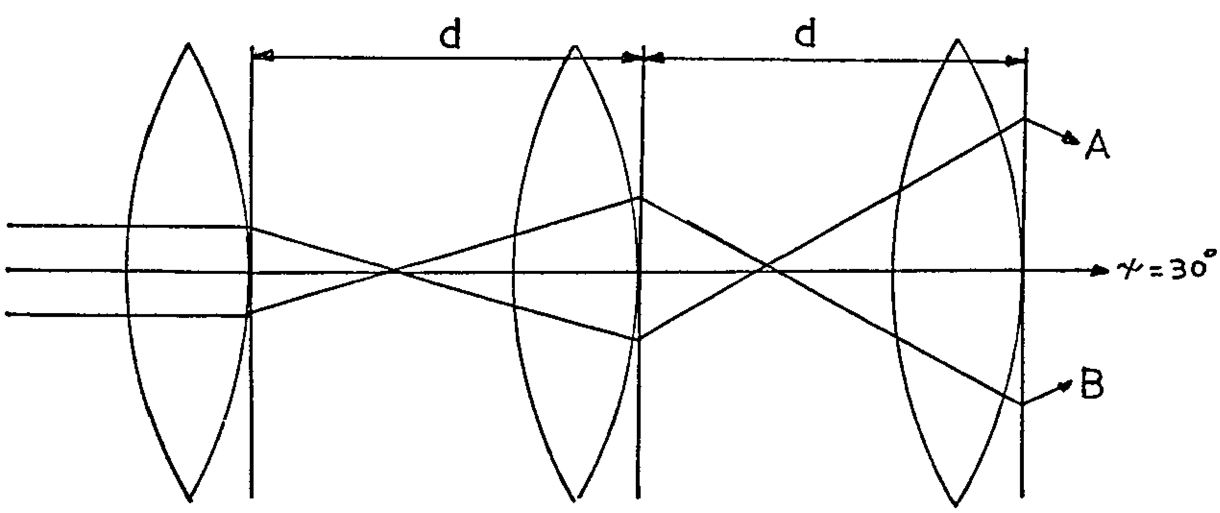


Figure 101. The lens sequence representation of ray paths in a circular cross section reflector.

6.6 Higher Order Trapping Modes in Elliptical Cross Section Waveguides

It has been shown in the earlier sections of this chapter that the high intensity patterns which appear about a diameter of the waveguides are accounted for by the ellipticity of the waveguides which causes a trapping condition to occur. The original patterns described in Chapter 1 also exhibited variations of intensity about other stationary skew plane paths of a cylindrical waveguide.

In this section the patterns which appear in positions corresponding to the $\chi = 30^{o}$ stationary skew plane ($p = 1$, $m = 3$, Figure 49) will be analysed under the assumption that the waveguide has an elliptical cross section as described in Section 6.4.

The approximation used in that section that $\sin \chi \simeq \chi$, is not acceptable for $\chi = 30^{o}$ and it was also found by computer simulation that the β dependent change in χ due to the offset centre of curvature of the ellipse (expressed by Equation (262)) was of the same order as that due to the variations in "radius" of the ellipse (Equation (256)).

A ray analysis similar to that given in Section 6.4 would therefore appear to be complex for higher order skew planes. A visual examination of the patterns suggests that since little fine detail is in evidence the lens sequence analysis will provide sufficient information as to the conditions for the appearance of the higher order skew ray patterns.

Consider first the application of the lens sequence theory to the $\chi = 30^{o}$ stationary skew ray whose path in a circular cross section waveguide is shown in Figure 100 together with the paths of two other skew rays. Ray A has $\chi > 30^{o}$ and ray B has $\chi < 30^{o}$. The corresponding sequence is shown in Figure 101 where the spacing d between the lenses

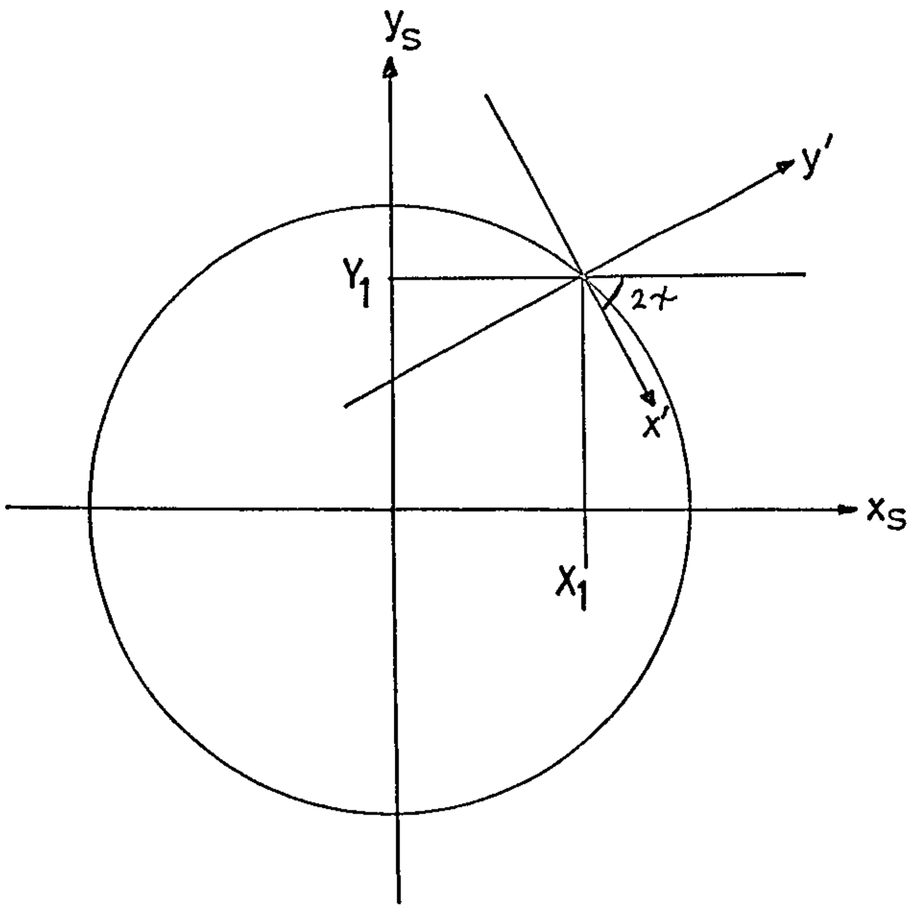


Figure 102. The origin shift for the derivation of the normalised focal length of a circular reflector.

is given by

$$d = 2a \cos 30^{o} \qquad 320.$$

where a is the nominal radius of the waveguide core.

A normalised expression for the focal length of a skew ray reflected by a cylindrical reflector may be obtained from the caustic equations (169a-b). With reference to Figure 102 the origin of the co-ordinate system x_s y_s of the caustic equations is shifted to X_1Y_1 and the axis turned through an angle of -2χ to give new co-ordinates x'y'. By co-ordinate geometry

$$x' = (x_s - X_1) \cos 2\chi - (y_s - Y_1) \sin 2\chi \qquad 321a.$$

$$y' = (y_s - Y_1) \cos 2\chi + (x_s - X_1) \sin 2\chi \qquad 321b.$$

Substituting for X_1Y_1 using Equation (166a) with m = 1 and for x_s,y_s using Equation (169a-b) also with m = 1 gives

$$x' = 0 \qquad 322a.$$

$$y' = -\frac{a}{2} \cos \chi \qquad 322b.$$

where y' is now the focal length $f(\chi)$ of each skew ray measured along its own path.

The spacing d between the lens of Figure 101 is given by Equation (320) as

$$d = 4f(30) \qquad 323.$$

and skew rays with $\chi > 30^{o}$ show the same divergence away from the $\chi = 30^{o}$ path as do the skew rays with $\chi > 0^{o}$ away from the $\chi = 0^{o}$ path since neither set of rays are able to satisfy the stability condition expressed by Equation (309). Unfortunately skew rays with $\chi < 30^{o}$,

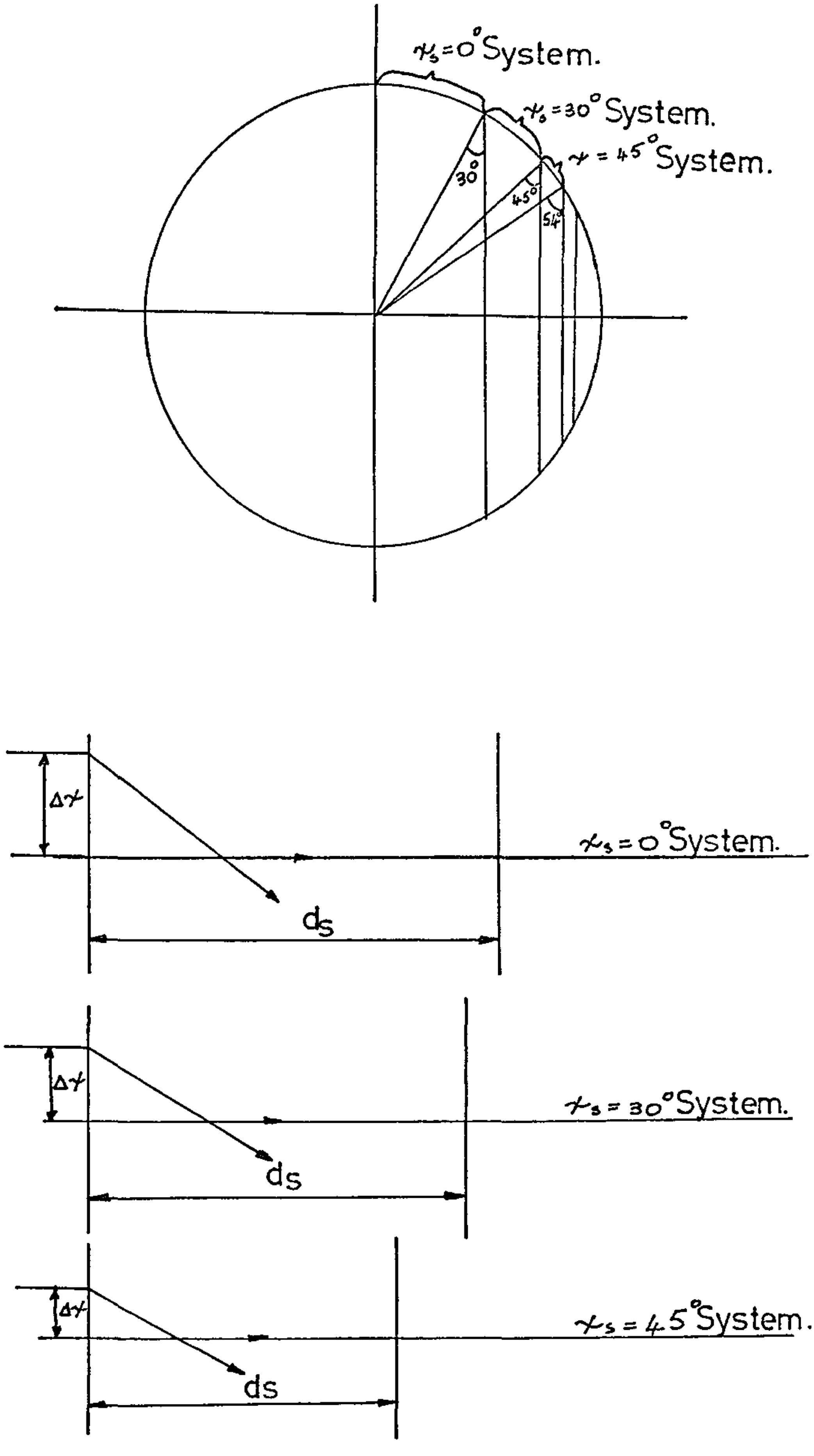


Figure 103. Examples of stationary skew ray lens sequences.

have focal lengths which do satisfy this stability condition and should therefore be trapped within the paraxial region of the $\chi = 30^{\circ}$ skew ray path, also diverge away from this path.

This difficulty does not arise with the $\chi = 0$ system of lenses since there are obviously no rays with $\chi < 0^{\circ}$, and it will be resolved by only including, in each lens sequence representation, rays with $\chi > \chi_s$ where χ_s is the stationary skew ray which forms the axis of the equivalent lens sequence.

The propagation of rays in a cylindrical waveguide may therefore be represented by their propagation through a series of lens sequences as shown in Figure 103, where the range of skew rays initially within each lens sequence, due to an incident plane wave, will depend upon the number of lens sequences chosen to represent the waveguide.

The values of χ_s for various combinations of p, $\overline{m}$ of equations (149), (150) are shown in Table 9 together with the range $\Delta\chi$ of χ in each lens sequence, where $\Delta\chi$ is given by

$$\Delta\chi = \chi_s(p,\overline{m}) - \chi_s(p,\overline{m} + 1) \qquad 324.$$

The lens spacing for all the lens sequences is given by

$$d_s = 4f(\chi_s) \qquad 325.$$

and as such the rays within each system will diverge from their respective axes. This is the result expected for ray propagation in cylindrical waveguides.

If the cross section of the waveguide is elliptical, then instead of the incident angle χ of each ray remaining constant as in a circular cross section, χ will now vary as a function of β. This will alter the values of χ_s for each p,$\overline{m}$ (other than p = 1, m = 2) and possibly place restrictions on the orientations of the closed figures.

M	p = 1		p = 2		p = 3		p = 4	
	$\chi_s{}^o$	$\Delta\chi^o$	$\chi_s{}^o$	$\Delta\chi^o$	$\chi_s{}^o$	$\Delta\chi^o$	$\chi_s{}^o$	$\Delta\chi^o$
2	0	30						
3	30	15						
4	45	9	0	18				
5	54	6	18	12				
6	60	4.3	30	9	0	12.8		
7	64.3	3.2	39	6	12.8	9.7		
8	67.5	2.5	45	5	22.5	7.5	0	10
9	70	2	50	4	30	6	10	8
10	72		54	3.3	36	4.9	18	6.5
11			57.3	2.7	40.9	4.1	24.5	5.5
12			60		45	3.5	30	4.6
13					48.5	3	34.6	4
14					51.5	2.5	38.6	3.4
15					54		42	3
16					54		45	3
17							48	2
18							50	2.1
19							52.1	1.9
20							54	

Table 9. Stationary skew plane systems.

This is demonstrated by considering the possible orientations of the $\chi_s = 0$ ($P = 1, \bar{m} = 2$) closed figure (straight line). The ray with $\chi_s = 0$ must coincide with the normal to the reflecting surface at each reflection point and in the ellipse this is only possible when $\beta = 0$ or $\frac{\pi}{2}$. The spacing d_o between the reflectors when $\beta = 0$ is

$$d_o = 2a \qquad 326.$$

and the radius of curvature $\bar{r}_o$ about this axis is given by Equation (250) as

$$\bar{r}_o = \frac{b^2}{a} \qquad 327.$$

The focal length f_o of the reflectors is given by Equation (305) by substituting $\bar{r}_o$ in places of the radius a of the circle, to give

$$f_o = \frac{b^2}{2a}$$

Substituting f_o and d_o into the stability condition, Equation (309) gives

$$(1 - \frac{a^2}{b^2}) < 0$$

since $a > b$, demonstrating that this is a non-stable system, in agreement with the results of the rigorous theory of Section 6.3.

A similar procedure for the axis at $\beta = \frac{\pi}{2}$ gives $d_{\pi/2} = 2b$, $\bar{r}_{\pi/2} = \frac{a^2}{b}$ to give a stability condition of

$$(1 - \frac{b^2}{a^2}) > 0$$

which corresponds to the trapping system described in Section 6.3.

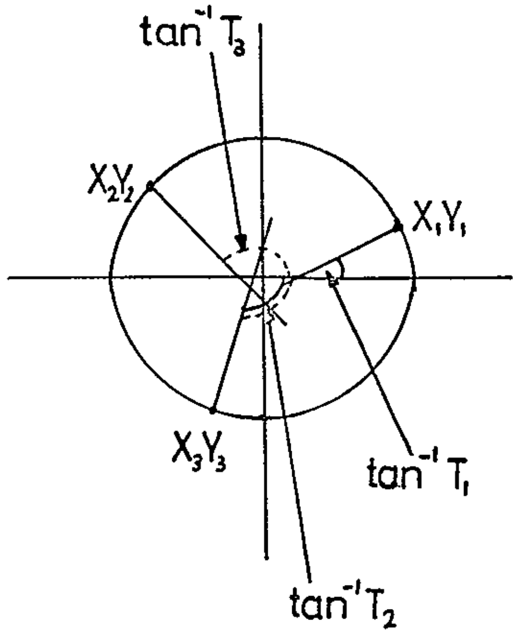


Figure 104. The co-ordinates of three arbitrary points on an ellipse.

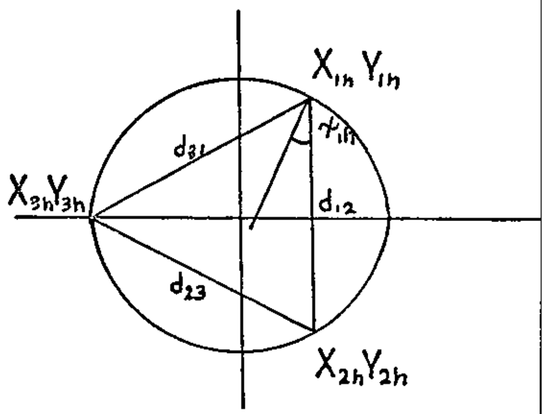


Figure 105a.

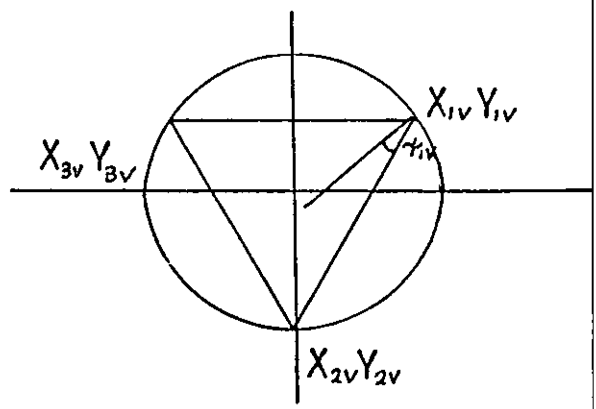


Figure 105b. The vertical and horizontal triangu ray paths.

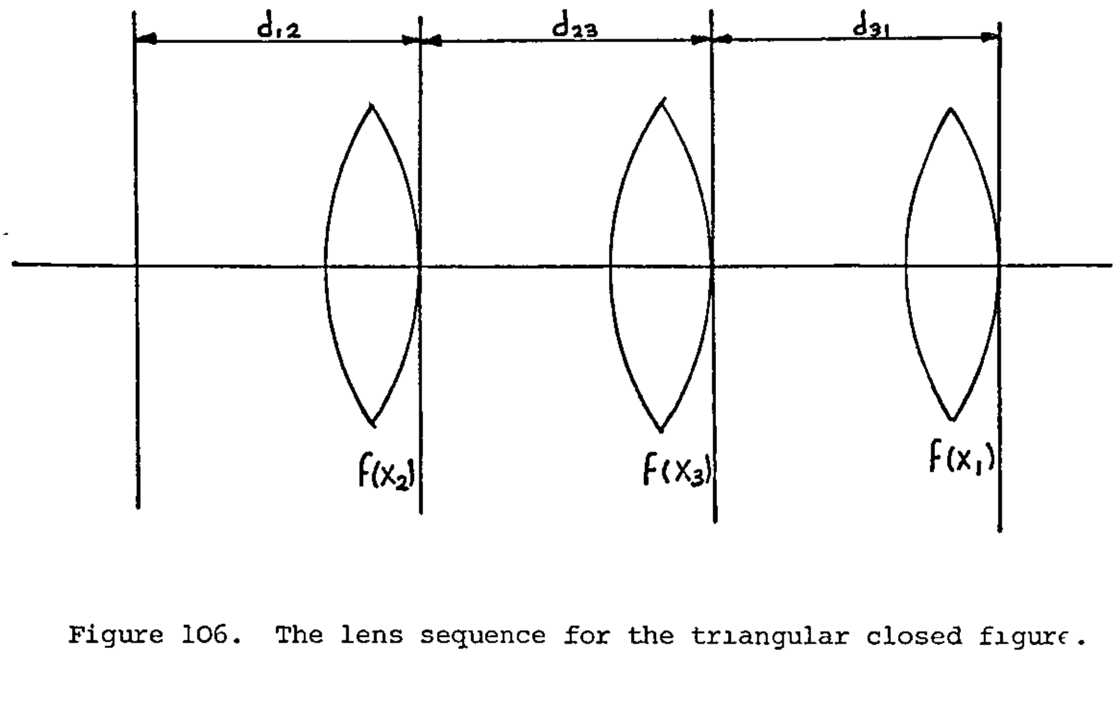


Figure 106. The lens sequence for the triangular closed figure.

The possible orientations and stability of the $\chi_s \sim 30^{\circ}$ ($p = 1$, $m = 3$) closed figures may be obtained using similar arguments. Consider the three points on the periphery of the ellipse shown in Figure 104 with co-ordinates $X_1\, Y_1$, $X_2\, Y_2$, $X_3\, Y_3$. The normals to the reflecting surface at these points have tangents T_1, T_2, T_3 as shown. If the three points are joined by rays to form a triangle, then the normals bisect the internal angles of the triangle. A well known theorem states that bisectors of the internal angles of a triangle intersect at a point which is the centre of the inscribed circle of the triangle.

This theorem, together with the expressions describing the points of intersection of two lines, Equations (164a-b), are used in Appendix D to show that there are two possible orientations for each of two triangular closed figures within the ellipticel cross section.

The two closed figures for $p = 1$ $\bar{m} = 3$ are shown in Figure 105 and their variables will be denoted by the subscript $_h$ for horizontal (Figure 105a) and $_v$ for vertical (Figure 105b). The second orientation of each triangle follows from the symmetry of the ellipse.

The lens sequence equivalent to a single trip round the triangular path is shown in Figure 106. This lens sequence may be represented by a single ray transfer matrix $\begin{vmatrix} A & B \\ C & D \end{vmatrix}$ where

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} 1 & d_1 \\ -\dfrac{1}{f(X_1)} & 1 - \dfrac{d_1}{f(X_1)} \end{vmatrix} \begin{vmatrix} 1 & d_2 \\ -\dfrac{1}{f(X_2)} & 1 - \dfrac{d_2}{f(X_2)} \end{vmatrix} \begin{vmatrix} 1 & d_3 \\ -\dfrac{1}{f(X_3)} & 1 - \dfrac{d_3}{f(X_3)} \end{vmatrix} \qquad 328$$

The multiplication of the matrices on the right hand side of the Equation is simplified by noting that

$$f(X_1) = f(X_2) \equiv f_1 \qquad 329. \qquad d_2 = d_3 \qquad 330.$$

and the subscripts may be simplified to give

$$f(X_3) \equiv f_3$$

The multiplication of the right hand side of Equation (328) gives

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} 1 - \dfrac{3d_3}{f_1} + \dfrac{d_3^2}{f_1^2} & d_1 + 2d_3 - \dfrac{3d_1d_3}{f_1} - \dfrac{d_3^2}{f_1} + \dfrac{d_1d_3^2}{f_1^2} \\ -\dfrac{1}{f_3} - \dfrac{2}{f_1} + \dfrac{3d_3}{f_1f_3} + \dfrac{d_3}{f_1^2} - \dfrac{d_3^2}{f_1^2f_3} & 1 - \dfrac{2d_3}{f_3} - \dfrac{2d_1}{f_1} - \dfrac{d_3}{f_1} - \dfrac{d_1}{f_3} + \dfrac{3d_1d_3}{f_1f_3} + \dfrac{d_1d_3}{f_1^2} + \dfrac{d_3^2}{f_1f_3} - \dfrac{d_1d_3^2}{f_1^2f_3} \end{vmatrix} \qquad 331.$$

The stability condition for rays in a sequence of lens systems represented by this matrix is given by Equation (307) as before and where

$$\tfrac{1}{2}(A+D) = \left(1 - \frac{(f_1+2f_3)(2d_3+d_1) - d_3(3d_1+d_3)}{2f_1f_3} + \frac{d_3(f_3d_1+f_3d_3-d_1d_3)}{2f_1^2f_3}\right) \qquad 332.$$

The co-ordinates X_1Y_1, X_2Y_2 of each triangle may be found by using the property that the normal at the point of reflection is also the bisector of the total angle subtended by the incident and reflected rays. If χ_{1v} is the incident angle of a ray at X_{1v} Y_{1v} in Figure 105b then

$$\mathrm{TAN}\ \chi_{1v} = T_{1v} \qquad 333.$$

and

$$\mathrm{TAN}\ 2\chi_{1v} = \frac{Y_{1v} + b}{X_{1v}} \qquad 334.$$

Using the identity $TAN\ 2x = \frac{2TANx}{1-2TAN^2x}$ and substituting for T and x using Equations (246), (240a) gives

$$TAN\ 2\chi_{1v} = \frac{b(Y_{1v}-b)}{a\sqrt{b^2-Y_{1v}^2}} \qquad \text{from Equation (334)} \qquad 335.$$

and

$$TAN\ 2\chi_{1v} = \frac{2aY_{1v}b(b^2-Y_{1v}^2)}{\sqrt{b^2-Y_{1v}^2}\left(b^2(b^2-Y_{1v}^2)-a^2Y_{1v}^2\right)} \qquad \text{from Equation (333)} \qquad 336.$$

Equating the right hand sides of Equations (335) and (336) gives a quadratic equation in Y_{1v} whose solution is

$$Y_{1v} = \frac{-a^2b \pm ab^2\sqrt{\frac{a^2}{b^2} + \frac{b^2}{a^2} - 1}}{(a^2 - b^2)} \qquad 337.$$

and X_{1v} may be found by using Equation (240a).

The co-ordinates X_{1h} Y_{1h}, X_{2h} Y_{2h} of the horizontal triangle may be found in a similar manner by noting that

$$TAN\ \chi_{1h} = \frac{1}{T_{1h}} \qquad 338.$$

and

$$TAN\ 2\chi_{1h} = \frac{a + X_{1h}}{Y_{1h}} \qquad 339.$$

A quadratic equation is formed in terms of X_{1h} and which has the solution

$$X_{1h} = \frac{b^2a \pm a^2b\sqrt{\frac{a^2}{b^2} + \frac{b^2}{a^2} - 1}}{(b^2 - a^2)} \qquad 340.$$

and Y_{1h} may be found using Equation (240b).

Table 10. Stability conditions for stationary skew plane systems.

Horizontal Triangle

e.	f_1h	f_3h	d_1h	d_3h	½(A+D)
1.00	.433	.433	1.732	1.732	-1.00
1.01	.434	.434	1.727	1.740	-1.002
1.02	.4349	.4350	1.723	1.749	-1.007
1.03	.4357	.4357	1.719	1.758	-1.020
1.04	.4360	.4364	1.7146	1.7666	-1.044
1.05	.4375	.4374	1.7104	1.775	-1.0298
1.10	.440	.4390	1.690	1.82	-1.143

Vertical Triangle

e	f_1v	f_3v	d_1v	d_3v	½(A+D)
1.00	.433	.433	1.732	1.732	-1.00
1.01	.434	.434	1.745	1.732	-1.003
1.02	.435	.435	1.757	1.732	-1.005
1.03	.436	.436	1.770	1.732	-1.006
1.04	.4365	.4367	1.7832	1.732	-1.0244
1.05	.4375	.4374	1.7958	1.7325	-1.03
1.10	.4401	.4399	1.8578	1.734	-1.132

Sensitivity of circular cross section

f	d	½(A+D)	
.433	1.732	-1.0	Stable
.432	1.732	-1.04	Unstable
.434	1.732	-0.959	Stable
.433	1.733	-1.010	Unstable
.433	1.731	-0.989	Stable
.433	1.730	0.979	Stable

The radii of curvature $\bar{r}(x)$ at each point of reflection are given by substitution of X into Equation (248) and the focal lengths obtained by substitution of $\bar{r}(x)$ for a in Equation (305). The lens spacings are given by

$$d_{1h} = 2Y_{1h} \qquad 341h. \qquad\qquad d_{1v} = 2X_{1v} \qquad 341v.$$

$$d_{3h} = \sqrt{Y_{1h}^2 + (a+X_{1h})^2} \qquad 342h. \qquad\qquad d_{3v} = \sqrt{X_{1v}^2 + (b+Y_{1v})^2} \qquad 342v.$$

Finally, the relevant values of f and d may be substituted into Equation (332) to give the stability of each closed figure. Table 10 gives the values of parameters d, f, ½(A+D) for various values of e where $e = a^2$ and $b = 1$.

When e = 1, corresponding to a circular cross section waveguide, the normalized radius of curvature and lens spacing become

$$\bar{r}_1 = \bar{r}_3 = 1.0 \qquad\qquad d_1 = d_3 = \sqrt{3}$$

and the stability is unity.

Table 10 demonstrates the sensitivity of the stability of this system to variations in the radius of curvature or lens spacing.

It is clear from the results shown in Table 10 that the p = 1, $\bar{m} = 3$ systems in an ellipse are unstable, but on the basis of the sensitivity of the e = 1 system, small deviations of the waveguide cross section from the elliptical cross section could produce a stable system. Such a stable system would confine rays of light about the stationary skew ray path and they will form regions of relatively high intensity compared to neighbouring less stable or unstable systems.

Figure 107.

Figure 108.

Figure 109a.

Figure 109b.

Figure 109c.

An example of a stable triangular system is shown in Figure 5c, where the orientation of the triangular pattern of high intensity corresponds approximately with the vertical triangle system. An example of an unstable triangular system is shown in Figure 5d, where shadow regions appear in the horizontal triangle positions. In Figure 109c, both the vertical and horizontal triangular systems appear unstable and form a hexagon shadow region. This is a result obtained from a different fibre in the same bundle as the fibre used for Figure 5. The patterns, evident in Figures 107 to 109 were all obtained with white light illumination of the bundle at an axial angle of incidence of 25^{o}. Figures 107 and 108 are from two further fibres from the same bundle as that used for Figure 109 and are included here to illustrate the variety of patterns obtained from fibres of the same length and under similar illuminating conditions. In Figure 107, two trapped skew plane caustics are inclined at 60^{o} to each other, suggesting that this fibre cross section has more than one trapping axis and is clearly not elliptical. In Figure 108, a stable system corresponding to the $p = 2$, $m = 2$, stationary skew plane (shown in Figure 49) appears as a pentagon figure of higher intensity than its immediate surroundings

The fibre patterns shown in Figures 109 a, b, c, show a complex frequency selective trapping system. In Figure 109a, $\alpha \simeq 0$ and the white caustic shown in the microphotograph has a strong blue colouration. At approximately right angles to it (that is in the position of the 'white' caustic shown in Figure 109c) there is a further caustic with red colouration. When $\alpha \simeq 90^{o}$ the two coloured caustics change position as shown in Figure 109c. (The Polaroid photographic material used in this study is less sensitive to the red end of the spectrum than the blue). When $\alpha \simeq 50^{o}$ the pattern shown in Figure 109b is obtained where the positions of the two coloured caustics now show as shadow regions. The appearances of these patterns are sensitive to azimuthal angles of incidence to within 2^{o}. The cross sectional geometry which could form this system has not been determined.

6.7 Conclusions

In this chapter the caustics formed in elliptical cross section waveguides have been examined and features of the theory illustrated by experimental results. The main conclusion to be drawn from the experimental results is that the waveguides which from their p = 1, m = 2 caustics appear to be elliptical in cross section also exhibit higher order caustics which are not found in the elliptical caustic theory. Thus the expected result that the waveguides are neither circular nor elliptical is confirmed and this suggests that a different approach to the problem of relating the pattern formation observed within any given length of waveguide, to its cross sectional geometry is required. A possible alternative approach is described in Chapter 9.

However, those fibres which do exhibit features of the elliptical cross section caustic theory may well have cross sectional geometries which differ from elliptical by very small amounts, in which case it may be that for certain purposes a perfectly elliptical cross section may be assumed. A major problem encountered in the measurement of the cross sectional geometries using the pattern observation method, is that there is no other known method of measuring the waveguide dimensions within the tolerances of dimensions which will produce the patterns. It is therefore not possible to assess the accuracy of the results by an independent measurement.

Because of the many unknown features of propagation of light in the dielectric waveguides used for the experiments, further experiments were conducted, as described in the following two chapters, in an attempt to obtain more information about other features of the waveguides which may influence the formation of the caustic patterns.

CHAPTER 7

7.1 Introduction

It appears from the results presented in Chapter 6 that the variations of intensity observed in a waveguide are primarily caused by its non-circular cross section. In particular the analysis of light propagating in an elliptical cross section waveguide has provided fairly conclusive evidence that the variations of intensity are directly related to caustics produced within the waveguide cross section.

It is more difficult to interpret the variations of the visibility of the patterns from waveguide to waveguide in terms of waveguide geometry, material homogeneity or other defects. In the analysis of the caustics formed in circular or elliptical cross section waveguides it was assumed that the waveguides were of constant cross section, straight, and free of all defects. On this basis, experimental results were obtained to demonstrate the agreement between the theoretical positions of the caustics and the variations of intensity within the waveguides. It is natural that waveguides exhibiting variations of intensity which show the clearest agreement with the theoretical caustic positions should be presented.

However, this does not mean that the assumptions used in the caustic analysis are applicable to the waveguides used to demonstrate the theory. This can only be shown by examining the effects of these assumptions on the caustics and if it is found that relaxation of the assumptions causes observable changes in the caustic positions then these changes can be investigated experimentally.

Numerous experiments were conducted in an attempt to detect such variations in the visibility or the form of the patterns observed within a particular waveguide as the length of the waveguide was varied or as the waveguide was bent. These experiments and the results are described in this chapter and the behaviour of the skew rays when a waveguide suffers a sharp bend is considered theoretically.

7.2 Length Dependent Variations in Waveguide Characteristics

The optical waveguides used for the experiments may have any or all of the following defects.

a) Variations in the cross sectional geometry and dimensions of the waveguide core and/or cladding.

b) Variations in the refractive index profile.

c) Randomly positioned scattering centres (air bubbles etc.)

The above waveguide defects are primarily introduced during manufacture and generally cannot be modified afterwards. Two further parameters which may be experimentally varied within limits are

d) Bends and twists in the waveguides.

To analyse the effects of (a) - (d) above upon the visibility and form of the caustic patterns requires prior knowledge of the distribution and magnitude of the defects and this information is not available. However, the following comments may assist in determining the possible effects of each kind of defect.

The factors which may cause defects (a) and (b) during manufacture (temperature, tension, and speed of pulling) are likely to vary slowly with respect to the speed of pulling of the fibre. It would seem likely that the fibres will maintain their geometry and dimensions over lengths of at least 1 m when the pulling rate is typically 1 metre per second.

If the scattering centres are caused by defects in the preforms then it is likely that they too will have a length dependent distribution where the distribution will also vary slowly with respect to the pulling rate.

Information obtained from the manufacturers of the fibres revealed that although the nominal loss of the fibres is quoted as 1 dB/m, certain fibres may have losses over short lengths ($\simeq$ 1m) considerably less than this figure. No explanation for this was offered other than the statistical probability that a few fibres will have losses showing large variation from the mean.

These comments suggest that the number of fibres within a bundle which will exhibit good contrast patterns may depend upon the length of the bundle, and the following experiment was conducted to test this proposal. A 4m length of fibre bundle was cut into lengths of 2m, 1m, 0.5m, 0.25m and .125m and each end of these bundles was embedded in resin, ground and polished as described in Chapter 2.

Using the experimental arrangement shown in Figure 20 each bundle was illuminated with plane waves from a white light source. The number of fibres within each bundle exhibiting good contrast, poor contrast or no patterns were counted and the results are shown in Table 11.

Fibre length (m)	Number of fibres with		
	good contrast patterns,	poor contrast patterns	no patterns
2	2	10	388
1	4	16	380
0.5	30	80	290
0.25	75	95	230
0.125	140	100	160

Table 11.

These results demonstrate a strong correlation between the length of the fibre bundles and the number of fibres within the bundles exhibiting good contrast patterns.

To investigate further the length dependence of the contrast of the patterns of a single fibre the following experiment was conducted. A 1m length of a single fibre which exhibited a poor contrast pattern was shortened in 1 cm steps down to a length of 0.25 m. There was a gradual improvement in the contrast of the patterns as the length of the fibre was reduced although at a length of 0.25m the patterns were still of relatively poor contrast. Similar tests were performed on five other single fibres and similar gradual changes in pattern contrast were observed.

Considerable care was taken to ensure that all the end terminations of the single fibres were of comparable quality since it was found that a poor end termination reduced the visibility of the patterns.

These subjective tests yield results which only permit tentative conclusions to be drawn about the length dependence of the contrast of the patterns. It would appear that there are no rapid length dependent changes in the waveguide geometry since no variations in the form of the patterns were observed as the length of the fibres were changed. The results of the tests reveal no information as to the source of the variations in the contrast of the patterns although it is possible that the degradation of the contrast of the patterns is related to random scattering of light within the waveguide.

To investigate this possibility an alternative measurement of the scattering of light within a waveguide is required. Other experimenters have measured the axial angle dependent distribution of the radiated cone of light and have attempted to relate this function to the length

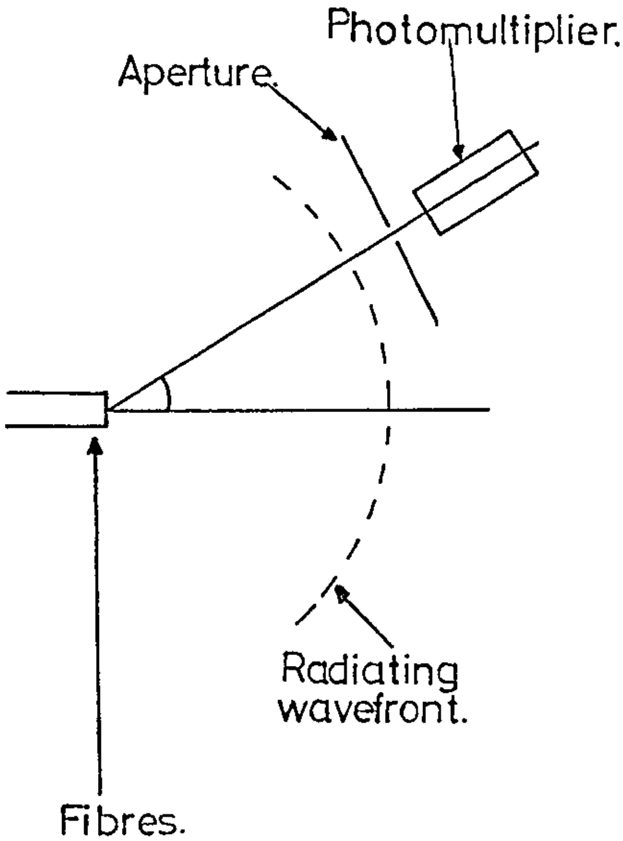


Figure 110. A typical arrangement for the measurement of the radiation field of fibres.

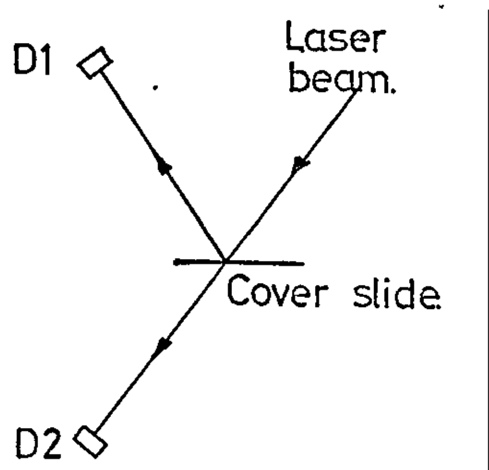


Figure 112. The experimental arrangement for t measurement of T coefficients of r during coating.

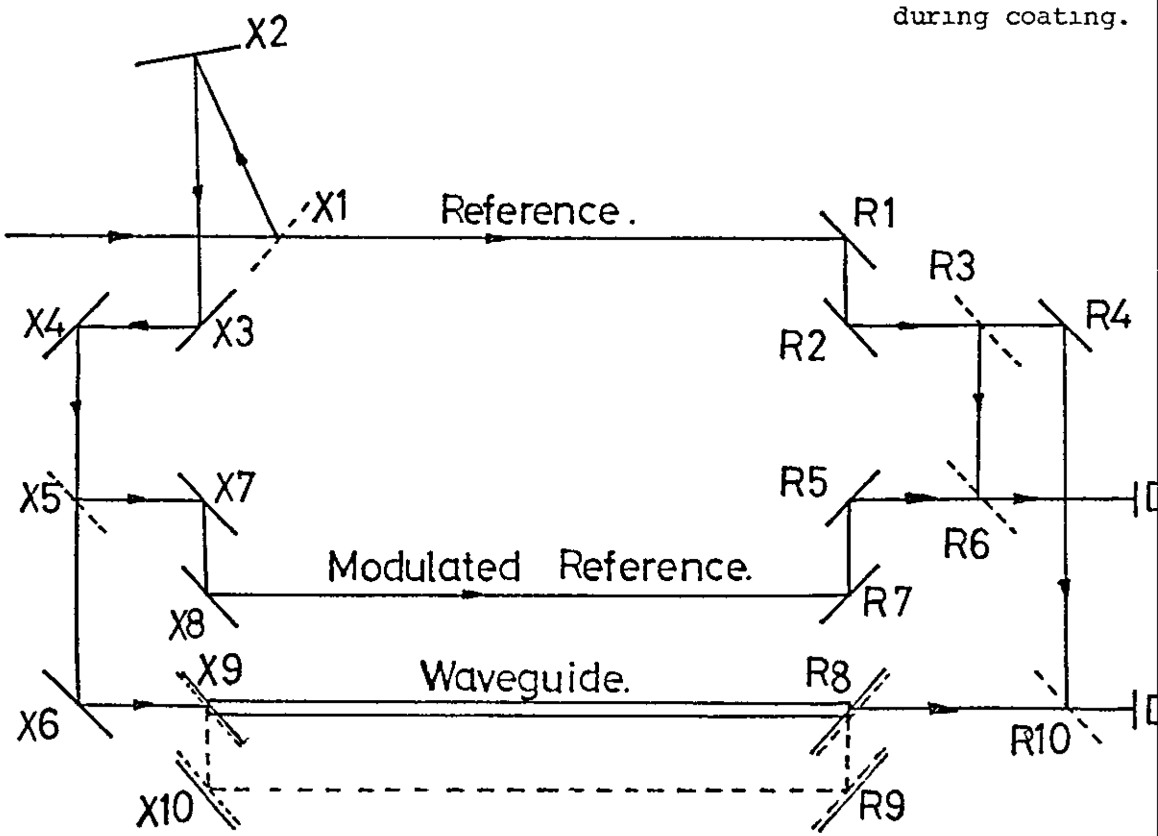


Figure 111. The diagrammatic representation of the heterodyne scanning system.

of the waveguide and its temporal dispersion. It is proposed to use a similar measurement to indicate the magnitude of the scattering and to relate the results to the contrast of the patterns.

Although the equipment for these experiments was designed and built, insufficient time prevented the production of results. However, because of its novel design, the test equipment and its proposed mode of operation are described in the next section.

7.3 Heterodyne Scanning System

A typical experimental arrangement for measuring the angular distribution of radiation from a single fibre or fibre bundle is shown in Figure 110. The fibres are illuminated by plane waves from a white light or laser source and their radiation field is scanned with a photomultiplier to measure the angular distribution of radiated light for each input angle. The photomultiplier provides sufficient gain to give good angular resolution of the radiated field of the waveguide.

Since no photomultiplier was available, a heterodyne detection system was proposed, firstly to increase the gain of the detection system and secondly to provide phase information for use in later experiments. The system designed and constructed is shown in diagrammatic form in Figure 111.

A laser beam is divided by the partially silvered mirror X1 and the transmitted beam forms a reference beam which passes to the two photodiodes via the silvered mirrors R1, R2, R4 and the partially silvered mirrors R3, R6 and R10. The laser beam reflected from mirror X1 is phase modulated by reflection from the silvered mirror X2 which is attached to the cone of a loudspeaker. This modulated beam passes through the transmitter arm via silvered mirrors X3 and X4 and is split by the partially silvered mirror X5. The reflected beam from mirror X5 is passed to photodiode D2 via the silvered mirrors X7, X8, R7, R5 and the partially silvered mirror R6. The transmitted beam from mirror X5 illuminates the waveguide entrance aperture via silvered mirror X6 and the radiation field of the waveguide is detected by photodiode D1 via the partially silvered mirror R10. To align the system the waveguide may be replaced by silvered mirrors X9, X10, R8, R9 to form a path similar to that of the reflected beam from mirror X5 to photodiode D2.

Mirrors X3 to X6 (and X10) are mounted on the swivelling transmitter arm and mirrors R2 to R6 (and R9) and the two photodiodes are mounted on the swivelling receiver arm. The two arms are rotated by stepper motors via worm drives and the position of each arm is measured using a potentiometer. The sequence of movements of the two arms is automatically controlled such that the receiver arm makes a sweep of the waveguide radiation field for each position of the transmitter arm.

The two light beams incident upon photodiode D1 may be represented by their E field components E_{R1}, E_{WG} where E_{R1} represents the reference field and E_{WG} represents the waveguide radiation field. If it is assumed that both of these fields have constant phase over the sensitive area of the photodiode, then the output of the photodiode will be proportional to the incident intensity I_1 where from Equation 33)

$$I_1 = \overline{(E_{R1} + E_{WG})^2} \qquad 343.$$

The E fields E_{R1}, E_{WG} may be represented by

$$E_{R1} = A \cos \omega t \qquad 344.$$

$$E_{WG} = B \cos (\omega t + \phi(t) + \gamma(t) + \alpha) \qquad 345.$$

where A, B are the amplitudes of the fields, ωt is the time dependence of the optical carrier, $\phi(t)$ is the phase modulation due to the movement of mirror X2 and $\gamma(t)$ is a phase shift due to mechanical vibrations of the test rig. The term α represents the phase shift introduced by the waveguide.

Substituting equations (344), (345) into (343) gives

$$I_1 = \overline{A^2 \cos^2 \omega t} + \overline{B^2 \cos^2(\omega t + \phi(t) + \gamma(t) + \alpha)}$$

$$+ \overline{AB \cos(2\omega t + \phi(t) + \gamma(t) + \alpha)} + \overline{AB \cos(\phi(t) + \gamma(t) + \alpha)} \qquad 346.$$

The first three terms of Equation (346) vary at or above the optical carrier frequency $\frac{\omega}{2\pi}$ and their time average will give a DC term in the output of the photodiode. The amplitude AB of the fourth term is the required amplitude B of the radiation field multiplied by the amplitude A of the reference field, the latter term forming the gain factor achieved by heterodyne detection. A detailed study of the system would consider the alignment requirements of the two beams to achieve the constant phase requirement over the sensitive area of the photodetector as well as optimising the ratio of the amplitudes of the two beams by varying the silvering of the mirrors to give the maximum signal amplitude.

This last requirement may be calculated for the output of the second photodiode if it is assumed that the shape of the beams are unaltered by the multiple reflection process and any divergence of the beams is neglected. Denoting the amplitude of the electric field of the original laser beam as D and the transmission and reflection coefficients of the partially silvered mirrors as T( ), R( ), respectively then the amplitude C of the modulated reference beam at photodiode D2 is given by

$$C = D(R(X1).\ R(X5).\ T(R6)) \qquad 347.$$

where it is assumed that the silvered mirrors have R = 1.0. Denoting the amplitude of the reference beam at photodiode D2 as E then

$$E = D(T(X1).\ R(R3).\ R(R6)) \qquad 348.$$

and the amplitude of the envelope at the output of the photodiode will be proportional to CE where

$$CE = D^2.\ TR(X1).\ TR(R6).\ R(X5).\ R(R3) \qquad 349.$$

The products of the coefficients for mirros X1 and R6 are maximised when T = R = 0.5 since

$$T + R = 1.0$$

then $$TR = T(1 - T)$$

and $$\frac{d(TR)}{dT} = 1 - 2T$$

Maximum of TR occurs when $\frac{d(TR)}{dT} = 0$ so that

$$(1 - 2T) = 0$$

and $$T = 0.5$$

To assist in the initial alignment of the system all the partially silvered mirrors were manufactured with T = R = 0.5 so that when the mirrors X10, X9, R8, R9 were used in place of a waveguide, the amplitudes of the envelopes in the outputs of the two photodiodes are equal.

However, since

$$AB = F.D^2. \; TR(X1). \; TR(R10) \; T(X5) \; T(R3) \tag{350.}$$

where F represents the reduction in amplitude due to propagation in a waveguide (F = 1 during alignment with mirrors) some compensation for F << 1 may be obtained by making T > R for mirrors X5 and R3.

The mirrors were manufactured by evaporation of aluminium onto glass microscope cover slides and the R and T coefficients were monitored during deposition using the arrangement shown in Figure 112.

The reason for the phase modulation provided by the loudspeaker becomes apparent when the form of the output given by Equation (346) is considered in the absence of the $\phi(t)$ term. The DC contribution of the first three terms will be as before but the amplitude of the fourth term will now vary with time as $\gamma(t)$ varies. The random nature of $\gamma(t)$ (unless the whole system is mounted on a vibration free table) would require an unacceptably long time averaging period to obtain an amplitude measurement and the envisaged motion of the receiver swivel arm during measurement would be intolerable. By introducing a time varying phase factor whose period will generally be very much shorter than that of the vibration function $\gamma(t)$, a measurement of the amplitude of the envelope of $\phi(t)$ may be made over a few cycles of $\phi(t)$ during which period $\gamma(t)$ will only change the instantaneous phase of the carrier.

The E fields incident upon the photodiode D2 may be represented by

$$E_{R2} = E \cos(\omega t + \beta + \gamma(t)) \tag{351.}$$

$$E_{mR} = C \cos(\omega t + \phi(t) + \delta) \tag{352.}$$

where E_{R2} is the E field of the reference beam and E_{mR} is the E field

Figure 113a. The transmitter arm assembly.

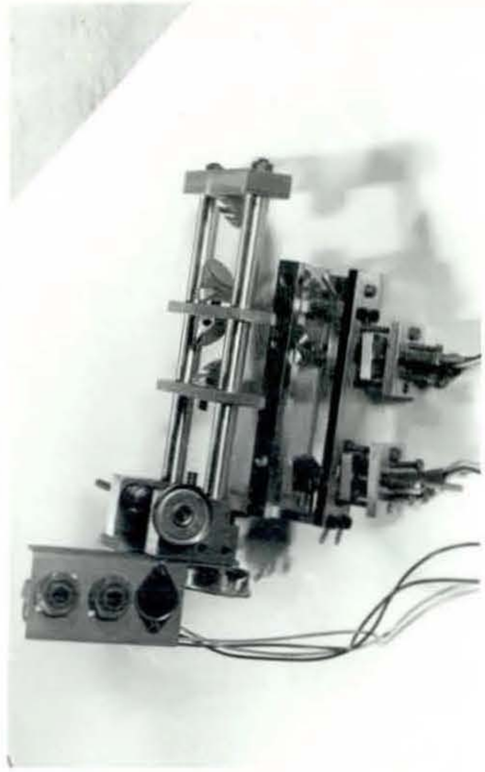


Figure 113b. The receiver arm assembly.

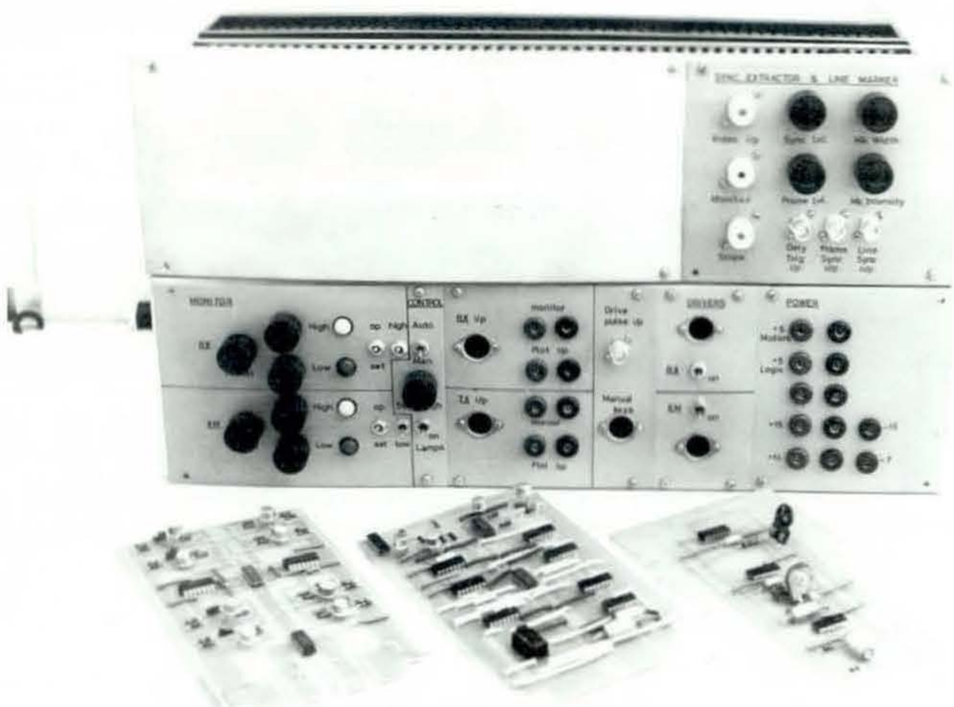


Figure 114. Printed circuit boards and control panel for the heterodyne scanning rig. (Also the sync extractor and line marker)

of the modulated reference beam. The phase shifts β and δ are arbitrary and assumed to be constant. The output of the photodiode D2 will be proportional to the incident intensity I_2 where

$$I_2 = \text{DC TERM} + \text{CE} \cos(\phi(t) + \gamma(t) + \beta + \delta) \qquad 353.$$

Assuming that the vibration term $\gamma(t)$ is the same for both photodiodes, the phase shift α may be extracted from Equation (346) by comparing the phase of the output from D1 with that from D2.

Further experimental work is required to find the variation of the phase shift α as the input or output light field is scanned, and also to ascertain the validity of the assumptions about the vibration term $\gamma(t)$ and the phase shifts β and δ.

Photographs of the transmitter and receiver arms are shown in Figure 113. These two components were designed for use in conjunction with a straight fibre mount which is described in the next section. The printed circuit boards and control unit designed to drive and control the swivel arms are shown in Figure 114.

7.4 Straight Fibre Mounting

Consideration was given to the problem of mounting a length of fibre in such a manner that its orientation in space could be controlled and measured.

The basic problem is illustrated in Figure 115 where the position of a fibre within a volume will be given by a set of XY co-ordinate values taken at intervals along the Z axis. The minimum size of the measurement intervals is determined by the rate of change of the

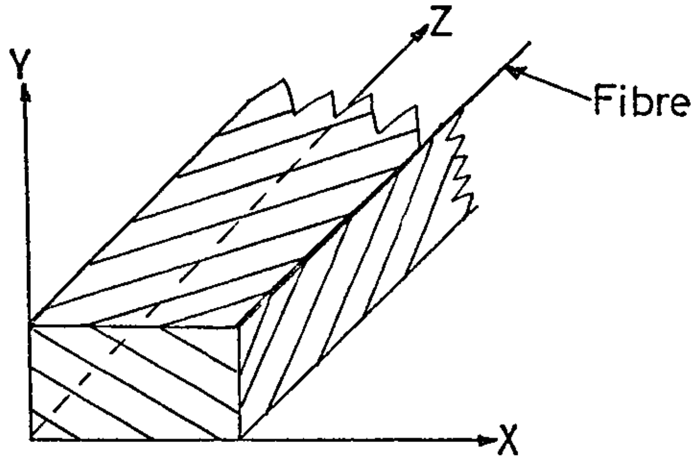


Figure 115. The allocation of space for mounting fibres.

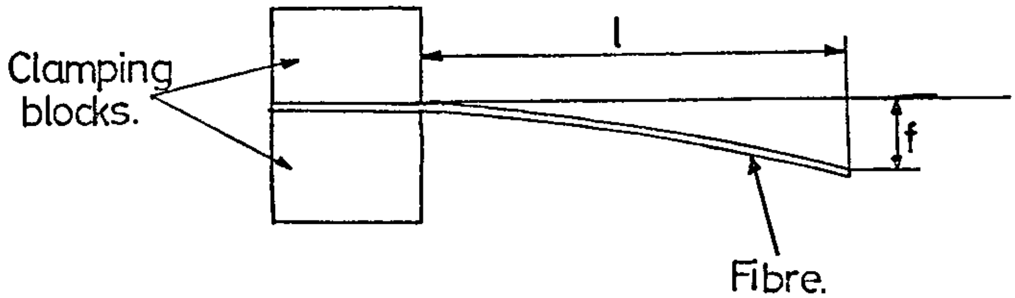


Figure 116. The experimental arrangements for measuring the deflection of self loaded fibres.

l (cms)	3.3	2.9	2.7	2.5	2.3	2.2
f (μ)	930	550	340	250	50	50

Table 12. The deflections of self loaded fibres.

position of the fibre within the XY plane as a function of Z, such that

$$\frac{\partial(X)}{\partial Z} \text{ or } \frac{\partial(Y)}{\partial Z} < \text{ the required resolution of position.}$$

In practical terms the measurement of the position of the fibre in space will require the use of a travelling microscope whose optical axis may be either parallel to the X axis for measurements of displacement in the Y direction or parallel to the Y axis for measuring X direction displacements. The short working distances of the large magnification microscope objectives required for such measurements means that the unshaded region of the plane XY shown in Figure 115 would be required for access to the fibre by the microscope. This leaves only the shaded region in which to provide a supporting structure which will also permit adjustment of the fibre position to within the required resolution.

The high tensile strength (Ref 61) of glass fibres suggests that it may be possible to support a fibre at discrete intervals along its length and still maintain the deflection of the fibre due to its own weight within the required tolerance in position. To measure the deflection of a fibre due to its own weight the experimental arrangement shown in Figure 116 was used. The deflection f of the fibre was measured for each length ℓ of overhanging fibre to give the results shown in Table 12.

A double fulcrum experiment based upon the above results suggested that the fibre position could be maintained within its own diameter by supporting the fibre at discrete intervals along its length, if the intervals did not exceed 4 cms for the fibre under test.

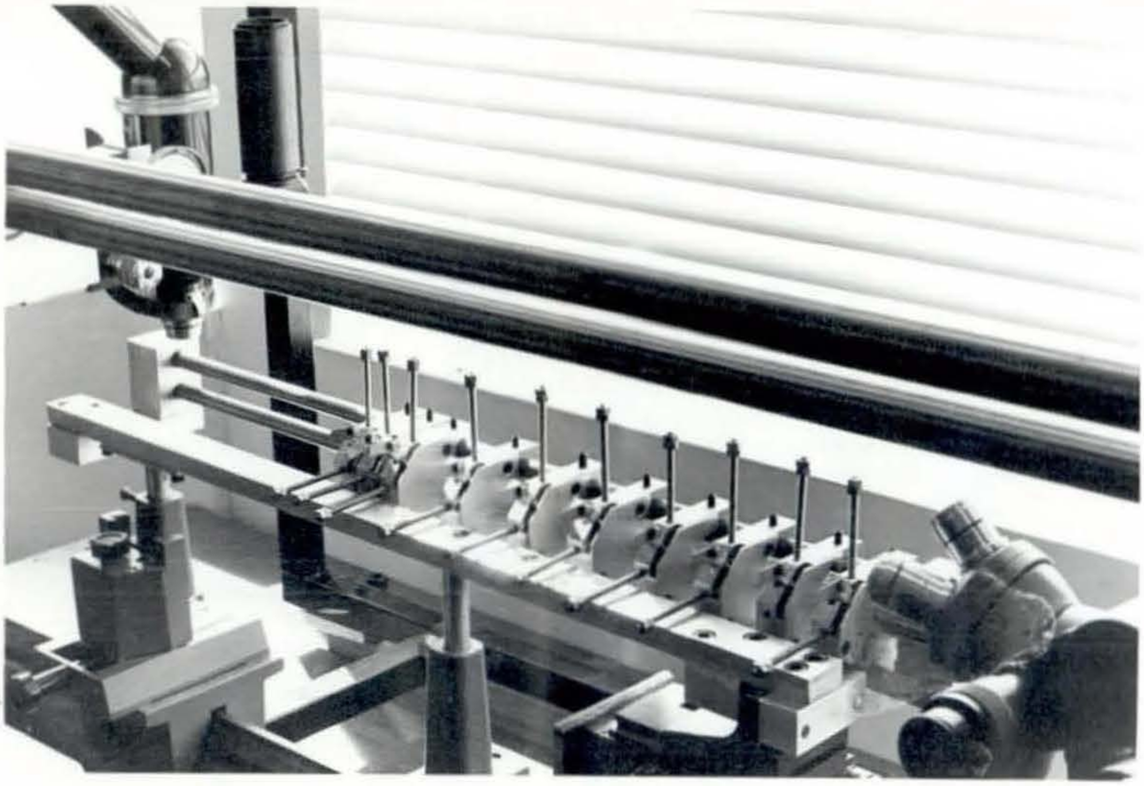


Figure 117. The straight fibre mounting system.

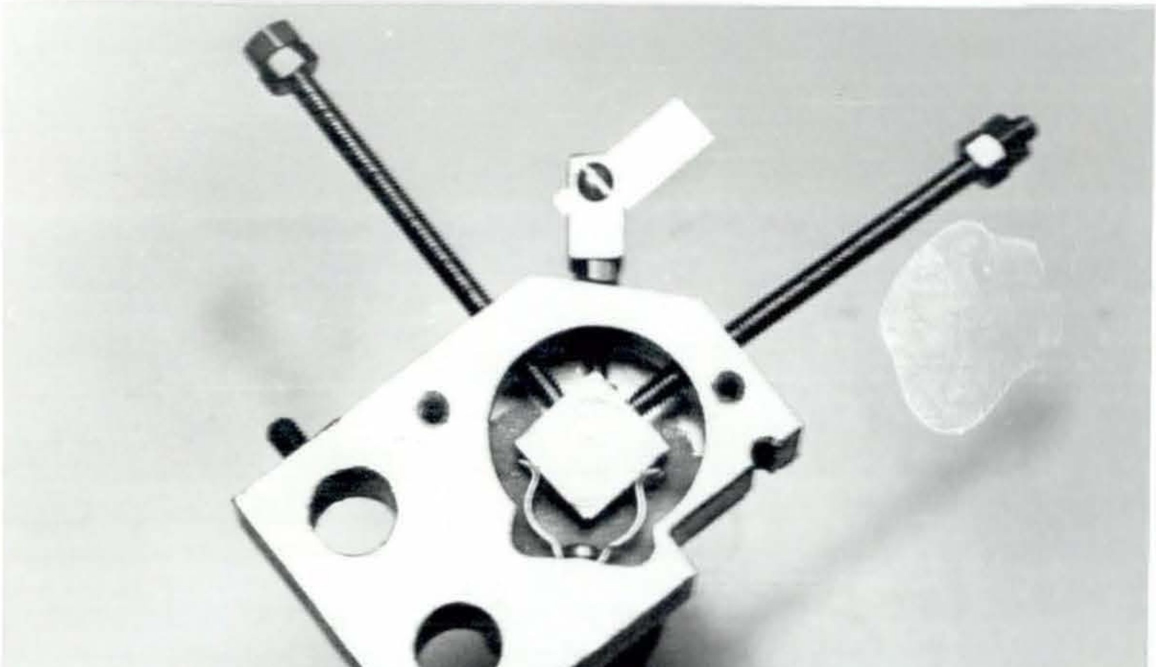


Figure 118. Detail of one of the supports of the straight fibre mounting system.

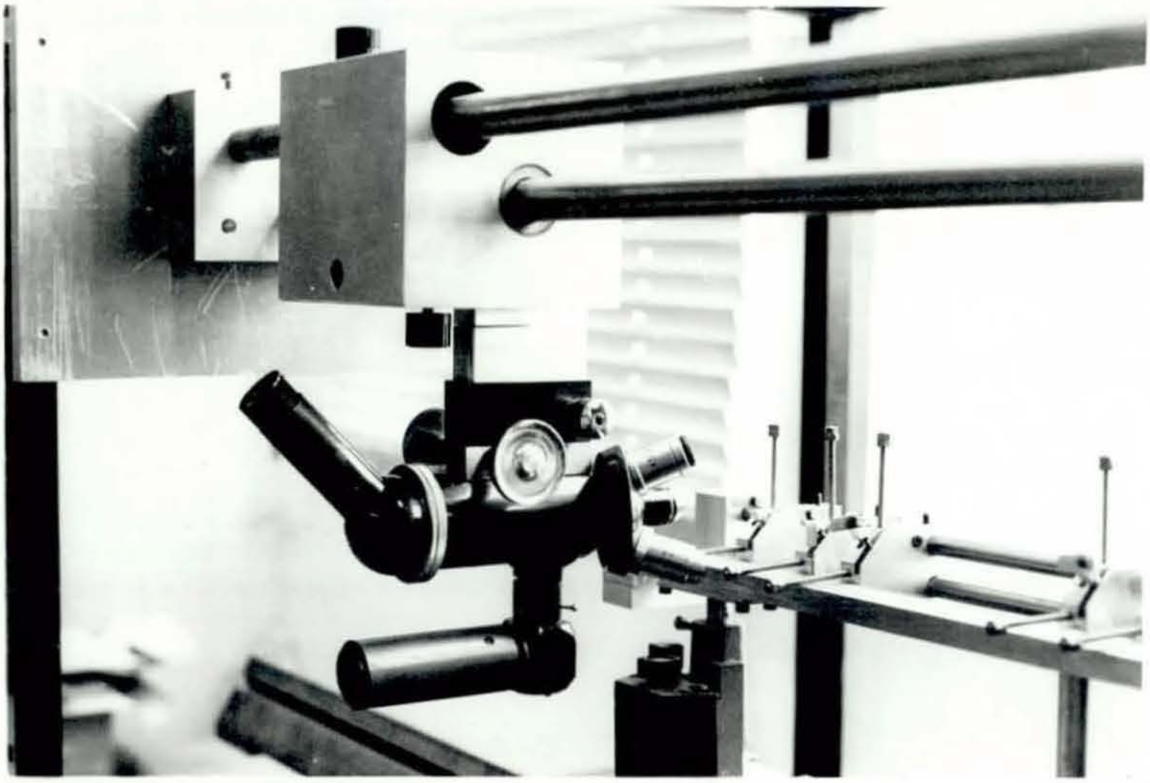


Figure 119. The travelling microscope.

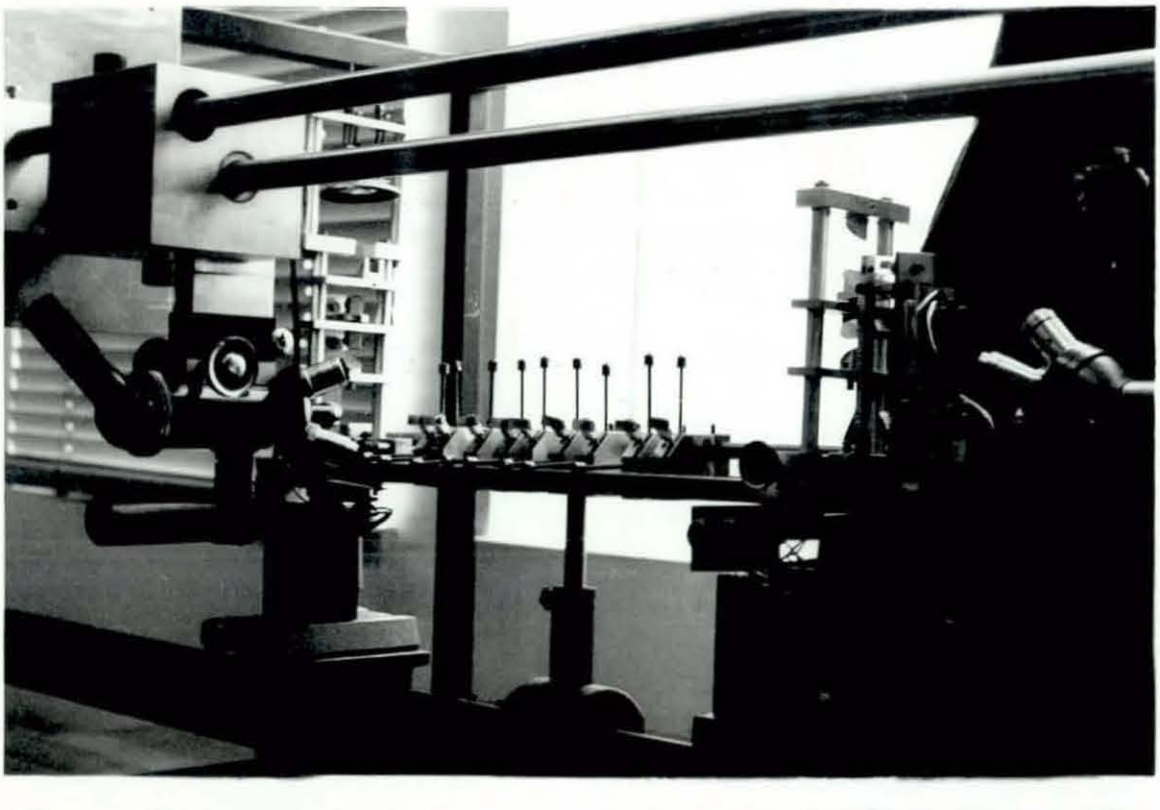


Figure 120. The heterodyne scanning system in position on the straight fibre mount.

The fibre supporting device which was subsequently designed and built is shown in Figure 117 and the detail of one of the supports is shown in Figure 118. The fibre is passed through the holes in the plates at the end of each support rod, although a better method would be to etch a slot in each plate to avoid damaging the fibre end during the threading process.

A travelling microscope was built for use with the straight fibre mount and this apparatus is shown in Figure 119. The scanning equipment and/or the microscope system were also designed to fit onto the straight fibre mount and these are shown in position in Figure 120.

7.5.1 Propagation of Skew Rays in Circular Cross Section Waveguides with Large Radius Bends.

It is convenient to define the radius R of a bend in a waveguide as

$$R = qa \qquad 354.$$

where a is the radius of the waveguide core and q is an arbitrary number > 100 for large radius bends.

A section of a waveguide with a constant radius bend is shown in Figure 121 where planes 1 and 2 are separated by the distance ℓ shown, where ℓ is the distance between adjacent points of reflection of a meridional ray propagating in a straight waveguide and is given by

$$\ell = \frac{2a}{\tan\theta_1} \qquad 355.$$

The axial angle of the meridional ray is θ_1. Plane 2 is rotated by an angle ε with respect to plane 1 and the projection of the centre of the waveguide in plane 2 onto plane 1 is shifted by an amount Δ from its position in plane 1. Using the equation given by reference (49 Eq.642)

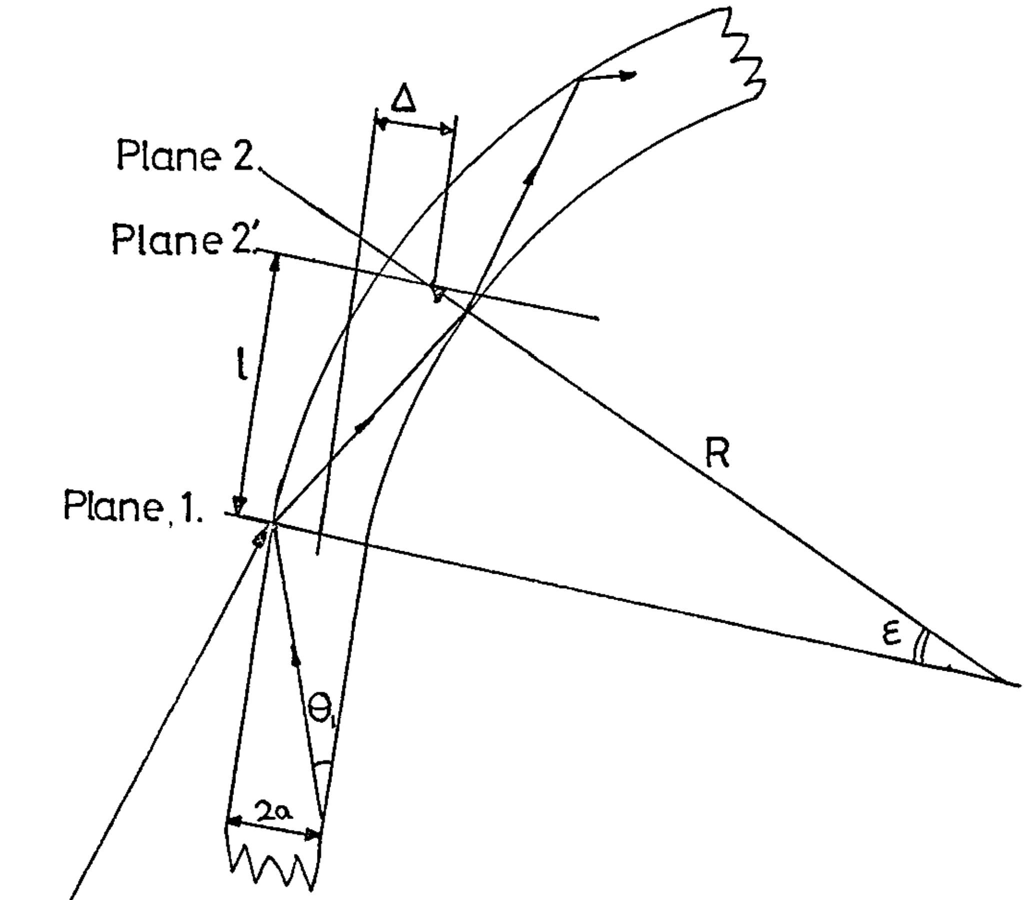


Figure 121. A constant radius bend waveguide.

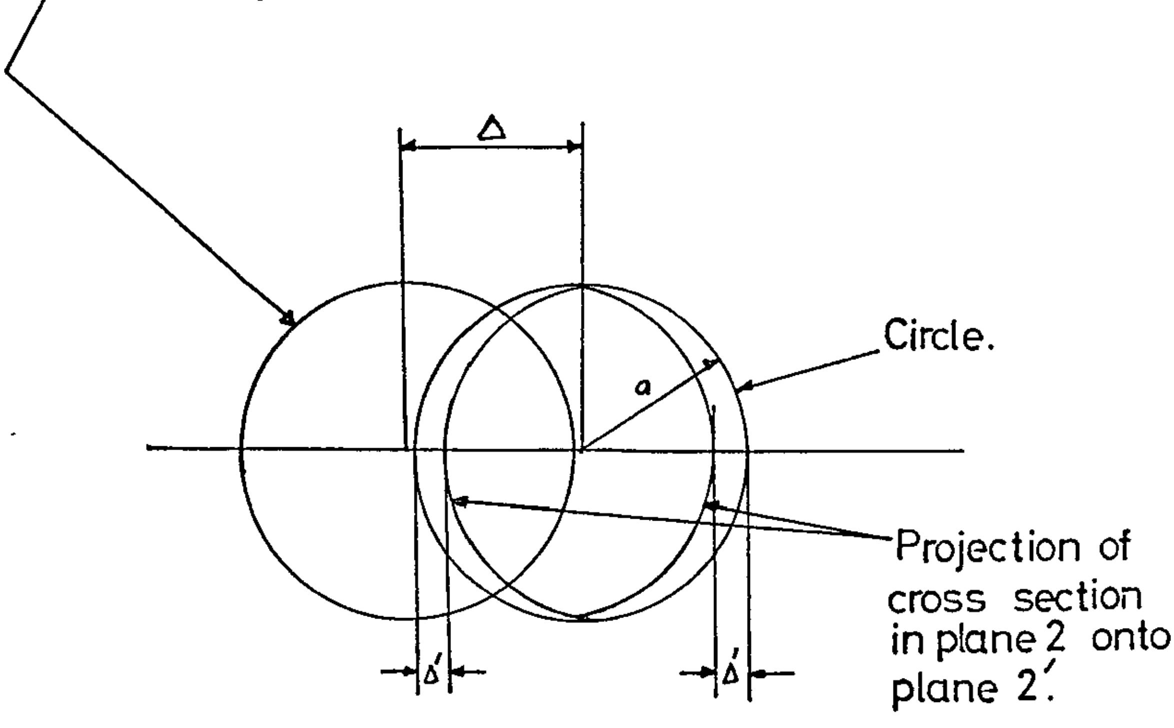


Figure 122. The projection of cross sections of bent waveguide onto a single cross section.

the relationship between R, ℓ and Δ is given by

$$R = \frac{\Delta}{2} + \frac{\ell^2}{2\Delta} \qquad 356.$$

Equation (356) forms a quadratic equation in terms of Δ, the solutions of which are

$$\Delta = R(1 \pm (1 - \frac{\ell^2}{R^2})^{\frac{1}{2}}) \qquad 357.$$

The sensible solution to Equation (357) is the negative solution which after substitution of Equations (354) and (355) and using the first two terms in the series expansion of

$$(1 - \frac{\ell^2}{R^2})^{\frac{1}{2}}$$

becomes

$$\Delta = \frac{2a}{q \tan^2 \theta_1} \qquad 358.$$

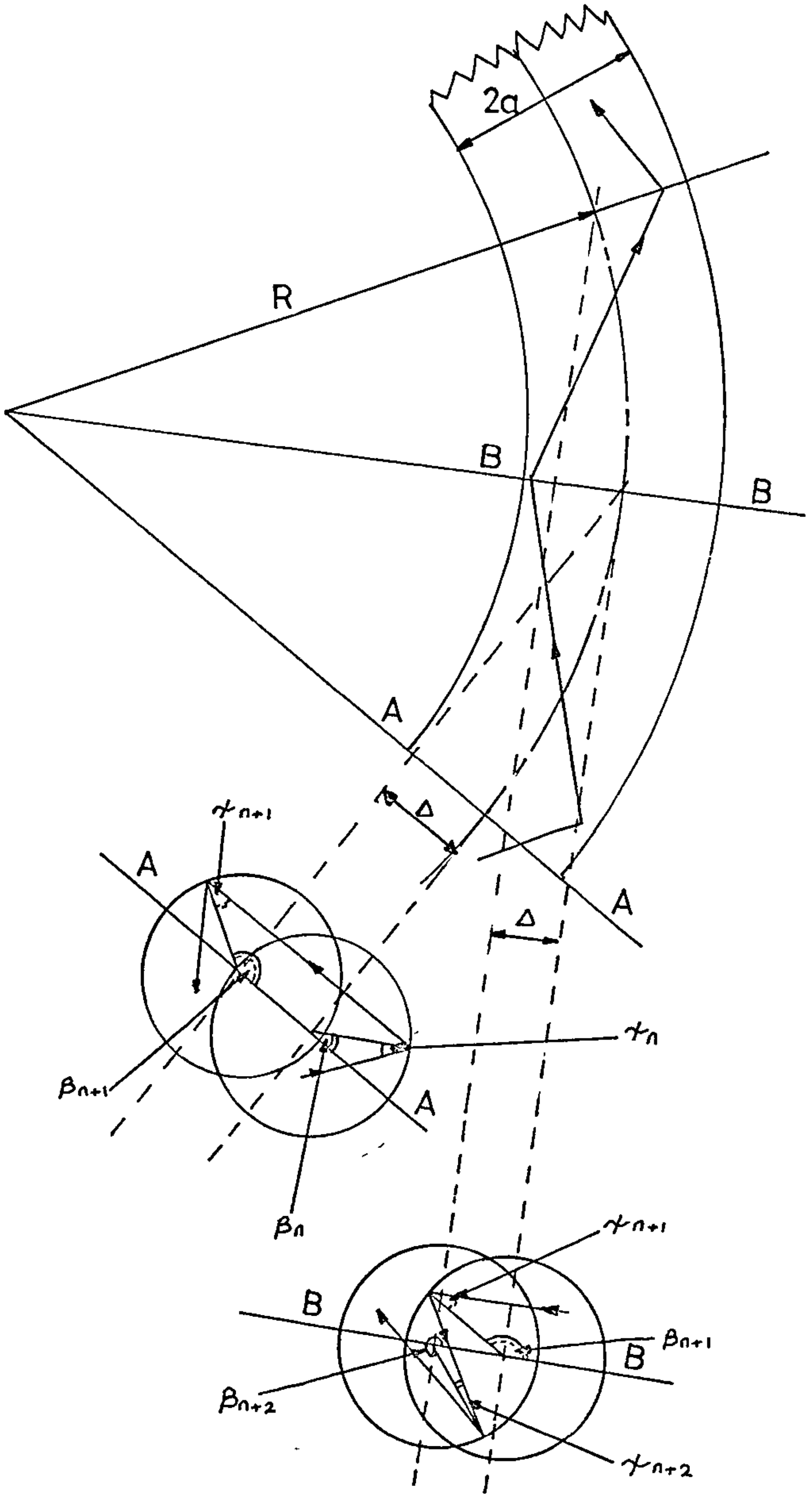
Assuming $R >> \ell$ the angle ε is given approximately by

$$\tan \varepsilon \simeq \frac{2}{q \tan \theta_1} \qquad 359.$$

If the waveguide cross section is assumed to be circular when it is projected onto plane 2', where plane 2' is parallel to plane 1, (See Figure 122) then the position of the walls of the waveguide will be displaced from their true position by a maximum of Δ' where

$$\Delta' \simeq a(1 - \cos \varepsilon) \qquad 360.$$

Figure 123a,b. The longitudinal section and the cross sections of a bent waveguide at the points of reflection of an arbitrary ray.



Substituting for cos ε using Equation (359) and assuming that $\sin\varepsilon \approx \tan\varepsilon$ Equation (360) becomes

$$\Delta' \approx \frac{2a}{q^2 \tan^2\theta_1} \qquad 361.$$

Since Δ' is a factor $\frac{1}{q}$ times smaller than Δ, it may be neglected.

In a rigorous analysis of a general ray propagating in a bent waveguide the variables of interest would be the axial and azimuthal angles of incidence of the ray at each of its reflection points. The variations in these angles will be a function of the initial position of the ray within the cross section (Reference 4), and the distance along the bent waveguide between adjacent reflections.

In the following simplified analysis, only the variations in the azimuthal angle of incidence will be calculated using the assumption that reflections occur at equal distances down the waveguide. This assumption results in a constant displacement Δ between the cross sections of the waveguide at adjacent reflection points and the variations in azimuthal angles of incidence may then be calculated by considering the path of rays in circular cross sections with centres displaced by Δ.

In Figures 123a and 123b the longitudinal and cross sections of an arbitrary ray propagating in a bent cylinder are shown. If a is the radius of the waveguide then the following relationships may be obtained from Figure 123b.

$$\sin\chi_n = \sin\chi_{n+1} + \frac{\Delta}{a}\sin(\beta_n - \chi_n) \qquad 362.$$

$$\sin\chi_{n+1} = \sin\chi_{n+2} + \frac{\Lambda}{a}\sin(\chi_n + 2\chi_{n+1} - \beta_n) \qquad 363.$$

$$\beta_{n+1} = \pi + \beta_n - \chi_n - \chi_{n+1} \qquad 364.$$

$$\beta_{n+2} = 2\pi + \beta_n - \chi_n - 2\chi_{n+1} - \chi_{n+2} \qquad 365.$$

By cross substitution between Equations (362)-(363) and (364)-(365) the following equations are obtained, where χ_n, χ_{n+1}, χ_{n+2} are all assumed to be small so that the sines may be replaced by the angles (in radians).

$$\chi_{n+2} = \chi_n - \frac{\Delta}{a}(\sin(\beta_n - \chi_n) - \sin(\beta_n - \chi_n - 2\chi_{n+1})) \qquad 366.$$

$$\beta_{n+2} = 2\pi + \beta_n - 4\chi_n + \frac{\Delta}{a}(3\sin(\beta_n - \chi_n) - \sin(\beta_n - 3\chi_n)) \qquad 367.$$

Equation (366) may be expanded and restated in the form of a difference equation

$$\Delta\chi = \chi_n - \chi_{n+2} = \frac{2\Delta}{a} \chi_n \cos(\beta_n - 2\chi_n) \qquad 368.$$

Similarly Equation (367) gives

$$\Delta\beta = \beta_n \quad \beta_{n+2} = 4\chi_n - \frac{2\Delta}{a} \sin\beta_n \qquad 369$$

It will be instructive to tabulate the corresponding difference equations for rays propagating in circles and ellipses as discussed in Chapter 6, and to compare them with Equations (368) and (369) above.

	$\Delta\chi$	$\Delta\beta$
CIRCLE	0	$4\chi_n$
ELLIPSE	$4\gamma_n$	$4\chi_n - 12\gamma_n$

Table 13.

In the above table it has been assumed that $\gamma_n \approx \gamma_{n+1}$ where $\gamma_n = \frac{h}{2} \sin 2\beta_n$. To detect the visible effects of a bend it is likely that the coefficients $\frac{2\Delta}{a}\chi$, $\frac{2\Delta}{a}$ in the difference equations for offset circles will be at least of the same order as the coefficients which cause visible effects in an elliptical cross section waveguide.

Setting these two sets of coefficients equal gives for the $\Delta\chi$ difference equation

$$h = \frac{\Delta}{a}\chi_n \tag{370.}$$

and for the $\Delta\beta$ difference equation

$$3h = \frac{\Delta}{a} \tag{371.}$$

Substituting for Δ using Equation (358) gives

$$h = \frac{2\chi_n}{q \tan^2\theta_1} \quad \text{372a. or} \quad 3h = \frac{2}{q \tan^2\theta_1} \tag{372b.}$$

If θ_1 is taken to be 18° so that $\tan^2 \theta_1 \approx 0.1$ and if $\chi_n \approx 0.2$ it is found that the radius of bend R which gives coefficients of magnitudes similar to those for an elliptical deformation of the order of 2% ($h = .04$) is

$$R = 100a$$

As θ_1 is reduced this bend radius will increase but the larger values of θ_1 are of primary interest because they produce the most visible elliptical caustics in long lengths (> 10 cm.) of waveguide.

7.5.2 Experimental Results

The following observations were made in order to test the above propositions. A 40 cm. length of single fibre was illuminated in such a manner as to produce a caustic of the type shown in Figure 5b . Care was taken to ensure that both ends of the fibre were securely fixed, and a bend was inserted about the middle of the length of fibre whilst observing the caustic.

When $\theta_1 \simeq 18^\circ$ no changes were observed in the caustic until the radius of the bend was of the order of 50a . A reduction of θ_1 to 10° produced changes in the caustic when the radius of the bend was 200a . No attempt was made to analyse the precise nature of the changes produced in the caustic by the introduction of the bends and so the changes observed are not described here.

7.6 Conclusions

This chapter has outlined aspects of propagation in optical waveguides which require further study.

It may be that if the visibility of the caustics observed within the waveguides could be related to the scattering which is occurring, the observation of the caustics could form a useful scattering measurement technique. The simplicity of the experimental arrangements required to view the caustics suggests that such a measurement technique may find application in production control environments.

The heterodyne scanning system clearly requires more experimental work before an assessment of its usefulness can be made. The straight fibre mounting technique may be of use for measuring the microbending which has been suggested may occur in optical waveguides(51).

Theoretical studies[52] have been made of the effects on signal propagation of some forms of microbending but little experimental work has been reported on the measurement of microbending.

The behaviour of light in bent waveguides has been studied[53] from the signal distortion aspect. A bend in a waveguide causes 'mode mixing', which may be a desirable feature in certain waveguide systems in order to minimise pulse distortion. Further study of the visible effects of bends in waveguides may confirm the accuracy of the mathematical representations of light propagation in bent waveguides.

CHAPTER 8

8.1 Introduction

In his early work on optical waveguides Kapany[3] introduces a waveguide characteristic term R where

$$R = \frac{2\pi a}{\lambda} (n_1{}^2 - n_2{}^2)^{\frac{1}{2}} \qquad 373.$$

and R is a function of the physics and geometry of the waveguide only. Its value indicates the maximum number of waveguide modes which may propagate in the waveguide.

In later work[54,55] by other authors this term is called normalised frequency and is closely related to the arguments of the Bessel functions used to describe the modal field distributions.

In this chapter the waveguide characteristic term is derived from consideration of the fringe system formed by a skewplane which has undergone many reflections. This fringe system and that developed for the meridional skew plane are shown to be similar to the waveguide mode patterns of the cladded dielectric waveguide.

The trapped skew planes of an elliptical waveguide produce fringe systems which are shown to be similar to the fringe systems found in certain resonant cavities.

The relationship between the visibility of caustics within optical waveguides and the spatial coherence of the source is examined and it is found (experimentally) that a spatially incoherent monochromatic source produces effects similar to those produced by a polychromatic source.

The final topic to be considered in this thesis is the apparent increase in coherence observed within the waveguide when propagation occurs at high axial angles of incidence. The increase in coherence

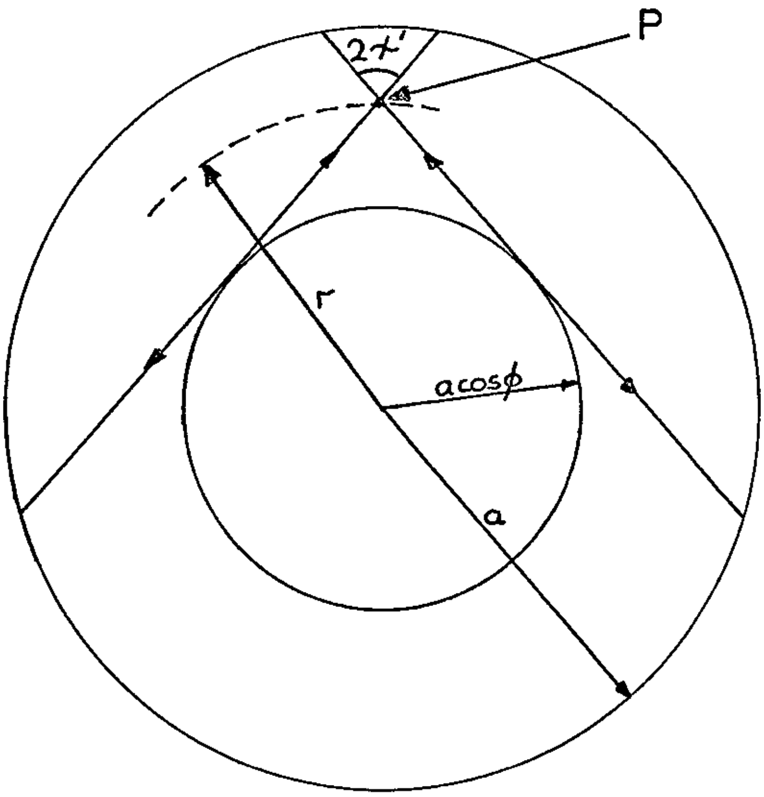


Figure 124. The cross section of a waveguide showing the four skew planes which contribute to the light at point P.

is suggested by the experimental observation of well defined, multifrequency fringes when the waveguide is illuminated by white light.

8.2 Interference Fringes in Circular Dielectric Waveguides

It was shown in Chapter 5 that the light initially contained within a single skew plane will, after many reflections, illuminate an annulus, inner radius a cos ϕ, outer radius a. It was also shown that the inner radius of the annulus represents the position of the caustic of the rays of light illuminating the annulus. The diffraction pattern produced by the caustic of the rays forms concentric fringes in the cross section of the waveguide with dimensions governed by the fringe factor F.

Consider the rays of light which may contribute to the disturbance at point P shown in Figure 124, where P is a radial distance r from the centre of the waveguide and $(a \cos \phi) < r < a$. If a plane wave at an axial angle of incidence θ_o illuminates the entrance aperture of the waveguide then two skew planes are excited for each value of ϕ, the positive ϕ skew plane propagates in the anticlockwise direction and that for $-\phi$ in the clockwise direction. Thus at point P there may be four rays of light, each tangent to the circle radius a cos ϕ, representing the contribution of the $\pm\phi$ skew planes to the disturbance at point P.

The two rays proceeding towards the circumference of the waveguide are inclined to each other in the cross sectional projection at an angle $2\chi'$ where

$$\sin \chi' = \frac{a \cos \phi}{r} \qquad 374.$$

The projection of the free space wavelength λ onto the cross sectional plane of the waveguide will give an equivalent wavelength λ' where

$$\lambda' = \frac{\lambda}{n_1 \sin \theta_1} \qquad 375.$$

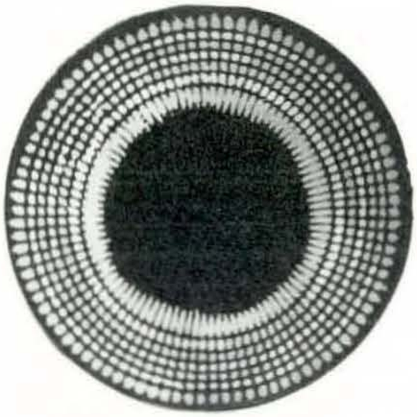
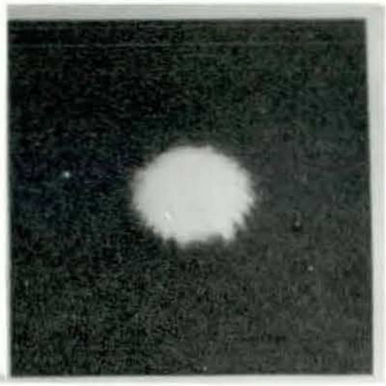
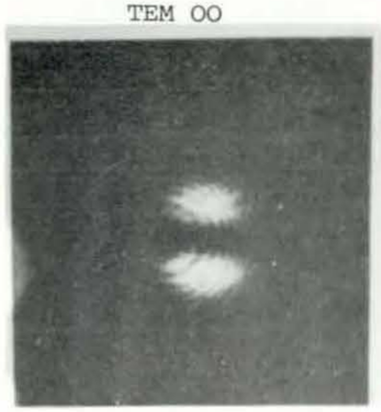


Figure 125



TEM OO



TEM 10

Figure 126. Low order modes of a typical laser cavity.

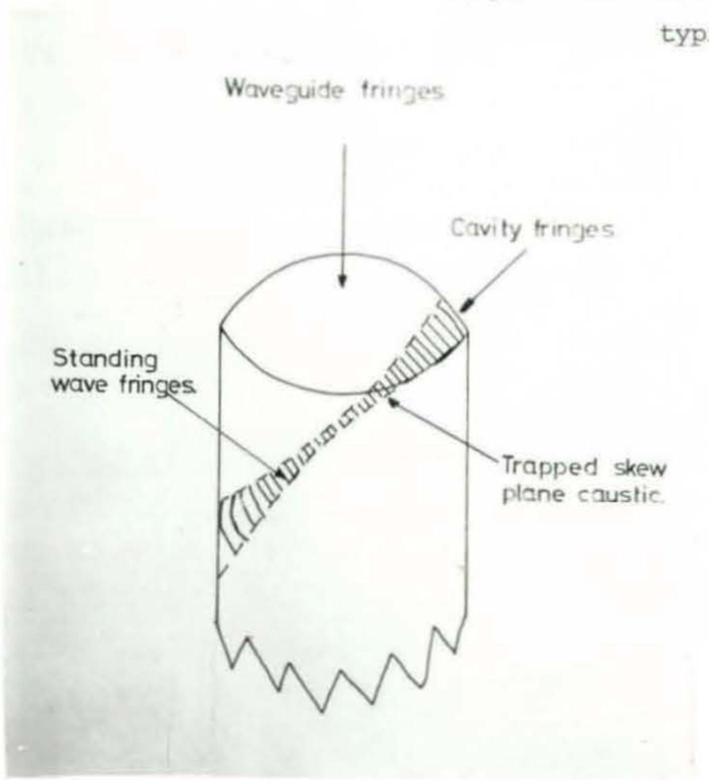


Figure 127. The relationship between waveguide modes and resonator cavity modes.

and

$$\sin \theta_1 = \frac{n_o}{n_1} \sin \theta_o . \qquad 376.$$

If the two rays are considered as plane waves at their point of intersection they will form an interference pattern similar to that for the plane waves intersecting as shown in Figure 26. Following the derivations given in Chapter 3 the spacing t between fringes will be given by Equation (43) by substituting for λ' from Equation (375) and for χ' from Equation (374) to give

$$t = \frac{\lambda r}{2 \sin \theta_1 \, a \cos \phi \, n_1} . \qquad 377.$$

Since similar interference patterns will be produced at all points around the circle, radius r, it will be assumed that a single continuous fringe system is formed where the total number of fringes formed around the circle radius r will be given by L where

$$L = \frac{2\pi r}{t} \qquad 378.$$

which on substitution for t from Equation (377) gives

$$L = \frac{4a\pi \sin \theta_1 \cos \phi}{\lambda} . \qquad 379.$$

Since L is now independent of the radius the same number of fringes will be formed at all radii within the annulus. This suggests that the combination of this interference pattern and that due to the caustic diffraction pattern will produce an interference pattern of the form shown in Figure 125.

The azimuthal and radial intensity distributions shown in Figure 125 may be denoted by $I\alpha$ and Ir respectively. Producing these intensity variations will be variations in electric field components which may be denoted as $E\alpha$ and Er where

$$I\alpha = \overline{E\alpha^2} \qquad 380a. \qquad Ir - \overline{Er^2} \qquad 380b.$$

The radial electric field Er will be given by the Airy function, Equation (197), with $v = r$ and its associated intensity function will be given by Equation (199).

With prior knowledge of the form of the modal fields obtained from a rigorous analysis of dielectric waveguides, it is assumed that the azimuthal electric field variations are of the form

$$E\alpha = A \sin (P\alpha) \qquad 381.$$

where P is an integer.

The corresponding intensity function is given by

$$I\alpha = \overline{E\alpha^2} = \overline{A^2 \sin^2(P\alpha)} \qquad 382.$$

and will form a total of 2P fringes per revolution of α. Since Equation (379) also represents the number of fringes per revolution, P and L are related by the expression

$$2P = L \cdot \qquad 383.$$

P may be expressed in the form

$$P = \frac{L}{2} = \frac{2a\pi \sin \theta_1 \cos \phi}{\lambda} , \qquad 384.$$

and will have a maximum value when $\cos \phi = 1$ and $\theta_1 = \theta_{1c}$ where θ_{1c} is the maximum value of θ_1 and is called the meridional critical angle.

Denoting the maximum value of P as P_{max} setting $\cos \phi = 1$ and using Equation (127) for θ_{1c} then

$$P_{max} = \frac{2a\pi}{\lambda} \left(n_1^2 - n_2^2\right)^{\frac{1}{2}} \qquad 385.$$

This value of P represents half the maximum number of fringes which may be formed at a radius r = a in the waveguide, when the skew planes corresponding to $\cos \phi = 1$ are illuminated at the maximum axial angle of incidence, θ_{1c}. P_{max} and R (Equation (373) are clearly equivalent.

In Chapter 2 of Reference (2) Marcuse shows that a modal field of a cladded dielectric waveguide may be represented by the superposition of four plane waves travelling in directions similar to those described above for a single skew plane ray. Further the description of the near field modal patterns given by Kapany[1] for the cladded dielectric waveguide results in patterns similar to that shown in Figure (125). These two results suggest that the geometrical approach to the problem of finding the waveguide modes of the cladded dielectric waveguide may produce useful results.

When only meridional rays are excited ($\phi = \frac{\pi}{2}$), there will be no azimuthal fringes produced, only radial variations in intensity, a distribution which corresponds to the waveguide mode family called HE 1m.

A detailed attempt to derive the waveguide modes of the cladded dielectric waveguide using the geometrical approach would have to consider the effects of the following assumption which have been made in deriving the fringe patterns described earlier in this section.

In general there will be a phase shift of field components on reflection at the core cladding interface. This phase shift will depend upon the polarisation of the incident fields, and may produce finite values of field components on both sides of the interface. Having considered the effects of these phase shifts on the fringe systems, the waveguide modes could be determined using the condition that a fringe system belonging to a given mode would have to satisfy the boundary conditions at the core cladding interface and also have an integral number of fringes in the azimuthal and radial directions. These conditions are exactly those described by Marcuse in his simplified approach to the eigenvalue equation for the assymetric dielectric waveguide given in reference (2).

8.3 Interference Fringes in Elliptical Dielectric Waveguides

In Section 8.2 it was shown that interference fringes are formed parallel and normal to geometrical caustics, (radial and azimuthal variations of intensity respectively). Similar fringes are observed for the trapped skew plane ray caustics produced in an elliptical cross section waveguide.

The diffraction pattern of a trapped skew plane caustic has a fringe spacing determined by the fringe factor as before and the fringes will be formed parallel to the caustic.

It was mentioned in Chapter 6 that the trapped skew planes were trapped within a form of resonant cavity and it is therefore likely that the mode patterns formed in resonant cavities will be similar to the fringe systems observed within the waveguides.

The mode patterns formed in spherical reflector resonant cavities have been extensively studied[56,57] typical low order mode patterns are shown in Figure 126. The relationship between the fringe systems formed in spherical reflector resonant cavities and those observed within a waveguide is shown diagramatically in Figure 127. It appears from Figure 127 that the caustic fringe patterns correspond to the normally observed resonant cavity mode patterns. There is a second set of fringes formed in spherical cavities as a result of the standing waves set up between the two reflectors. In the elliptical waveguide both the 'standing wave' and the 'mode pattern' fringe systems are visible as shown in Figure 133a.

8.4 Spatial Coherence and Caustic Visibility

In Chapter 3 it was shown that the coherence of a source was measured by observing the contrast of interference fringes formed by the superposition of light either from different areas of the source (to measure spatial coherence) or from the same area of the source but with beams of light arriving at different times (to measure temporal coherence).

In this section the influence of the propagation of light in a waveguide upon its coherence characteristics will be briefly examined. The analysis is based upon experimental observations of fringe contrast and caustic visibility under varying coherence illuminating conditions.

The majority of the illustrations of caustics in this thesis were obtained using a white light source which as noted in Chapter 3 has the shortest coherence length of the sources used in this study. When the caustic fringe factor F is small, the different frequency components of the white light source will have maxima of intensity close to the geometrical caustic. These components will add incoherently

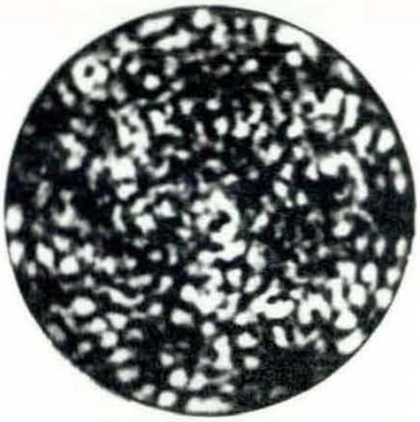


Figure 128.

Figure 129.

Figure 130a.

Figure 130b.

Figure 130c.

to produce a region of white light of higher intensity than the surrounding region and which will approximately correspond to the position of the geometrical caustic. The remaining multifrequency fringes of the caustic fringe pattern will have different spatial positions which, depending upon the frequency dependence of the fringe factor F, may overlap to produce no visible fringes, (incoherent illumination) or may produce white light fringes as shown in Figure 65 (coherent illumination).

The caustics produced in circular cross section waveguides are all circular and adjacent skew planes form adjacent caustics. The fringe patterns associated with each skew plane ray caustic will be similar and since they are immediately adjacent to their neighbours the fringe patterns will overlap and produce nominally 'uniform white illumination' across the cross section of the waveguide. The absence of interference fringes is the experimental result which leads to the statements made by Kapany and Burke referred to in Chapter 1.

If a laser source is used to illuminate the waveguide with plane waves then apparently randomly positioned interference fringes are produced as shown for example in Figure 128 where $\theta_o = 10^o$. The size of the fringes is inversely proportional to θ_o as demonstrated by comparing the size of the fringes shown in Figure 129, where $\theta_o = 20^o$ and the fringes shown in Figure 128. By careful inspection of these interference patterns it is possible to observe incomplete concentric fringes similar to the radial variations of intensity described in Section 8.2 The spacing of these concentric fringes is also inversely dependent upon θ_o.

It was found that the visibility of these concentric fringes could be improved if the laser was moved so that θ_o was oscillated about its mean value by $\sim \pm 2^o$. This had the effect of moving some of the fringes very rapidly about the cross section of the waveguide which made them disappear, leaving complete concentric fringes visible although of lesser contrast than with the laser stationary. The occurrence of this phenomenon indicates that the concentric fringes have a smaller dependence upon θ_o for their position than the remaining apparently randomly positioned fringes.

The fading of the fringes described above may be interpreted as a reduction in coherence of the source, and the technique has certain similarities with speckle pattern interferometry. In speckle pattern interferometry[58,59] the wavelength of the source is changed so that the superposition of random speckle patterns will reveal variations in surface geometry of the illuminated object. In these experiments the axial angle of incidence of the illumination is varied so that random fringes appear to be incoherent because of their rapidly changing position and so enhance the visibility of the concentric fringes which move relatively slowly.

Another method of changing the coherence of a laser source was described in Chapter 3, where the spatial coherence of the laser was confined within an area whose dimensions are determined by the relative positions of a rotating ground glass screen and the plane of observation. This variable coherence arrangement, shown in Figure 19, was used to illuminate a 40 cm. length of fibre bundle and where the plane of observation is taken to be the entrance apertures of the fibres which are then a distance R from the laser point source.

When $R_s << R$ (where R_s is the distance of the rotating ground glass screen from the laser point source) the interference pattern observed at the radiation end of the fibres was the same as that obtained when the fibres were illuminated with a stationary laser. However, as R_s approached R the fringes at the circumference of each fibre core began to fade leaving a uniform illumination at the wavelength of the laser light (red light for the He-Ne laser used here). This effect moved progressively towards the centre of each fibre until the rotating ground glass screen was within 1 mm of the entrance apertures of the fibres. The interference fringes obtained under these conditions are shown in Figure 130a and may be compared with those shown in Figures 130b and 130c where the illumination is at the same angles of incidence but with stationary laser light and white light respectively.

It would appear that the minimum spatial coherence illumination condition, $R_s \simeq R$, results in an intensity distribution within the waveguide cross section after propagation in 40 cm of waveguide similar to that obtained with white light illumination.

The radial dependence of the fading of the fringes is related to the contribution made by each skew plane to the cross sectional intensity distribution in the waveguide as described in Chapter 5. It was shown there that after many reflections each skew plane ray will illuminate an annulus inner radius a cos ϕ outer radius a. The effects of a reduction in the spatial coherence of the light within the entrance aperture of the waveguide will appear first where the light from skew plane rays furthest apart (i.e. $\phi = \frac{\pi}{2}$, $\phi = 0$) in the entrance aperture superimpose, which must be at the circumference of the waveguide core. The centre of the waveguide is illuminated only by the skew plane ray with $\phi = \frac{\pi}{2}$ and since it is not possible to make the light within this single skew plane incoherent (following the finite volume theory of electromagnetic waves given in Chapter 3) the centre of the waveguide is always coherently illuminated. This proposition is supported by the presence of interference fringes at the centre of the waveguide as shown in Figures 130c and 133a even though the illumination is white light which is both spatially and temporally incoherent (after propagation in 40 cms of waveguide).

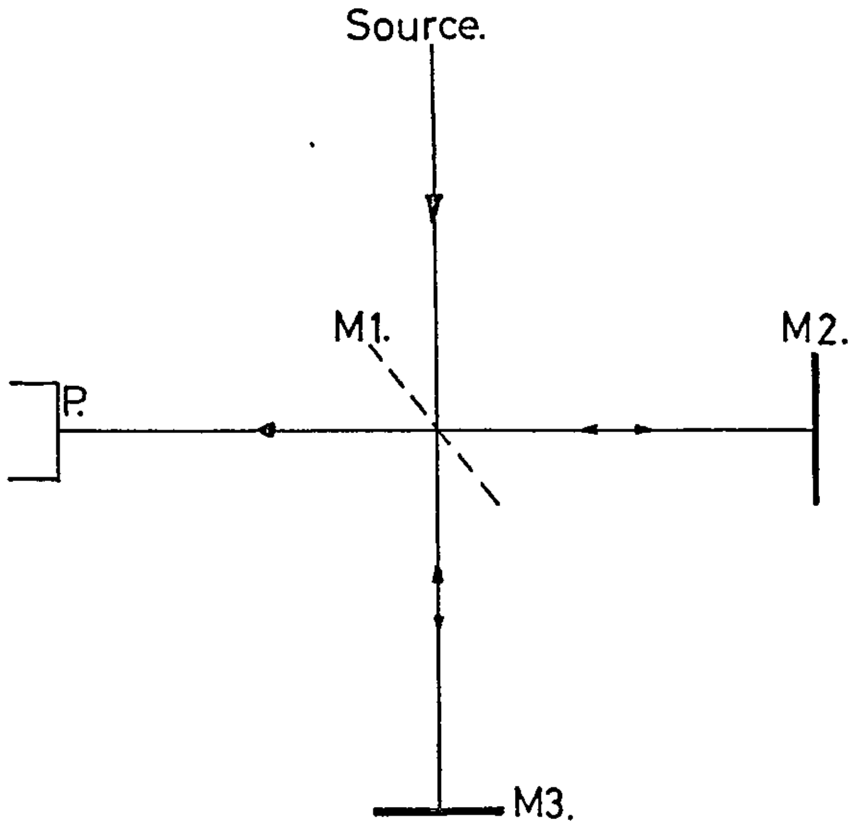


Figure 131. A dıagram of the Mıchelson ınterferometer.

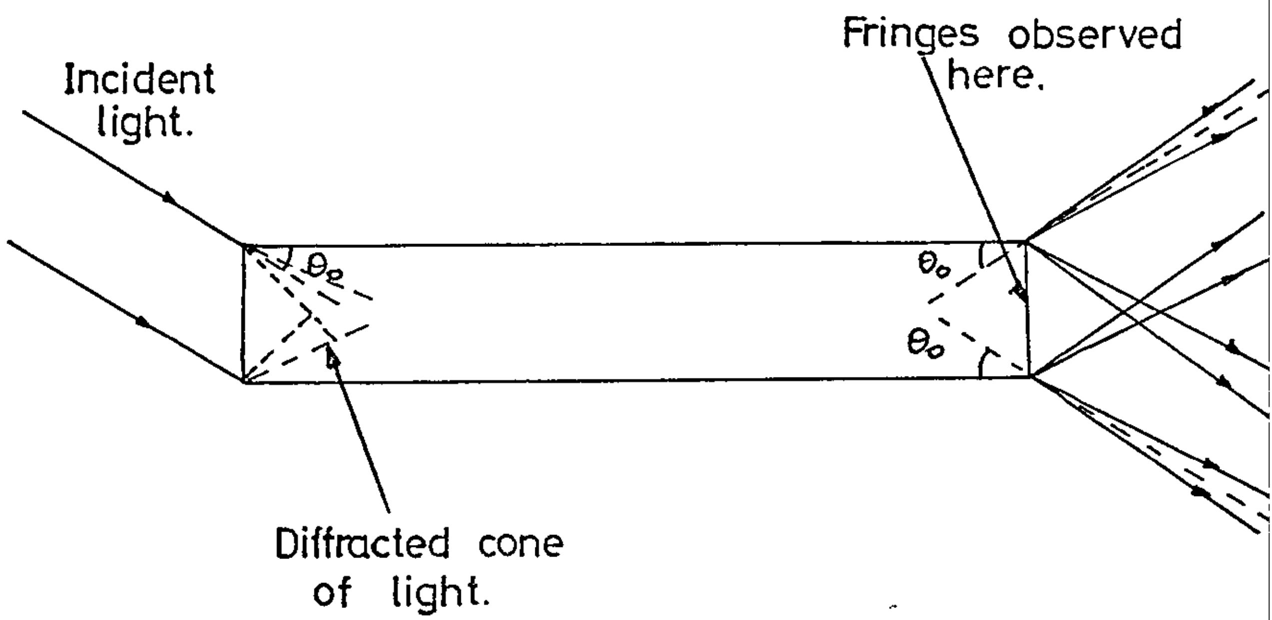


Figure 132. Interference in a dielectric cladded waveguıde.

8.5 Enhancement of Coherence by Propagation in Elliptical Waveguides

The short coherence length of white light may be demonstrated using a Michelson interferometer[31] which is shown diagrammatically in Figure 131. The light from the white light source is divided in amplitude by the half silvered mirror M1 and each half of the beam is totally reflected by mirrors M2 and M3. The two reflected beams are then superimposed and observed at point P. When the optical distances S1 and S2 differ by less than the coherence length of the white light ($\simeq 3\mu$) a set of fringes will appear to be formed at P, the centre fringe being white and the adjacent fringes coloured.

Consider now the experimental arrangement shown in Figure 132 where a 40 cm length of cladded optical waveguide, of elliptical cross section, is illuminated by plane waves from a white light source. Wnen the angle of incidence θ_o of the plane waves is $\simeq 24^o$ a set of white fringes as shown in Figure 133a are formed at right angles to the quasi-hyperbolic caustic as described in Section 8.2 If θ_o is increased these fringes fade and reappear when $\theta_o \simeq 32^o$ but now with a smaller fringe spacing, as shown in Figure 133b.

The formation of fringes is to be expected when the illumination is monochromatic since as suggested in Section 8.2, the elliptical cross section forms a resonant cavity and these fringes are the standing waves set up between two reflectors. It is surprising, however, that fringes should appear when the illumination is white light, since the difference in path length for adjacent reflections of the plane waves radiating from the waveguide is of the order of 12 microns when $\theta_o \simeq 24^o$ and 16 microns when $\theta_o \simeq 32^o$. If it is assumed that the plane waves suffer diffraction at the entrance aperture as if they were normally incident (in fact they will suffer greater diffraction than this) then the plane wave will be

Figure 133a.

Figure 133b.

formed into a cone of light of semi angle given by Equation (198). After propagation in 40 cms of waveguide at an incident angle of $\theta_o \simeq 24^o$, this cone will form into approximately sixty beams of light superimposed in the radiation aperture to give the interference fringes observed. The number of beams increases as θ_o is increased.

Clearly these fringes are not formed by the same process as in the Michelson interferometer since the difference in optical distance from the entrance aperture to the point of superposition of adjacent beams exceeds the coherence length of the white light and further the fringes formed in the waveguide are all white. A possible explanation is offered in part by Streifer[41] in his concluding remarks in which he states "The methods described ..... apply to waveguides or dielectric rods which allow a mode description of electromagnetic wave propagation. The "mode selective" (my italics) properties of such systems could act to produce virtually complete coherence from initially incoherent radiation".

Each monochromatic component of the white light beam is likely to be coherent within the trapped skew plane caustic because of the small aperture the trapped skew planes present within the entrance aperture of the waveguide. The coherence of the light within a small number of skew planes follows from the comments made in the latter part of Section 8.3 It may be that the mode selection which occurs in the resonant type of structure in which the trapped skew planes propagate, results in mode patterns which are insensitive to frequency when the illumination is incident on the waveguide at certain angles. Further work is required to examine these propositions.

8.6 Conclusions

Various phenomena which were observed during the experimental investigation of the formation of caustics within cladded optical waveguides have been reported in this chapter. The explanations offered for the occurrence of the phenomena are not rigorous but are theories which could form the basis for future study.

Perhaps the most interesting topic for further study is an investigation of the coherence of the radiation field of the optical waveguides under varying spatial coherence illuminating conditions. It may be possible to obtain information from such measurements about scattering within the waveguides and the effects of bends and micro-bending of the waveguide, and it was for this purpose that the heterodyne scanning system described in Chapter 7 was built.

CHAPTER 9

Conclusions

The major objective of the experimental investigation described in this thesis was to find the cause of the 'patterns' produced in cladded optical waveguides under certain conditions of illumination. It has been shown by computer simulation and analysis of ray propagation in cladded waveguides that the 'patterns' are caustics produced by non-circular cross section waveguides. The propagation of rays in an elliptical cross section waveguide has been studied in detail and experimental results have illustrated the main results of this analysis.

The diffraction at the entrance aperture of a cladded waveguide has been investigated and experimental results confirm that the waveguide appears to cause diffraction of the incident light field as if the waveguide aperture were a simple pinhole. The experimental results also suggested a non-stepped refractive index profile between core and cladding glasses and this was further investigated. The variation of illuminated core diameter and the 'ring effect' are both results of a non-stepped refractive index profile but neither method permits accurate determination of the profile.

The thick lens behaviour of a cladded waveguide was investigated A novel use of such a lens in an optical communications system was reported recently[60] where an uncladded fibre was used to increase the light collecting property of a small diameter core waveguide. It may be that improvements in the performance of this system can be achieved by using a cladded waveguide lens where the refractive indices of the core and cladding glasses and that of the surrounding media are selected to give a cylindrical lens with a long focal length. A long focal length lens will launch light into the transmission waveguide with a small conical semi angle and this will reduce the temporal dispersion of the transmitted signal since the dispersion due to the waveguide geometry is proportional to the axial angle of incidence of the light.

The analysis of ray propagation in circular cross section waveguides confirms the 'uniform cone' property of large diameter waveguides when they are illuminated at a single azimuthal angle of incidence. The use of this radiation cone in the alignment of waveguides prior to polishing has been described. The effects of a sloping (but flat) end face on the radiation cone suggested a method of progressively normalising this slope when it appears at the radiation end of the waveguide. The ray analysis also predicts the 'black hole' effect which was used for alignment of waveguides and as a measure of the axial angle of propagation of light within the waveguide.

The use of a visual computer simulation of the propagation of rays has been particularly successful when applied to ray propagation in elliptical cross section systems where the imagination is unable to supply the necessary 'mental picture' of the transition from trapped to non-trapped modes of ray propagation.

The results of the analysis of higher order stationary skew plane ray systems within an elliptical cross section waveguide suggested that the observed variations of intensity which formed triangles, squares and other regular multi-sided figures are produced by non-circular, non-elliptical cross section waveguides. This suggests that a precise determination of the cross sectional geometry which produces these patterns would require computer simulation of ray propagation within a generalised cross section system.

This means that instead of proposing specific cross section geometries and then concluding from the simulation which patterns are likely to occur within real waveguides, the patterns observed within any specific waveguide would be used as the initial conditions for the computer simulation and the cross section of the computer program representation of the waveguide will be modified until it too predicts a similar arrangement of patterns.

A major difficulty with such a proposal is that there is no alternative method of measuring the waveguide cross sections within the variations of dimensions which have been shown may produce patterns.

There are, however, alternative methods for measuring the scattering within optical waveguides and this suggests that the observation of patterns may be developed into a method of measuring scattering and microbending if the magnitude of these properties can be shown to be directly related to the visibility of the caustic patterns. Here it is assumed that because of the sensitivity of the caustic formation system, all cladded waveguides will exhibit caustics related to non-circular cross sections.

Finally the observations made in Chapter 8 on the coherence of the light propagating in the cladded waveguides suggests that optical waveguides may be a useful experimental media in the study of the enhancement of coherence by propagation in bounded media. This subject is of significance in the design of laser resonator cavities[39].

APPENDIX A.

Loss of Fibre Optic Cables

In the paper* 'Design and performance of Optical Fibre Cables' presented at the International Conference on Optical Communications held in London in September 1975, T. Nakahara et al describe results of transmission loss measurements on cladded fibre cables. The minimum loss measurement made by them is quoted as 1.6 dB/km.

* Design and performance of optical fibre cables, T. Nakahara et al, I.E.E. Conference Publication No. 132, Optical Fibre Communication 1975., pp 81.

APPENDIX B.

The following two pages are copies of the manufacturer's specification for the fibre optic cable used in the experiments described in this thesis.

Specification No MD 690

This specification is chiefly concerned with Glass Fibre Optic components and sub-assemblies incorporating **non-coherent** Fibre Optics in both solid and flexible form

1. DIMENSIONS AND TOLERANCES

Dimensions and tolerances will be specified on the drawing

2. MECHANICAL QUALITY

All items will be made to our normal standards of engineering quality and finish, and will generally have a good appearance

3. OPTICAL QUALITY

3.1 Light Transmission This will be as shown on Graph 1, and will be better than the lower line on this graph The shaded area shown will be the possible variation of the transmission and will take account of packing factor, broken fibres, interstitial losses, Fresnel reflections and optical face polishing

3.2 Fibre Packing The fibre packing within any optical area will be better than 80% and any dark area will be less than 0 6 mm dia

3.3 Broken Fibres A small percentage of broken fibres are present in most Fibre Optics The percentage of broken fibres is not specified, but will never be so large as to prevent the component meeting specification 3 1 for light transmission.

3.4 Optical Faces All optical faces will be ground and polished to optimise the light transmission and to ensure that the item meets specification 3 1 for light transmission

3.5 Fibre Size Nominal fibre diameter will be stated on the drawing

3.6 Numerical Aperture This will be approximately 0 54* unless otherwise specified A typical polar diagram is shown on Graph 2

3.7 Temperature Range In general components of standard design are capable of withstanding temperatures from −20°C to +105°C* for long periods without deterioration Standard components will withstand temperatures outside these limits for shorter periods (temperature requirements should be specified if outside the standard limits above)

3.8 For flexible Fibre Optic components it is suggested that the minimum bend radius is not exceeded or damage to fibres may occur, and light transmission will be reduced The recommended minimum bending radius for a flexible Fibre Optic unit depends on **The Optical Diameter** as shown in the table below

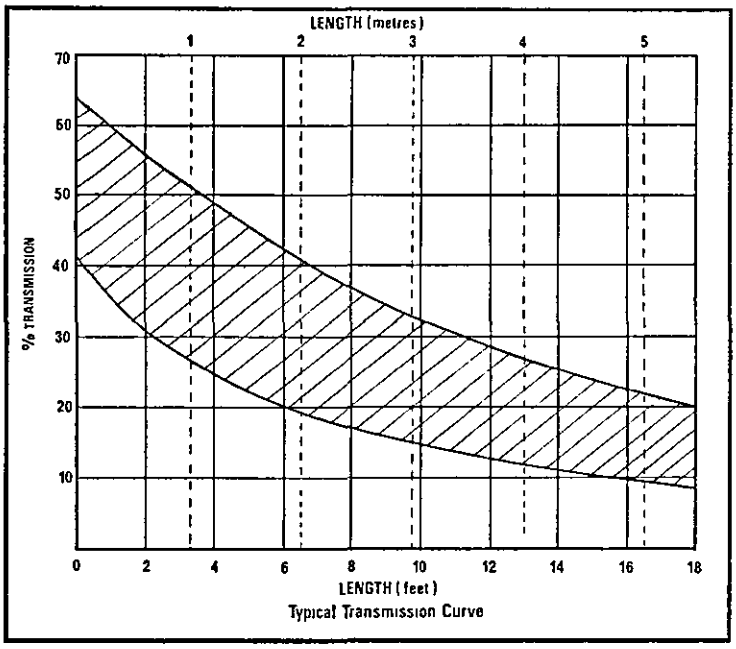
Optical Diameter	Bend Radius
1½ mm	19 mm
3 mm	32 mm
4½ mm	64 mm
6 mm	64 mm
9 mm	89 mm

3.9 Spectral Transmission Typical spectral transmission is shown on Graph 3

** Fibre Optic components with numerical apertures between 24 and 77 can generally be made specially if required Also, Fibre Optic components can generally be made specially to withstand temperatures between −200°C and +250°C, if required*

RAPH 1
RANSMISSION CURVE

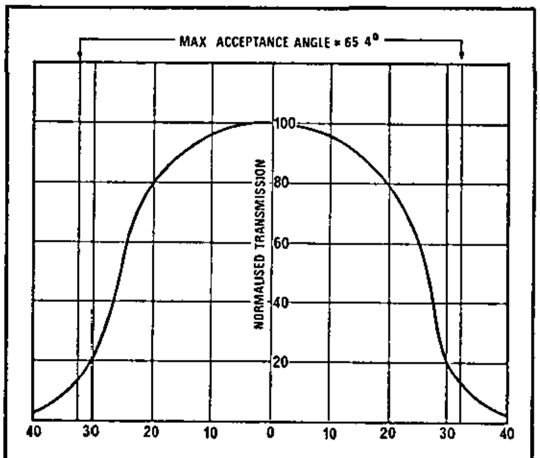
ıs graph shows typical percentages of ht transmission through various lengths fibre optics The light transmission ll be in the shaded area on the graph



Typical Transmission Curve

RAPH 2
OLAR DIAGRAM

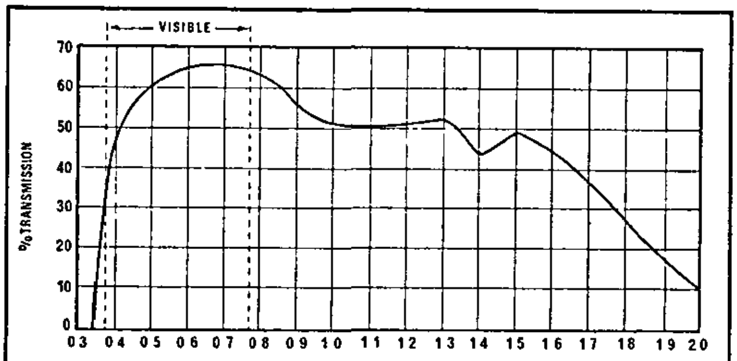
ıs curve shows a typical distribution of ht output from a fibre optic according to angle to optical axis N A 0 54



Typical Normalised Polar diagram for a 0 043 ᵐ/m dia fibre subjected to Lambertian Input

RAPH 3
PECTRAL TRANSMISSION

s curve shows a typical spectral trans-ssion curve for a standard fibre optic It ws that little transmission occurs in the a-violet region



Typical Spectral Transmission Curve (For 45 cm Length of Fibre)

APPENDIX C

Pulse Extractor and Line Markers Circuit

The circuit diagram is shown in Figure C1. R1, C1 and D1 form the input matching circuit. Comparator 1 detects all negative pulses and the line pulses are extracted at the output of this comparator. The line pulses are filtered from the output of comparator 1 by the low pass filter R2 and C2. Comparator 2 detects the remaining frame pulses and they are reshaped by the three 'and' gates 7400 and the R3, C3 integrator network. The line selected by the delay circuit of the tektronix oscilloscope provides a trigger signal on the 'A Gate In' line which together with the relevant line trigger pulse, which indicates the start of the line, trigger the monostable 74121 whose output is a variable length pulse. The pulse length is adjustable by VR4 and the amplitude of the pulse applied to the video signal is determined by VR3. The selected line is 'brightened' by this pulse, and the length of the pulse determines the distance along the line of the 'bright up'.

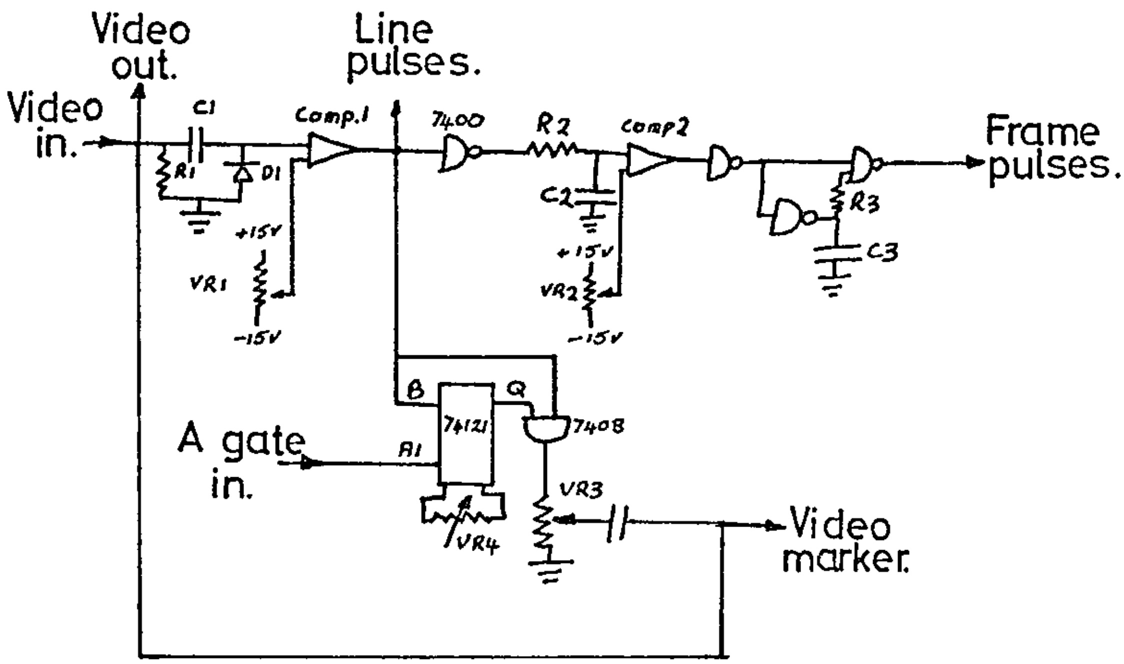


FIGURE C1.

APPENDIX D

Derivation of the possible orientations of the triangular closed figures within an ellipse.

Consider three points (X_1Y_1, X_2Y_2, X_3Y_3) on the periphery of an ellipse. The normal to the ellipse at each point is defined by its tangent T_n. In order for the three points to form the apexes of an inscribed triangle within the ellipse the normals to the ellipse at the three points must intersect at a single point which is then the centre of the inscribed circle of the triangle (since the normals bisect the internal angles of the triangle).

From the theory of caustic formation (Chapter 5.4) two straight lines intersect at a point $\bar{x}$, $\bar{y}$ where

$$\bar{x} = \frac{(K_2-K_1)}{P} \qquad \text{(Equation 164a)} \qquad \text{D1}$$

$$\bar{y} = \frac{T_1K_2-T_2K_1}{P} \qquad \text{(Equation 164b)} \qquad \text{D2}$$

where K_n and P are given by equations 163a-b.

The X co-ordinate of a point on the ellipse may be expressed in terms of its Y co-ordinate and the constants of the ellipse, where from equation 240a

$$X_n = \frac{a}{b} \quad b^2 - Y_n^2 \qquad \text{D3}$$

and from equation 246

$$T_n = \frac{a}{b} \quad \frac{Y_n^2}{b^2 - Y_n^2} \qquad \text{D4}$$

So K_n given by Equation 163 as

$$K_n = Y_n - X_nT_n \qquad \text{D5}$$

may be expressed purely in terms of the Y co-ordinate of the point and the ellipse constants, by substituting for X_n and T_n from equations D3 and D4 to give

$$K_n = Y_n \left(1 - \frac{a^2}{b^2}\right) \qquad \text{D6}$$

Taking the three normals in pairs, the co-ordinates $\bar{x}$, $\bar{y}$ of a point of intersection of the three normals will be given by substituting for K_n from equation D6 into D1 and D2 to give

$$\frac{\bar{x}}{\left(1 - \frac{a^2}{b^2}\right)} = \frac{(Y_2 - Y_1)}{(T_1 - T_2)} = \frac{(Y_3 - Y_2)}{(T_2 - T_3)} = \frac{(Y_1 - Y_3)}{(T_3 - T_1)} \qquad \text{D7}$$

$$\frac{\bar{y}}{\left(1 - \frac{a^2}{b^2}\right)} = \frac{(T_1Y_2 - T_2Y_1)}{(T_1-T_2)} = \frac{(T_2Y_3 - T_3Y_2)}{(T_2 - T_3)} = \frac{(T_3Y_1 - T_1Y_3)}{(T_3-T_1)} \qquad \text{D8}$$

Rearranging the three right hand terms of equation D7 gives

$$\left.\begin{aligned} T_2(Y_2 - Y_1) + Y_2(T_1 - T_2) &= \qquad \text{D9a} \\ T_1(Y_2 - Y_1) &= \qquad \text{D9b} \\ T_1(Y_2 - Y_1) + Y_1(T_1 - T_2) &= \qquad \text{D9c} \end{aligned}\right\} T_3(Y_2 - Y_1)+Y_3(T_1-T_2)$$

Subtracting equation D9b from D9a gives

$$Y_1(T_1 - T_2) = 0 \qquad \text{D10}$$

Let both Y_1 and Y_2 lie in the first quadrant of the ellipse so that Y_1, Y_2, T_1, T_2 are all positive.

From equation D10 either $Y_1 = 0$ or $(T_1 - T_2) = 0$. If $Y_1 = 0$ then $T_1 = 0$ and substituting these values into equation D9c gives

$$T_3Y_2 - Y_3T_2 = 0$$

which on substitution of equation D4 gives

$$Y_2 = Y_3$$

If $Y_2 = Y_3$ the two point are coincident and cannot form the apexes of a triangle.

Equally if $T_1 - T_2 = 0$ then $T_1 = T_2$ and $Y_1 = Y_2$ and these points are also coincident.

Let Y_1 and Y_2 be in adjacent quadrants of the ellipse i.e. Y_2 and T_2 are negative.

Substituting $Y_2 = -Y_2$ and $T_2 = -T_2$ into equation D10 gives

$$Y_1 = 0 \quad \text{or} \quad T_1 = -T_2$$

If $T_1 = -T_2$ then $Y_1 = -Y_2$ and from the symmetry of the ellipse the two normals passing through points X_1Y_1, X_2Y_2 where $Y_1 = Y_2$ must intersect along the X axis of the ellipse, between the two quadrants containing the points. Since the third point must lie somewhere in the opposite two quadrants, in order for its normal to intersect also along the X axis, its normal and the X axis must coincide to give $Y_3 = 0$, these three point forming the triangle shown in Figure 105a.

A similar process, starting with the assumption that Y_1 and Y_2 are in adjacent quadrants but now with Y_2 positive and T_2 negative, gives the result

$$Y_1 = Y_2 \quad \text{and} \quad T_1 = -T_2$$

The normals passing through these two point now intersect along the Y axis, again between the two quadrants containing the points, and result in the orientation of the triangle shown in Figure 105b.

APPENDIX E.

The Computer Simulation Program

The program written to simulate the propagation of rays in an elliptical cross section reflecting system is not described in detail because of the variations in display hardware available in different institutions. The following is an outline of the major steps in the program.

1. The starting position of the first ray is defined together with the desired direction of propagation expressed in cartesian co-ordinates.

2. The increments in the starting positions of successive rays is defined.

3. The starting position of the final ray of the simulation is defined.

4. The number of reflections to be displayed and the number of reflections that must occur before the displayed reflections is defined.

5. Each ray is then incremented in the desired direction, testing at each increment to determine if the reflecting boundary has been reached. When the boundary is reached, the normal to the reflecting surface is calculated, and the direction of the incident ray is altered in accordance with the law of reflection. The reflection number counter is incremented and the ray incremented in its new direction until the boundary is reached again.

6. If the reflection number lies within the range of those to be displayed, then each increment of the ray is displayed.

7. Where a large number of reflections occur before the displayed reflection, where the program may run for a long time (½ hour for the m = 40 caustic shown in Figure 59) a paper tape is produced of the co-ordinates of the required reflection and is used for display at a later time.

8. The simulation in circular reflectors is achieved by setting the constants of the ellipse equal.

9. The program was written in FORTRAN 2 and the simulation run on the Department's Modular One Computer and displayed on a Tektronix Storage System 611.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to the following members of the Department for their various contributions, either to the research work or in the production of this thesis.

To Dr. C. Wilson for his constant guidance and encouragement throughout the period of this research.

To Professor J.W.R. Griffiths for permitting the research to be carried out in the Department and for providing the equipment described in the thesis.

To the members of the Mechanical Workshop for the construction of the majority of the mechanical devices described in the thesis.

To Mr. P. Atkinson for his assistance in the production of the photographs in the thesis.

To Mrs. S. Peach for typing the thesis and labelling the diagrams and photographs.

The author also wishes to thank the Science Research Council for the award of a Studentship for the duration of the research project.

Finally the author wishes to record his gratitude to members of the various industrial research departments who have provided both material and ideas. In particular he wishes to thank Mr. M. Ramsay of S.T.L. (Harlow) and his staff, and Mr. W. Stewart of Plessey (Towcester).

REFERENCES

1. KAPANY, N.S. & BURKE, J.J. 'Optical Waveguides', Academic Press, 1972.
2. MARCUSE, D. 'Theory of Dielectric Optical Waveguides', Academic Press, 1974.
3. KAPANY, N.S. 'Fibre Optics, Principles and Applications', Academic Press, 1967.
4. ALLAN, W.B. 'Fibre Optics, Theory and Practice', Plenum Press, 1973.
5. TIEDEKEN, R. 'Fibre Optics and its Applications', Focal Press 1972.
6. RAMSAY, M.M. 'Fibre Optical Communications within the United Kingdom', Opto-electronics 5 (1973) p 261-274.
7. SUNAK, H.R.D. 'Advances in Fibre-Optical Communication Systems', Electronics & Power, Vol. 19, 29th November, 1973, p 528-532.
8. KAO, K.C. & HOCKHAM, G.A. 'Dielectric-Fibre Surface Waveguides for Optical Frequencies, Proc. IEE, Vol. 113, No. 7, July 1966,p 1151-1158.
9. HOPKINS, H.H. & KAPANY, N.S. ' Transparent Fibres for the Transmission of Optical Images', Opt. Acta., Vol. 1, No. 4, Feb. 1955.
10. UCHIDA, T. et al, 'A Light Focussing Fibre Guide', IEEE Conference on Laser Engineering and Applications, Washington D.C. 1969.
11. KAWAKAMI, S. & NISHIZAWA, T. 'An Optical Waveguide with the Optimum Distribution of the Refractive Index with reference to Waveform Distortion', IEEE Trans. MTT-16, 1969, p 814-818.
12. GLOGE, D. & CHINNOCK, E.L.'Study of Pulse Distortion in Selfoc Fibres', Electronics Letters, Vol.8, No. 21, 1972, p 526-527.
13. DAKIN, J.P. et al 'Theory of Dispersion in Lossless Multimode Optical Fibres', Optics Communications, Vol.17, No. 1, 1973, p 1-5.
14. MAURER, S.J. & FELSEN, L.B. 'Ray Optical Techniques for Guided Waves', Proc. IEEE, Vol.55, No. 10, 1967, p 1718-1729.

15. GAMBLING et al 'Propagation Model for Multimode Optical Fibre Waveguide', Electronics Letters, Vol. 8, No. 10, 1972, p 260-263.

16. BOUILLIE, R. & ANDREWS, J.R. 'Measurement of Broadening of Pulses in Glass Fibres', Electronics Letters, Vol. 8, No.12, 1972, p 309-310.

17. SODHA, M.S., CHAKRAVARTI, A.K. & GAUTAMA, G.D. 'Propagation of Optical Pulses through Cladded Fibres : Modified Theory', Applied Optics, Vol. 12, No. 10, 1973, p 2482-2485.

18. POTTER, R.J. 'Transmission Properties of Optical Fibres', J. Opt. Soc. Am., Vol.51, No.10, 1961, p 1079-1089.

19. BROWN, R.G. & DERICK, B.N. 'Plastic Fibre Optics. /3 Off Axis Transmission Characteristics', Applied Optics, Vol.8, No. 2, 1969, p 437-442.

20. STEWART, W.J. 'Asymmetries in Optical Fibres', Optics Communications, Vol. 9, No.1, 1973, p 98-102.

21. WILSON, C. & STEVENS, P.J.'Polishing Technique for Optical Waveguide Terminations', Journal of Physics E., Vol.7, 1974, p 614-615.

22. FRANKE, G 'The Production of Optical Parts', Chapter 8 of Advanced Optical Techniques, edited by A.C.S. Van Heel, North Holland Publishing Co., 1967.

23. GLOGE, D. et al 'Optical Fibre end Preparation for Low Loss Splices', B.S.T.J., Vol.52, No. 9, 1973, p 1579-1588.

24. DAKAS, M.L. & BRIDGER, A. 'Plug-in Fibre to Fibre Coupler', Electronics Letters, Vol.10, No.14, 1974, p 280-281.

25. DALGLEISH, J.F. & LUKAS, H.H. 'Optical-Fibre Connector', Electronics Letters, Vol.11, No. 1, 1975, p 24-26.

26. SOMEDA, C.G. 'Simple Low-loss Joints between Single Mode Optical Fibres', B.S.T.J., Vol. 52, No. 4, 1973, p 583-596.

27. BISBEE, D.L. 'Optical Fibre Joining Technique', B.S.T.J., Vol.50, No. 10, 1971, p 3153-3158.

28. DYOTT, R.B., STERN, J.R. & STEWART, J.H. 'Fusion junctions for glass fibre waveguides', Electronics Letters, Vol. 8, No.11, 1972, p 290-292.

29. BURRUS, C.A. & STANDLEY, R.D. 'Viewing Refractive - Index Profiles and Small Scale Inhomogeneities in Glass Optical Fibres: Some Techniques', Applied Optics, Vol.13, No.10, 1974, p 2365-2369.

30. SNYDER, A.W., MITCHELL, D.J. & PASK, C. 'Failure of Geometric Optics for the Analysis of Circular Optical Fibres', J.Opt.Soc. Am., Vol.64, No.5, 1974, p 608-614.

31. BORN, M & WOLF, E. 'Principles of Optics', Pergamon Press 1964.

32. MARCUSE, D. 'Light Transmission Optics', Van Nostrand Reinhold Co. 1972.

33. STRONG, J. 'Concepts of Classical Optics', W.H. Freeman & Co., 1958.

34. BROWN, E.B. 'Modern Optics', Reinhold Publishing Corporation 1965.

35. 'The Services Textbook of Radio'. Volume 5 'Transmission and Propagation' E.V.D. GLAZIER & H.R.L. LAMONT, H.M. Stationery Office, 1958.

36. BADEN FULLER, A.J. 'Microwaves', Pergamon Press 1965.

37. NUSSBAUM, A. 'Electromagnetic Theory for Engineers and Scientists', Prentice Hall Inc. 1965.

38. BLEANEY, B.I. & BLEANEY, B. 'Electricity and Magnetism', Clarendon Press, Oxford 1965.

39. ALLEN, L., GATEHOUSE, S. & JONES, D.G.C. 'Enhancement of Spatial Coherence during Light Propagation in Bounded Media', Optics Communications, Vol. 4, No.2, 1971, p 168-171.

40. WOLF, E. 'Spatial Coherence of Resonant Modes in a Maser Interferometer' Physics Letters, Vol. 3, No. 4, 1963, p 166-168.

41. STREIFER, W. 'Spatial Coherence in Periodic Systems', J.Opt.Soc. Am., Vol.56, No. 11, 1966, p 1481-1489.

42. SNYDER, A.W. 'Surface Waveguide Modes along a Semi-infinite Dielectric Fibre Excited by a Plane Wave', J. Opt. Soc. Am., Vol. 56, No. 5, 1966, p 601-606.

43. SNYDER, A.W., MITCHELL, D.J. 'Leaky Rays on Circular Optical Fibres', J. Opt. Soc. Am., Vol. 64, No. 5, 1974, p 599-607.

44. LONGHURST, R.S. 'Geometrical and Physical Optics', Longmans, Green and Co.Ltd., 2nd Edition 1967.

45. LANDAU, L.D. & LIFSHITZ, E.M. 'The Classical Theory of Fields', Pergamon Press, Third English Edition 1971.

46. ABRAMOWITZ, M. & STEGUN, I.A. 'Handbook of Mathematical Functions' Dover Publications Inc., 1965.

47. LONEY, S.L. 'The Elements of Co-ordinate Geometry', Macmillan & Co. Ltd., London 1947 Edition.

48. EDWARDS, J. 'Differential Calculus for Beginners', Macmillan & Co.Ltd., London 1945.

49. GIECK, K. 'A Collection of Technical Formulae', English Translation by F. Catty, Heilbronn 1967. (West Germany).

50. KOGELNIK, H. & LI, T. 'Laser Beams and Resonators', Applied Optics, Vol. 5, No. 10, 1966, p 1550-1566.

51. MARCUSE, D. 'Coupled Mode Theory of Round Optical Fibres', B.S.T.J., Vol. 52, No. 6, 1976, p 817.

52. WOOLRIDGE, A.B. 'The Reduction of Microbending Losses by Coating Optical Fibres', Paper presented at the Colloquium on Optical Fibre Waveguides at Queen Mary College, London, 8th January, 1976.

53. GAMBLING, W.A. et al 'Gigahertz Bandwidths in Multimode Liquid Core Optical Fibre Waveguides', Optic Comm., Vol. 6, No. 4, 1972.

54. SNYDER, A.W. 'Asymptotic Expressions for Eigenfunctions and Eigenvalues of a Dielectric or Optical Waveguide', IEEE Trans. MTT-17, p 1130-1138, December 1969.

55. GLOGE, D. 'Weakly Guiding Fibres', App. Optics, Vol. 10, 1971, p 2252-2258.

56. YARIV, A. 'Quantum Electronics', John Wiley & Sons, Inc. 1967.

57. MAITLAND, A. 'Laser Physics', North Holland Publishing Co. 1969.

58. LEENDERTZ, J.A. 'Interferometric Displacement Measurement on Scattering Surfaces using Speckle Effect', Jnl. Physics E., Vol. 3, p 214-218, 1970.

59. ASAKURA, T. et al 'Measurement of Spatial Coherence using Speckle Patterns', Optica Acta, 1972, Vol 19, No. 4, p 273-290.

60. DAKIN, J.P. 'Glass Fibre Lenses for Increased Launching Efficiency into Fibre Optic Waveguides', Paper presented at the Colloquium at Queen Mary College in January, 1976.

61. ALLEN, W.B. Reference 4, p 51.

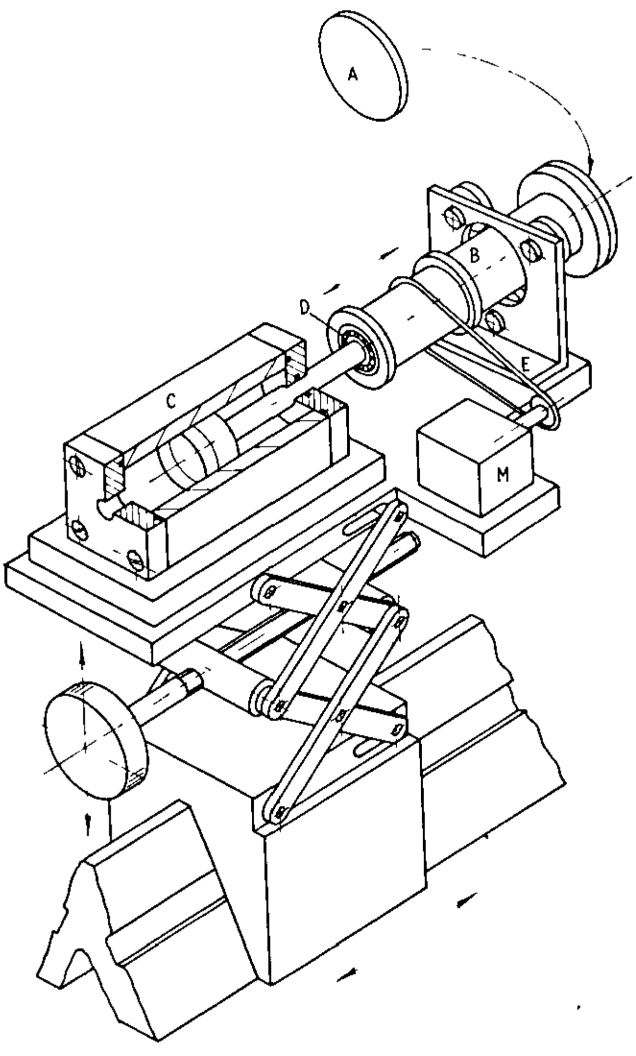


Figure 2 The grinding and polishing machine. A, aluminium mounting plate, B, chuck; C, pneumatic piston, D, ball race, E, belt drive; M, motor

variable between 0 and 3000 rpm The air supply is obtained from a large manually actuated piston which forms a closed system with cylinder (C)

4 Procedure

The fibre is clamped in the vice jaws with the fibre end approximately level with the jaw face Normal grinding and polishing procedures are used to obtain the final finish The curvature of the finished surface is demonstrated in figure 3 The interference fringes were obtained using a standard interference microscope objective and a sodium light source The fibre was adjusted to give the minimum number of fringes across the end face For many experiments a high optical finish is not required and a flatter surface is obtained by the grinding process alone Since the fibre end is proud of the jaw surface, the polishing compound tends to round off the edges of the fibre

5 Conclusion

A technique is described which terminates glass optical waveguides without the necessity of permanent embedding in a holding material For many experiments a raw end may be prepared in 5 min, thus allowing rapid inspection of different lengths of the same waveguide

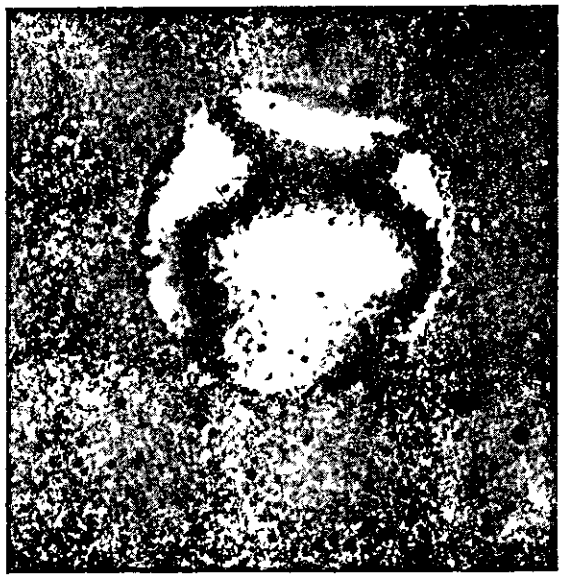


Figure 3 Photomicrograph of a 50 μ diameter cladded construction glass fibre showing interference fringes across the end of the fibre to demonstrate the slope of the end face

The termination is nominally flat and may be at any desired angle to the fibre length Methods of alignment of the fibre to give a specific slope have not been discussed and will be published later.

Journal of Physics E: Scientific Instruments 1974 Volume 7
Printed in Great Britain © 1974

Polishing technique for optical waveguide terminations

C Wilson and P J Stevens
Department of Electronic and Electrical Engineering, Loughborough University of Technology, LE11 3TU

Received 31 January 1974, in final form 15 May 1974

Abstract A technique is described which produces a good optical finish on the end of an optical waveguide at any required slope, without permanently embedding the waveguide in a holding medium

1 Introduction

To investigate the propagation of energy in optical waveguides, a method of terminating these devices is required, which has a precision compatible with the parameters to be investigated A typical glass waveguide of cladded construction has a core of 40 μm diameter and a cladding of 5 μm thickness

The usual method of producing high quality terminations is to embed the end of the glass fibre (waveguide) in a resin compound, which then allows conventional grinding and polishing machines to be used

The major drawback of this method is that the length of the fibre is fixed unless either the fibre is embedded for the whole of its length, or the end is re-embedded for each change of length.

The ideal technique would be to hold the fibre in such a manner that it may be ground and polished, and then released This paper describes such a technique. The equipment consists of a vice and an all-purpose grinding and polishing machine

2 The vice and vice jaws

The vice jaws must be of a material which grips the 50 μm diameter fibre firmly, but whose hardness is less than the glass of the fibre, so that during grinding the jaws will also be ground, but at a greater rate than the fibre end. The jaws are manufactured from an embedding resin casting, which is cut as shown in figure 1(a) The mould is a 7 mm diameter gelatine

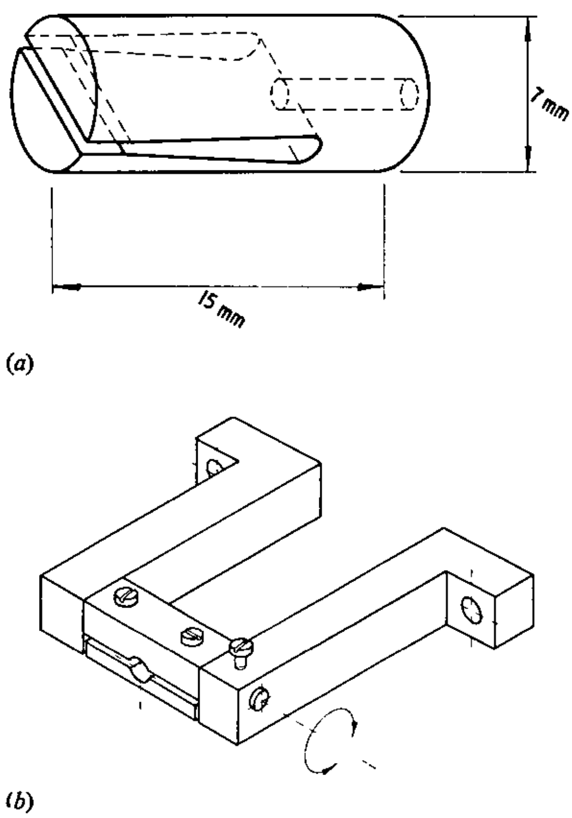


Figure 1 (a) The vice jaws, (b) the mounting vice

capsule The fibre to be polished is passed through the hole in the base of the jaws and then clamped between the jaws by the mounting vice shown in figure 1(b) The clamping pressure must be sufficient to hold the fibre, but not great enough to encourage chipping of the leading edge of the fibre by the coarser grades of grinding The mounting vice is secured to an adjustable table which is mounted on an optical bench Various adjustments allow the fibre to be positioned at any angle to the grinding plane

3 Grinding and polishing machine

The optical bench mounted machine is shown in figure 2

The various grinding and polishing compounds are mounted on aluminium plates (A) These plates are fixed in turn to the rotating chuck (B) The chuck is mounted on the piston rod of the pneumatic cylinder (C) via two ball races (D) The motor (M) rotates the chuck via the belt drive (E) The plates (A) are thus rotated by the motor (M) and moved along the Z axis by the pneumatic piston action When the plates are in contact with the workpiece, pressure between the two may be adjusted by varying the air pressure in the pneumatic cylinder (C). The motor (M) is a single pole induction motor and is driven by a variable frequency supply Its speed is