USE OF THESES

This copy is supplied for purposes of private study and research only. Passages from the thesis may not be copied or closely paraphrased without the written consent of the author.
THE THEORY OF UNBOUND MODES ON CIRCULAR DIELECTRIC WAVEGUIDES

ROWLAND ALEXANDER SAMMUT

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

Canberra,
December, 1975
PREFACE

This dissertation is an account of work carried out in the Department of Applied Mathematics of the Institute of Advanced Studies in the Australian National University between February, 1973 and November, 1975 under the supervision of Drs. Allan W. Snyder and Colin Pask.

While I have benefited immeasurably throughout this time from discussions with members of this department and other workers in the field, the work described in the following pages is, unless explicitly disowned in the text, my own.

None of this work has been submitted to any other institution of learning for any degree.

Rowland A. Sammut

Rowland A. Sammut
PUBLICATIONS


ACKNOWLEDGEMENTS

It is undoubtedly bad form but I must begin this thesis by making a complaint against the members of the department in which it was written and particularly against its head, Barry Ninham. For while their constant friendship, help and encouragement have made the writing of most of the thesis more enjoyable and rewarding than it would otherwise have been, they have made the writing of these acknowledgements an impossibly difficult task to perform adequately.

From the time of my first brief stay in his department as a Summer-vacation student, Barry Ninham has taken an interest in my welfare, both personal and academic, which extended far beyond the call of duty and, at times, certainly beyond what I have deserved. Through his influence, the department has developed an atmosphere of friendship and enthusiasm whose equal is difficult to imagine.

I am no less indebted to my supervisors, Allan Snyder and Colin Pask, who have generously provided, as the occasion required, advice, encouragement, enthusiasm, experience, long (but enlightening) arguments, metaphorical shoulders to cry on and a great deal of their time. My completion of this work without their help is also difficult to imagine.

John Love and John Mitchell helped greatly by allowing me to pick their mathematical brains at various times in the course of the work and Dieter Marcuse by cheerfully allowing me to distract him from more important things to search for the ubiquitous factor-of-two mistake.

I am sincerely grateful to my fellow students Peter McIntyre and Dave Carpenter for their personal and scientific companionship over the past three years and particularly to Ian White for most of his suggestions on the original draft of the thesis. The fact that the thesis has appeared in public indicates, however, that not all of these suggestions were treated equally seriously.

Finally, I must express my gratitude to Norma Chin for the admirable skill and good humour with which she converted a very difficult manuscript into a readable form, and the Australian Government and the Australian National University for their financial support.
ABSTRACT

The electromagnetic field propagating along an ideal dielectric waveguide may contain both bound and unbound (radiation) components. In a transverse spectral representation, the bound field is represented by a finite sum of discrete modes but the unbound field requires integration over a continuum of modes. This integration, while it provides an exact description of the radiation field, can be very slowly convergent. An asymptotic representation using improper (leaky) modes is more rapidly convergent but is only valid in a restricted region of space.

The purpose of this thesis is to study both the spectral and improper unbound modes, and the connection between them.

Chapter 1 provides some general background to the modal theory of electromagnetic-field propagation in dielectric waveguides, with particular emphasis on the unbound modes. It defines the physical system to be studied and the basic notation to be used in the remainder of the thesis and, along the way, re-derives some well-known results of electromagnetic theory for later use.

In Chapter 2, the bound modes of a circular dielectric rod are listed, and the spectral radiation modes derived, in a form found most useful for later analysis. Orthogonality and normalization conditions for the radiation modes are derived using a general formalism which yields a relationship between free-space and modal fields valid in a much wider class of systems than those of specific interest here.

These orthonormality conditions are used in Chapter 3 to study the excitation and propagation of the radiation modes within a circular dielectric rod. The first source used is a truncated plane wave. The effects of changing the angle of incidence or the extent of the source field are studied. A close analogy is found between the far-field
radiation pattern of a dielectric waveguide with small dielectric-
constant-difference between core and cladding and the diffraction
pattern of a circular aperture. Similar calculations are then repeated
with a quasi-monochromatic, totally incoherent source.

In Chapter 4, the transformation is made from the transverse
spectral representation of the waveguide field to a longitudinal
representation. A steepest-descent approximation is then made by which
the integral over the continuum of radiation modes is replaced, within a
restricted region, by a sum over discrete, improper, unbound modes —
leaky modes — plus a saddle-point contribution. The leaky modes are
shown to be the analytic continuation of bound modes below cutoff and
the range of significance of the saddle-point term is estimated. Two
methods for determining the orthogonality and excitation of leaky modes
are investigated.

The detailed characteristics of leaky modes are studied, both
numerically and asymptotically in Chapter 5. Using these
characteristics, the excitation and propagation of the radiation field
are studied in terms of leaky modes under the same excitation conditions
as used in Chapter 3. These results are used to estimate the range of
validity of the leaky ray theory. The number of leaky modes with a
given attenuation and the effect on this attenuation of including
material absorption are also discussed.

In Chapter 6, some of the above ideas are applied to the
optical waveguides in visual photoreceptors. The contribution of
unbound modes to light absorption in these structures is investigated
and an ambiguity in the method of determining photoreceptor parameters
by observation of mode cutoffs is pointed out.

The major conclusions of the thesis are summarized in
Chapter 7.
Notes on the text:

(i) References are numbered consecutively and listed at the end of each chapter.

(ii) Equations are numbered consecutively within each section. In the notation \( (\ell.m.n) \), \( \ell \) is the chapter number, \( m \) the section number and \( n \) the equation number. Within a particular chapter, equation numbers are abbreviated to \( (m.n) \).
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>ii</td>
</tr>
<tr>
<td>PUBLICATIONS</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>1. INTRODUCTION AND PRELIMINARY ELECTROMAGNETIC THEORY</td>
<td></td>
</tr>
<tr>
<td>1.1 Modal Expansions in Waveguide Theory</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Some Preliminary Results and Notation</td>
<td>6</td>
</tr>
<tr>
<td>References</td>
<td>14</td>
</tr>
<tr>
<td>2. THE PROPER MODES OF A DIELECTRIC WAVEGUIDE OF CIRCULAR CROSS SECTION</td>
<td></td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>17</td>
</tr>
<tr>
<td>2.2 Bound Modes</td>
<td>18</td>
</tr>
<tr>
<td>2.3 Radiation Modes</td>
<td>24</td>
</tr>
<tr>
<td>2.3.1 Properties and methods of derivation</td>
<td>24</td>
</tr>
<tr>
<td>2.3.2 Radiation modes from scattering</td>
<td>27</td>
</tr>
<tr>
<td>2.4 Orthonormality of the radiation modes</td>
<td>37</td>
</tr>
<tr>
<td>2.5 Résumé</td>
<td>45</td>
</tr>
<tr>
<td>Appendix 2A: Direct Calculation of Normalization and Orthogonality</td>
<td>47</td>
</tr>
<tr>
<td>Appendix 2B: Alternative Form for Transverse Components of Radiation Modes</td>
<td>53</td>
</tr>
<tr>
<td>References</td>
<td>55</td>
</tr>
<tr>
<td>3. EXCITATION AND PROPAGATION OF UNBOUND MODES</td>
<td></td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>57</td>
</tr>
<tr>
<td>3.2 Plane Wave Excitation</td>
<td>61</td>
</tr>
<tr>
<td>3.2.1 Amplitude coefficients</td>
<td>61</td>
</tr>
<tr>
<td>3.2.2 Power propagation</td>
<td>67</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.2.3 Comparison with diffraction at a circular aperture</td>
<td>80</td>
</tr>
<tr>
<td>3.2.4 &quot;Slab&quot; approximation to total unbound-mode power</td>
<td>84</td>
</tr>
<tr>
<td>3.3 Excitation by Quasi-monochromatic Incoherent Source</td>
<td>86</td>
</tr>
<tr>
<td>3.4 Generalization and Conclusions</td>
<td>93</td>
</tr>
<tr>
<td>Appendix 3A: Plane-wave Excitation Coefficients for Bound Modes</td>
<td>94</td>
</tr>
<tr>
<td>Appendix 3B: Cross-integrals for Calculation of Power in Waveguide</td>
<td>96</td>
</tr>
<tr>
<td>Appendix 3C: Coefficients for Incoherent Source</td>
<td>99</td>
</tr>
<tr>
<td>References</td>
<td>103</td>
</tr>
<tr>
<td>4. ASYMPTOTIC EXPANSION OF THE RADIATION FIELD WITHIN A</td>
<td></td>
</tr>
<tr>
<td>CIRCULAR DIELECTRIC WAVEGUIDE - LEAKY MODES</td>
<td></td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>105</td>
</tr>
<tr>
<td>4.2 Alternative Representations of the Waveguide Field -</td>
<td>109</td>
</tr>
<tr>
<td>Steepest-Descent Approximation</td>
<td></td>
</tr>
<tr>
<td>4.3 Pole Contributions</td>
<td>118</td>
</tr>
<tr>
<td>4.4 Saddle-point Contribution</td>
<td>125</td>
</tr>
<tr>
<td>4.5 Orthogonality and Excitation of Leaky Modes</td>
<td>129</td>
</tr>
<tr>
<td>4.5.1 Truncated modes and coupling</td>
<td>130</td>
</tr>
<tr>
<td>4.5.2 An exact orthogonality condition</td>
<td>135</td>
</tr>
<tr>
<td>4.6 Résumé</td>
<td>139</td>
</tr>
<tr>
<td>References</td>
<td>140</td>
</tr>
<tr>
<td>5. CHARACTERISTICS AND APPLICATIONS OF LEAKY MODES</td>
<td></td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>143</td>
</tr>
<tr>
<td>5.2 Approximate Eigenvalue Equation</td>
<td>148</td>
</tr>
<tr>
<td>5.2.1 Numerical solution</td>
<td>148</td>
</tr>
<tr>
<td>5.2.2 Asymptotic results</td>
<td>157</td>
</tr>
<tr>
<td>5.3 Exact Eigenvalue Equation - Detailed Mode Characteristics</td>
<td>164</td>
</tr>
<tr>
<td>5.4 Excitation and Propagation of the Unbound Field -</td>
<td>183</td>
</tr>
<tr>
<td>Leaky Mode and Leaky Ray Analyses</td>
<td></td>
</tr>
<tr>
<td>5.4.1 Leaky mode analysis</td>
<td>183</td>
</tr>
<tr>
<td>5.4.2 Leaky ray analysis</td>
<td>186</td>
</tr>
<tr>
<td>5.4.3 Numerical results</td>
<td>188</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION AND PRELIMINARY

ELECTROMAGNETIC THEORY

Only a portion of the energy launched into the core of a dielectric waveguide remains within the core indefinitely, even in an ideal, dissipationless system. The remainder is gradually radiated by one of two mechanisms which can best be described in ray-optical terms.

Rays which strike the inner surface of a waveguide at an angle to the normal smaller than the critical angle given by Snell's law lose energy to the cladding (the outer medium) due to refraction. On the other hand rays which would be totally reflected if the surface were planar, but which have an angle to the waveguide axis greater than the complement of this critical angle, lose energy due to curvature effects.

An electromagnetic analysis of waveguide propagation is often most conveniently carried out by studying the normal modes of the structure and then superposing these modes to form any required electromagnetic field. Rays which are totally internally reflected then correspond to bound modes and the other two classes of rays to unbound modes.

It is the purpose of this thesis to study the unbound modes of an ideal, circular dielectric rod (with application to optical fibres and visual photoreceptors) and to determine, amongst other things, the
circumstances under which the ray-optical description is not only qualitatively helpful but also quantitatively accurate.

The unbound modes are studied firstly in an exact form valid everywhere (proper modes) and then in an approximation which only has physical significance within a restricted region (leaky modes).

1.1 MODAL EXPANSIONS IN WAVEGUIDE THEORY

In conventional parlance, a "proper" mode of a waveguide is a characteristic field of the structure which satisfies

1. the source-free Maxwell equations,
2. boundary conditions at the waveguide surface, and
3. the Sommerfeld radiation condition at infinity.

In a closed waveguide, an arbitrary electromagnetic field can be expanded in a complete spectrum of proper, discrete modes. However, in a dielectric waveguide (indeed in any open, guiding structure), only a finite number of such discrete modes exist at a given frequency so that these cannot form a complete set.

This difference arises from the fact that a dissipationless, open waveguide cannot support proper, evanescent modes while the closed structure supports both propagating and evanescent modes. Because evanescent modes correspond to field variations of spatial period smaller than the wavelength of light (page 9 of Ref. 2), the absence of these modes on open waveguides is physically equivalent to the statement that a proper, discrete mode on such a waveguide must have a phase velocity smaller than that of a plane wave in the external medium. 1,3
1.1

Fields with phase velocities greater than that of a plane wave in the external medium represent radiation.

Such fields can be found which satisfy conditions (1) and (2) above but, while they remain everywhere finite, these fields do not individually satisfy condition (3) because they are formed by a superposition of both incoming and outgoing waves. However, since these "radiation modes" are not discrete but form a continuum, any finite source will always excite a number of these modes and their superposition will satisfy condition (3). The term "proper mode" will, therefore, also be applied to the modes of the continuum.

The combination of the mutually orthogonal discrete (bound) and continuous (radiation) mode sets forms the complete mode spectrum for the open waveguide. In Chapter 2, the bound modes for a uniform, circular, dielectric cylinder embedded in an infinite, uniform medium of lower refractive index are listed. The radiation modes for this structure are then derived and their orthogonality and normalization is investigated. These radiation modes are used to investigate the excitation and propagation of the unbound field on such a waveguide in Chapter 3, emphasis being placed on that portion of the field which remains within the waveguide after a finite length. The sources used are a truncated plane wave (i.e. a coherent, or collimated, source) and a quasi-monochromatic, totally (spatially) incoherent source.

The spectral representation described so far, while it is mathematically complete, is not always particularly rapidly convergent. For this reason, workers in the field of antenna theory sought approximations to the radiation field by no longer requiring that a modal field necessarily remain finite everywhere."
restriction, one finds that open waveguides, in fact, have discrete modes which decay in the axial direction but that these are "improper" in the sense that while they again obey conditions (1) and (2) above, they grow exponentially beyond a certain radial distance from the waveguide.

The improper modes, labelled "leaky modes" by Marcuvitz can therefore only be used within restricted regions of space. Within their range of validity, however, they can be a very useful approximation to the radiation field. Moreover, while the leaky mode approximation was derived for, and has generally been used in, antenna problems, it is no less valid or useful in the study of the radiation field within a dielectric waveguide.

Detailed investigations concerned with leaky modes have previously been confined to planar interfaces or slab configurations. In Chapter 4, the mathematical details involved in the transformation from a spectral representation to a non-spectral one in terms of leaky modes will be presented for the circular dielectric waveguide described above. Methods of obtaining orthogonality and normalization conditions for these non-square-integrable modes are also discussed. Chapter 5 then describes the physical characteristics of the leaky modes and investigates their behaviour, numerically for low order modes and asymptotically for the higher-order modes.

The reason for this interest in the unbound modes (proper or improper) within a dielectric waveguide is their possible importance in the two fields of optical fibre communications and photoreceptor optics.

It seems inappropriate to attempt any general review of the literature connected with optical communication fibres, not only because
such a review might possibly have to begin with Newton's speculation on what is now known as the "Goos Hänchen shift" but also because several comprehensive reviews have recently appeared and others are in preparation. Moreover, as the authors of these reviews have pointed out, the field is a very rapidly changing one.

As a striking example of the rate of technological change, when Snitzer was making his mode observations in 1961, the losses of the fibres involved were approximately 1500 dB/km. At that time, he was "optimistic" that this figure could be improved by a factor of 10. By 1970, it had in fact been reduced by a factor of 75 and since then by another factor of 10 so that, at present, losses as low as the theoretical limit in bulk silica can occasionally be achieved.

As a result of these technological advances, experiments have been able to be done which suggest that unbound modes may still constitute a significant portion of the field within a waveguide even after kilometre lengths.

At the same time, there has been an accompanying proliferation of theoretical studies. One of the results of this has been a new look at leaky modes and the recognition that slowly radiating leaky modes should, indeed, persist over long distances in a low-loss fibre. Another result has been the development of modifications of classical ray-tracing techniques to take account of radiation from the curved surface of an optical fibre. In a central section of Chapter 5, we attempt to evaluate this ray-optical approximation by solving the excitation problem for the leaky modes using the two sources mentioned above and comparing the results with those obtained by ray-tracing.
The application of waveguide theory to photoreceptor optics, while it is by no means new, has also experienced an upsurge in interest in recent times. Developments in this field have lately been reviewed at a conference of its leading investigators. But the main property of visual photoreceptors of interest here is the fact that they are sufficiently short so that, even though the radiation field leaves them very quickly, this field may still contribute significantly to light absorption under certain circumstances. This contribution of unbound modes to absorption in photoreceptors is investigated in Chapter 6. In that chapter, the method of determining photoreceptor parameters by observation of mode cutoffs is also evaluated in the light of the presence of leaky modes.

The major conclusions of the thesis are summarized in Chapter 7.

But we begin by listing some well-known, preliminary results required in later chapters.

1.2 SOME PRELIMINARY RESULTS AND NOTATION

The fundamental quantities involved in electromagnetic phenomena are the electric field intensity \( \mathbf{E} \), the electric displacement \( \mathbf{D} \), the magnetic field intensity \( \mathbf{H} \) and the magnetic flux density \( \mathbf{B} \). The equations governing their behaviour in a linear, source-free medium of dielectric permittivity \( \varepsilon \) and magnetic permeability \( \mu \) are Maxwell's equations

\[
\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t},
\]

\[
\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t},
\]

(2.1)

(2.2)
\[ \nabla \cdot \mathbf{D} = 0, \quad (2.3) \]
\[ \nabla \cdot \mathbf{B} = 0, \quad (2.4) \]

and the constitutive relations
\[ \mathbf{D} = \varepsilon \mathbf{E}, \quad (2.5) \]
\[ \mathbf{B} = \mu \mathbf{H}, \quad (2.6) \]

where \( \varepsilon \) and \( \mu \) are tensors in an anisotropic medium.

At a dielectric interface (a discontinuity in the dielectric permittivity) the electromagnetic fields must also satisfy the boundary conditions

(i) components of \( \mathbf{E} \) and \( \mathbf{H} \) tangential to the interface are continuous across the interface, and

(ii) components of \( \mathbf{D} \) and \( \mathbf{B} \) normal to the interface are continuous.

In the absence of static electric charges or magnetic fields, Eqs. (2.3-4) and boundary conditions (ii) are a direct consequence of Eqs. (2.1-2) so that, in the time-varying situations with which we are concerned, only the latter equations and boundary conditions (i) need be considered.\(^2\)

Any well-behaved complex field \( \mathbf{F}(\mathbf{r}, t) \) (electric or magnetic) which satisfies these equations and conditions can be Fourier-transformed in time to give

\[ F(\mathbf{r}, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\mathbf{r}, t) \, e^{-i\omega t} \, dt \quad (2.7) \]
\[ F(\mathbf{r}, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\mathbf{r}, \omega) \, e^{i\omega t} \, d\omega \quad (2.8) \]
so that it is only necessary to consider time-harmonic fields of the form $F(x, \omega) e^{i\omega t}$, where $\omega$ is the angular frequency of the field.
Throughout this thesis, the fields to be studied will be complex, time-harmonic fields, it being understood that the physical field is given by the real part of this quantity.

For these harmonic fields, Eqs. (2.1 - 2) can be rewritten using Eqs. (2.5 - 6) to give

\[ \nabla \times \tilde{E} = -i\omega \mu \tilde{H} \quad (2.9) \]

and

\[ \nabla \times \tilde{H} = i\omega \varepsilon \tilde{E} \quad (2.10) \]

where $\tilde{E}$ and $\tilde{H}$ are the Fourier transforms (in time) of $E$ and $H$.

If $\mu$ is a homogeneous quantity, Eq. (2.9) can again be rewritten in a form more useful for the analysis of field propagation along a waveguide. Taking the curl of Eq. (2.9), substituting Eq. (2.10) and using the identity (in a cartesian coordinate system),

\[ \nabla \times (\nabla \times \tilde{E}) = \nabla (\nabla \cdot \tilde{E}) - \nabla^2 \tilde{E} \quad (2.11) \]

we find

\[ \nabla^2 \tilde{E} + \left( \nabla \left( \frac{\nabla E}{E} \right) \right) + \omega^2 \mu \varepsilon \tilde{E} = 0 \quad (2.12) \]

In graded-index fibres where $\varepsilon$ is not homogeneous, the second term in Eq. (2.12) may be of some significance if $\varepsilon$ varies on a scale comparable with the wavelength of light. In the case where $\varepsilon$ varies more slowly or in the case of step-index fibres, Eq. (2.12) reduces to the Helmholtz equation

\[ [\nabla^2 + k^2] \tilde{E} = 0 \quad (2.13) \]

where

\[ k = \omega (\mu \varepsilon)^{1/2} = \frac{2\pi}{\lambda} \quad (2.14) \]
1.2

is the magnitude of the propagation vector, $\lambda$ being the wavelength of light in the medium in question.

If we choose a coordinate system in which the $z$-axis lies in the direction of the waveguide axis, then provided the waveguide is translationally invariant in the axial direction, the field $\mathbf{E}$ can be written in the form $\mathbf{E}(x,y) e^{i(\omega t - \beta z)}$, where $x,y$ are rectangular coordinates in the transverse plane and $\beta$ is the $z$-component of the wave vector. The minus sign is chosen in the exponent to give a wave travelling in the positive $z$-direction (a forward wave) when the real part of $\beta$ is positive.

Substituting this form for the electric field in Eq. (2.13) we obtain the equation

$$[\nabla^2_t + (k^2 - \beta^2)] \mathbf{E} = 0,$$  \hspace{1cm} (2.15)

where $\nabla^2_t$ is the transverse part of the Laplacian operator. Taking the curl of Eq. (2.10) rather than of Eq. (2.9) and following through the above procedure in the case where $\varepsilon$ is homogeneous shows that the magnetic field vector satisfies an identical equation.

In rectangular coordinates, Eq. (2.15) separates so that each component of $\mathbf{E}$ (or $\mathbf{H}$) satisfies the scalar Helmholtz equation, $\nabla^2_t$ being replaced by the scalar operator

$$\nabla^2_t = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$ \hspace{1cm} (2.16)

In the cylindrical polar coordinate system more appropriate to our problem, the equations for the transverse components remain coupled, however, and it is only the $z$-component which satisfies the scalar Helmholtz equation, with
But the transverse components can be derived directly from the longitudinal (z) components.

To see this we write

$$\nabla^2 = \frac{\partial^2}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2}. \quad (2.17)$$

where \( \hat{z} \) is a unit vector in the z-direction and the subscript t again indicates the transverse part. Then substituting in Eqs. (2.9) and (2.10) and equating transverse components, we find

$$-i\omega \mu \mathbf{H}_t = \nabla \times \mathbf{E}_t + \frac{\partial}{\partial z} \left( \nabla \times \mathbf{E}_t \right) \quad (2.21)$$

$$i\omega \mathbf{E}_t = \nabla \times \mathbf{H}_t + \frac{\partial}{\partial z} \left( \nabla \times \mathbf{H}_t \right). \quad (2.22)$$

Substituting Eq. (2.22) into (2.21),

$$k^2 \mathbf{H}_t = i\omega \left( \nabla \times \mathbf{E}_t \times \hat{z} \right) + \nabla \times \left( \frac{\partial \mathbf{H}_t}{\partial z} \right) - \frac{\partial^2}{\partial z^2} \mathbf{H}_t. \quad (2.23)$$

Finally, using the fact that \( \frac{\partial}{\partial z} = -i\beta \) in Eq. (2.23) we have

$$\mathbf{H}_t = (k^2 - \beta^2)^{-1} \left[ -i\beta \nabla \mathbf{E}_t + i\omega \left( \nabla \times \mathbf{E}_t \right) \right] \quad (2.24)$$

and similarly

$$\mathbf{E}_t = (k^2 - \beta^2)^{-1} \left[ -i\beta \nabla \mathbf{H}_t - i\omega \mu \left( \nabla \times \mathbf{H}_t \right) \right]. \quad (2.25)$$

Equations (2.24) and (2.25) hold for arbitrary graded index profiles provided only that \( \varepsilon \) and \( \mu \) do not vary with z.
So in determining the fields propagating along a waveguide, we need only solve the scalar Helmholtz equation for the z-components of the electric and magnetic fields. The remaining components can be found from Eqs. (2.24 - 25).

The particular system we will be studying in the remainder of this thesis is that shown in Fig. 1.1.

The "core" of the waveguide has radius $\rho$, uniform refractive index $n_1$ (dielectric permittivity $\varepsilon_1 = n_1^2$) and magnetic permeability $\mu_1$. It is embedded in an infinite "cladding" of uniform refractive index $n_2 (< n_1)$ and magnetic permeability $\mu_2$. Both $\varepsilon_1$ and $\varepsilon_2$ are assumed real so that the system is dissipationless.

The assumption of an infinite cladding has been shown not to be particularly critical in determining the behaviour of the bound modes except very close to cutoff. It may be expected to be more important, however, in the case of leaky modes whose fields extend further from the core. In fact, if the medium outside a cladding of finite radius were of refractive index $n_3 < n_2$, then some of the leaky modes might become

$\mu_1$ and $\mu_2$ are allowed to differ in deriving the modal fields in order to retain symmetry in the expressions for electric and magnetic fields but all calculations will assume that $\mu_1 = \mu_2$. 
proper modes of the cladding structure (so-called "cladding modes\(^{40}\)).

On the other hand, if \( n_1 > n_2 \), all modes are leaky.\(^{20}\) But we will assume that for slowly radiating leaky modes, in which we are most interested, since the modal field has the same behaviour as a bound mode field out to very large distances from the waveguide, the effect of the finite cladding is not too significant. The validity of this assumption is yet to be examined.

As for the assumption of uniform refractive indices, this is essential to the details of the mode calculations but general considerations concerning the origin of leaky modes and their orthogonality and excitation to be presented in Chapter 4 apply equally well to graded index media.

In the system with cylindrical symmetry shown in Fig. 1.1, the requirement that the electromagnetic field be a single-valued function of position implies that the solution of the scalar wave equation is of the form

\[
f(r) \exp[i(\omega t - \beta z - \phi)] ,
\]

where \( \ell \) is an integer. The equation satisfied by \( f \), found by substituting Eq. (2.26) in Eqs. (2.15) and (2.17), is then Bessel's equation

\[
\frac{d^2 f}{dr^2} + \frac{1}{r} \frac{df}{dr} + \left( k^2 - \beta^2 - \frac{s^2}{r^2} \right) f = 0 ,
\]

where \( k = k_1 = \omega (\mu_1 \epsilon_1)^{1/2} \) in the core, and
\[
k_2 = k_2 = \omega (\mu_2 \epsilon_2)^{1/2} \text{ in the cladding.}
\]

In Eqs. (2.26 - 27), \( \beta \) may be either positive (forward modes) or negative (backward modes), the fields having reflection symmetry
about the plane $z = 0$. Since we shall be concerned with problems in which a source is placed in this plane with a semi-infinite cylinder extending in the positive $z$ direction, we shall assume from now on that $\beta > 0$. The magnitude of $\beta$ must be less than $k_1$ in order to give well-behaved solutions of Eq. (2.27) in the core. From the discussion in Section 1.1, we see that if $\beta > k_2$, the field solutions correspond to bound modes, and if $\beta < k_2$, the solution corresponds to a radiation mode.

We now have sufficient background to proceed to a discussion of these modes in greater detail.
REFERENCES


CHAPTER 2
THE PROPER MODES OF A DIELECTRIC WAVEGUIDE
OF CIRCULAR CROSS-SECTION

2.1 INTRODUCTION

A dielectric waveguide is characterized by a spectrum of resonant frequencies and by the modes which propagate at these frequencies. These modes are electromagnetic waves which travel down the waveguide with a well-defined phase velocity, group velocity, cross-sectional intensity distribution and polarization. Each component of their electric and magnetic fields is of the form \( f(r,\phi) e^{i(\omega t - \beta z)} \) (where the \( z \)-axis of the coordinate system coincides with the axis of the guide, \( \beta \) is the \( z \)-component of the wave vector and \( r,\phi \) are polar coordinates in the cross-sectional plane) and the field vectors satisfy the homogeneous wave equation in all the media that make up the guide, as well as the boundary conditions at the interfaces.\(^1\)

Modes which propagate unattenuated in the \( z \)-direction but which decay exponentially away from the waveguide core (and thus carry a finite amount of energy) are "guided", "trapped" or "bound" to the core, like the bound states of a quantum mechanical system. A superposition of these modes can completely describe any guided field distribution in the waveguide but is not sufficient to describe radiation phenomena which require a travelling-wave behaviour in the radial coordinate.\(^2\)
By analogy with the unbound states of a quantum mechanical system, however, field distributions which are not bound to the waveguide core can be completely described by a continuum of "radiation modes". These are modes which do not decay exponentially away from the core and, indeed, considered individually, carry an infinite amount of energy. But because no physical process can excite an individual radiation mode without, at the same time, exciting an infinite number of its neighbours, the resulting superposition of modes gives the desired outward-travelling wave with finite energy.

The excitation of modes (bound or unbound) will be considered in detail in Chapter 3 but first it is necessary to write down the modes themselves. The bound modes have been widely known and used (in a variety of forms, e.g. Refs. 1-8) for some time so we shall list them here without any derivation, both in their exact form and in a useful approximate form which corresponds to the "paraxial ray" approximation of geometric optics. The radiation modes, because they are central to a large part of this thesis, will be derived in some detail (Section 2.3.2).

2.2 BOUND MODES

If we assume an angular dependence of the form $e^{-i\ell z}$ (where $\ell$ is an integer) then the radial behaviour of each of the bound mode field components is as listed below (where the top line gives the behaviour in the core and the lower line in the cladding):

$$e_z = A_{\ell} \begin{cases} J_{\ell} (UR) & R < 1 \\ \frac{n_3}{n_2} K_{\ell} (WR) & R > 1 \end{cases}$$  \hspace{1cm} (2.1)
\[ e_r = \frac{\gamma_1}{\eta_3} \left( \frac{J_{\ell-1}(UR)/U}{K_{\ell-1}(WR)/W} \right) + \frac{\gamma_2}{-\eta_3} \left( \frac{J_{\ell+1}(UR)/U}{-K_{\ell+1}(WR)/W} \right) \]  

\[ e_\phi = -i\gamma_1 \left( \frac{J_{\ell-1}(UR)/U}{\eta_3 K_{\ell-1}(WR)/W} \right) + i\gamma_2 \left( \frac{J_{\ell+1}(UR)/U}{-\eta_3 K_{\ell+1}(WR)/W} \right) \]  

\[ h_z = B_\ell \left( \frac{J_\ell(UR)}{\eta_3 K_\ell(WR)} \right) \]  

\[ h_r = \gamma_3 \left( \frac{J_{\ell-1}(UR)/U}{\eta_3 K_{\ell-1}(WR)/W} \right) + \gamma_4 \left( \frac{J_{\ell+1}(UR)/U}{-\eta_3 K_{\ell+1}(WR)/W} \right) \]  

\[ h_\phi = -i\gamma_3 \left( \frac{J_{\ell-1}(UR)/U}{\eta_3 K_{\ell-1}(WR)/W} \right) + i\gamma_4 \left( \frac{J_{\ell+1}(UR)/U}{-\eta_3 K_{\ell+1}(WR)/W} \right) \]  

where

\[ R = r/\rho \]  

\[ U^2 = \rho^2 (k_1^2 - \beta^2) \]  

\[ W^2 = \rho^2 (\beta^2 - k_2^2) \]  

\[ k_1 = \omega (\mu_1 \epsilon_1)^{1/2} \]  

\[ k_2 = \omega (\mu_2 \epsilon_2)^{1/2} \]  

\[ \eta_3 = J_\ell(U)/K_\ell(W) \]  

\[ \gamma_1 = -\frac{\rho}{2} \left[ i\beta A_\ell + k_n \left( \frac{\epsilon_n}{\mu_n} \right)^{1/2} B_\ell \right] \quad n = \begin{cases} 1 & R < 1 \\ 2 & R > 1 \end{cases} \]  

\[ \gamma_2 = \frac{\rho}{2} \left[ i\beta A_\ell - k_n \left( \frac{\epsilon_n}{\mu_n} \right)^{1/2} B_\ell \right] \]
2.2

\[
\begin{align*}
\gamma_3 &= \frac{\rho}{2} \left[ k_n \left( \frac{\mu n}{\epsilon n} \right)^{1/2} \right. \\
\gamma_4 &= \frac{\rho}{2} \left[ k_n \left( \frac{\mu n}{\epsilon n} \right)^{1/2} \right. \\
A_n/B_n &= \frac{\ln}{\left( \frac{\mu n}{\epsilon n} \right)^{1/2}} = \frac{2 \beta}{\left( \frac{\mu n}{\epsilon n} \right)^{1/2}} - \frac{\ln}{\left( \frac{\mu n}{\epsilon n} \right)^{1/2}} \\
\nu^2 &= U^2 + W^2 = \rho^2 (k_1^2 - k_2^2) \\
n_1 &= \frac{K_L'(W)}{W K_L(W)} + \frac{\mu_1}{\mu_2} \frac{J_L'(U)}{U J_L(U)} \\
n_2 &= \frac{K_L'(W)}{W K_L(W)} + \frac{\epsilon_1}{\epsilon_2} \frac{J_L'(U)}{U J_L(U)} \\
\end{align*}
\]

and \(J_L, K_L\) denote the usual Bessel and modified Hankel functions, and the prime indicates differentiation with respect to the argument.

It remains to specify \(\beta\) and \(A_n\) to completely determine the modes. Because the boundary conditions give four linear, homogeneous equations in four unknowns, they lead to a consistency condition which must be satisfied by any mode. This condition, otherwise known as the eigenvalue equation, specifies the allowed values of \(\beta\) (and hence \(U, W\)) and, from Eq. (2.17), takes the form

\[
\eta_1 \eta_2 - \left( \frac{\beta}{k_2} \right)^2 \left( \frac{V}{U W} \right)^4 = 0 .
\]

Real solutions to Eq. (2.21) only exist for \(U \ll V\) so that the value of \(V\) at which \(U = V\) defines a "cutoff frequency" for the mode in question. Below this frequency, at which \(\beta = k_2\), we are in the radiation mode region. Thus for guided modes, \(\beta\) lies between \(k_2\) and \(k_1\).
The remaining factor, $A_k$, is found by imposing some normalization condition on the modes. For an ideal (i.e. non-absorbing) waveguide a condition commonly used is

$$\int_{S_\infty} e_p \times h_p^* \cdot \hat{z} \ da = 1 , \quad (2.22)$$

where $S_\infty$ represents an infinite cross-sectional area, $\hat{z}$ is the unit vector in the z-direction and the asterisk represents complex conjugation. $e_p$ and $h_p$ are the electric and magnetic fields, respectively, of mode $p$, where $p$ is a symbolic label containing all information required to specify a particular mode. Imposition of this condition leads to the equation

$$\frac{1}{U_p^2} \left[ \gamma_1^j(1) \gamma_3^j(1)* \left( J_{\ell-2}^2 - J_\ell J_{\ell-2}^* \right) - \gamma_2^j(1) \gamma_4^j(1)* \left( J_{\ell+1}^2 - J_\ell J_{\ell+2}^* \right) \right]$$

$$+ \frac{\eta_0^2}{W_p^2} \left[ \gamma_1^j(2) \gamma_3^j(2)* \left( K_\ell K_{\ell-2} - K_{\ell+1}^2 \right) - \gamma_2^j(2) \gamma_4^j(2)* \left( K_\ell K_{\ell+2} - K_{\ell+1}^2 \right) \right] = \frac{1}{2\pi i \rho^2} , \quad (2.23)$$

from which $A_k$ can be found. In Eq. (2.23) $\gamma_1^j(1)$ represents $\gamma_j$ in medium 1 and $\gamma_1^j(2)$ represents $\gamma_j$ in medium 2; the unstated argument of all Bessel functions of the first kind is $U_p$ and of the modified Hankel functions, $W_p$.

An individual bound mode is now completely specified. But it is also necessary to know the relationship between two different bound modes (and between a bound mode and the radiation field). In an axially uniform waveguide of inhomogeneous and arbitrarily-shaped cross section, this relationship is stated via the orthogonality condition

$$\int_{S_\infty} e_p \times h_q \cdot \hat{z} \ da = 0 \quad \text{if } p \neq q \quad (2.24)$$

where $p$ and $q$ stand for any two bound modes or for a bound mode and the radiation field.
In the ideal (lossless) guide with which we are concerned, a more useful orthogonality condition, power orthogonality, also holds. This takes the form

\[ \int_{S_\infty} \int e_p^* h_q \cdot \hat{E} \, da = 0 \quad \text{if } p \neq q \quad (2.25) \]

and states that, in an ideal guide, the time-averaged power carried by any field distribution can be considered to be divided among the modes (by contrast with the non-ideal guide in which each mode does not carry a fixed amount of power but there is a continual interchange among the modes).

Combination of Eqs. (2.22) and (2.25) gives, for the bound modes, the orthonormality condition

\[ \int_{S_\infty} \int e_p^* h_q^* \cdot \hat{E} \, da = \delta_{pq} \quad (2.26) \]

where \( \delta_{pq} \) is the Kronecker delta.

In many practical situations, it is possible to use a much simplified form of the modes presented in Eqs. (2.1-6). If we define the parameters \( \delta, \theta_p \) by

\[ \delta = 1 - \frac{\varepsilon_2}{\varepsilon_1} \quad (2.27) \]

and

\[ \theta_p = \delta^2 \frac{U_p}{V} \quad (2.28) \]

then most fibres of practical interest are those with \( \delta << 1 \) and hence (since for bound modes, \( U_p << V \)) \( \theta_p << 1 \). Under these circumstances, Snyder has derived the following simplified transverse modal fields (when \( \mu_1 = \mu_2 = \mu \)):

\[ \]
\[ \varepsilon_0 = \left( \frac{\mu}{\varepsilon_1} \right)^{\frac{1}{2}} \]  
\[ \mu_0 = \hat{\mu} \ 
\]

\[ \psi_p = \psi_p \left( \frac{\sin \phi}{\cos \phi} + \frac{\cos \phi}{-\sin \phi} \right) \]  
\[ f_\lambda(r) \]  
\[ f_\lambda(r) = f_0(R) \quad \text{TE}_{0m} \]  
\[ f_\lambda(r) = -f_0(R) \quad \text{TM}_{0m} \]  

where

\[ f_\lambda(r) = \frac{J_{\lambda+1}(UR)}{J_{\lambda+1}(U)} \quad R \leq 1 \]  
\[ = \frac{K_{\lambda+1}(WR)}{K_{\lambda+1}(W)} \quad R \geq 1 \]  
\[ \psi_p = \pi \rho^2 \left( \frac{\varepsilon_1}{\mu} \right)^{\frac{1}{2}} \left( \frac{V}{U} \right)^{\frac{1}{2}} \xi \]  
\[ \xi = \frac{K_\lambda^2(W) K_{\lambda+2}(W)}{K_{\lambda+1}(W)} \]  

The longitudinal fields are of order \( \theta_p \).

In the above equations TE\textsubscript{0m} and TM\textsubscript{0m} represent transverse electric and transverse magnetic modes respectively, and, when the double sign notation is used, the upper sign is for HE\textsubscript{lm} modes, the lower for EH\textsubscript{lm} modes. (The differences between these two classes of modes which are distinct solutions of Eq. (2.21) at the same value of V are described by Snitzer.\textsuperscript{3})

These simplified modes also satisfy the orthonormality condition, Eq. (2.26), \( \psi_p \) being the normalization factor.

As stated in Section 2.1, these modes correspond to the paraxial ray approximation in geometric optics because when \( \theta_p \ll 1 \), \( \beta_p \approx k_1 \). This means that the modal wave-vector lies almost parallel to
the waveguide axis. This condition is not as restrictive as it may sound, however, and it has been found that these simplified modes are very widely applicable. 7

Approximate forms for the bound modes have also been given by Gloge 8 and Kapany and Burke. 1 Since our major concern is with the radiation modes, however, we will leave the discussion of bound modes here and return to them only as they are needed in specific numerical calculations (Section 5.4), where Eqs. (2.29-35) will be used.

2.3 RADIATION MODES

2.3.1 Properties and Methods of Derivation

The modes of the continuous spectrum can be obtained in a number of ways. Briefly, these are:

(i) to consider the modes as the limit of fields in a dielectric rod of finite radius \( \rho \) surrounded by a concentric cylindrical metal wall of radius \( \rho_1 \). As \( \rho_1 \to \infty \), the complete set of properly normalized modes of the dielectric rod in free space is found. This method corresponds to the "normalization in a box" familiar in quantum mechanics 9 and has been widely used in waveguide theory (e.g. Refs. 10-13). An obvious disadvantage of the approach is that, because an additional boundary must be introduced, the algebraic complexities of the problem are escalated.

(ii) to integrate the dyadic Green's function in the complex eigenvalue plane, the continuous spectrum being identified with branch cut singularities. In principle this method displays the modes excited by an arbitrarily oriented source and it has therefore been widely used in studying the slab waveguide (see for example Ref. 9). In practice
for the circular fibre, however, the Green's function can only be obtained in closed form under very special circumstances\textsuperscript{14-16} so that this method is not yet practically useful.

(iii) to solve the homogeneous wave equation directly for the dielectric rod in an unbounded dielectric medium. This is the approach most recently used by, for example, Snyder,\textsuperscript{17} Kapany and Burke\textsuperscript{1} and Marcuse.\textsuperscript{2} The advantage of this method, particularly in the formulation used by Snyder, is that the resulting fields are such that a physically intuitive interpretation can be given to the modes and hence their normalization can be carried out with relatively little difficulty. This is the approach to be followed below.

In principle the last approach is identical with that used to find the bound modes, but there arises one complicating factor which is encountered neither in the latter problem nor in finding the radiation modes for the slab waveguide. This difficulty is most clearly shown in the solution presented by Marcuse,\textsuperscript{2} but briefly it can be explained as follows.

In the case of the bound modes, the "source" of the field is considered to be within the waveguide at $z = -\infty$ (or $+\infty$ for backward travelling modes) so that the field outside the guide is evanescent in its radial coordinate. Thus the radial behaviour of the $z$-component electromagnetic fields must be of the form $J_\ell(UR)$ inside the guide and $K_\ell(WR)$ outside. One therefore has four undetermined coefficients (coefficients of $J_\ell$ and $K_\ell$ for both electric and magnetic fields) and four boundary conditions with which to specify these (up to a normalization constant). For the radiation modes, however, because the fields extend to infinity in the radial direction with only an $r^{-\ell}$
2.3.1

decay, the field outside must be described in terms of the two linearly independent functions $J_{\ell}$ and $Y_{\ell}$, or combinations of these. Thus the number of undetermined coefficients has risen to six while the number of boundary conditions remains unchanged. One of the extra coefficients is related to the energy carried by the mode; but the remaining one remains arbitrary.

The reason this difficulty does not arise in treating the slab waveguide is that in that case, one can naturally separate the modes from the beginning into even and odd sets. Such a natural separation is not, however, immediately apparent for the circular fibre although there do exist two independent mode sets which need to be separated.

Marcuse\textsuperscript{2} overcomes this problem by superimposing pairs of modes and choosing the arbitrary constants in such a way as to form two mutually orthogonal sets of such composite modes. But, as stated above, the fields of the radiation mode can be given a clear physical interpretation and Snyder\textsuperscript{17} has overcome the difficulty of the arbitrary constant by invoking this physical picture. Thus if one writes the field in the cladding in terms of Hankel functions $H_{\ell}^{(1)}$, $H_{\ell}^{(2)}$ rather than Bessel functions, it is immediately clear that this field can be thought of as a superposition of incoming and outgoing cylindrical waves and the total field as that resulting from the scattering by a circular rod of an incident cylindrical wave. The separation into two distinct mode sets is then achieved by considering the polarization of the incident field – an incident transverse magnetic field leads to one mode set while an incident transverse electric field leads to another, the two sets being mutually orthogonal, as we shall prove in Section 2.4.
The fact that the radiation field decomposes naturally into mode sets with orthonormality conditions based entirely on the properties of extremely simple incident fields is a major advantage of this method.

Another practical advantage in approaching the radiation modes from this point of view is that because the scattering problem has already received a great deal of attention (e.g. Ref. 18), very little additional work is required to determine the modes. So we shall now trace through the derivation of the fields resulting from the scattering of an obliquely incident plane wave from a circular cylinder. The result, originally due to Wait, is now very well known and was first used in this form by Snyder. However, as we shall need to write the fields and establish a considerable amount of notation for later use anyway, the derivation is reproduced fairly completely.

2.3.2 Radiation Modes from Scattering

Consider then the system previously shown in Fig. 1.1 but now extended to $-\infty$ in the $z$ direction. We wish to find the electromagnetic fields in media 1 and 2 resulting from the scattering of a plane wave with angular frequency $\omega$ incident at angle $\theta$ to the cylinder axis.

We begin with the situation where the electric vector is parallel to the plane $\phi=0$ (so that the incident magnetic field is purely transverse — Fig. 2.1). The $z$-component of the electric field of the incident wave is then given by

$$E_z^i = E_0 \sin \theta \ e^{i\lambda R \cos \phi} \ e^{i(\omega t - \beta z)}, \quad (3.1)$$

where


\[ \lambda = \rho k_2 \sin \theta, \quad (3.2) \]

\[ \beta = k_2 \cos \theta, \quad (3.3) \]

and \( E_0 \) is a normalization constant which shall remain unspecified for the moment. \( k_2 \) and \( R \) are defined by Eqs. (2.11, 2.7) respectively.

Since, throughout this section, the cylinder is assumed infinite, all fields will contain the factor \( e^{i(\omega t - \beta z)} \). For simplicity of notation we therefore suppress this factor in the following equations.

Using the fact that\(^4\)

\[ e^{i\lambda R \cos \phi} = \sum_{\lambda = -\infty}^{\infty} i^\lambda J_\lambda(\lambda R) e^{-i\lambda \phi}, \quad (3.4) \]

we can rewrite Eq. (3.1) as
The $z$-component of the scattered electric field can then be expanded in the form

$$ E^s_z = E_0 \sin \theta \sum_{\ell=-\infty}^{\infty} i\ell J_{\ell}^{(2)}(\lambda R) e^{-i\ell \phi}, \quad (3.5) $$

where the Hankel function of the second kind is chosen to give an outward-going cylindrical wave at infinity.

Now whereas the field components $E^i_z$ and $E^s_z$ satisfy the scalar Helmholtz equation appropriate to medium 2, namely

$$ [\rho^2 V_t^2 + \lambda^2] f = 0, \quad (3.7) $$

the fields inside must satisfy the equation

$$ [\rho^2 V_t^2 + U^2] f = 0, \quad (3.8) $$

$V_t$ being the transverse component of the gradient operator and $U$ being defined by Eq. (2.8).

The longitudinal component of the electric field inside the cylinder is therefore written

$$ E^i_z = \sum_{\ell=-\infty}^{\infty} a_{\ell} J_{\ell}^{(1)}(UR) e^{-i\ell \phi}, \quad (3.9) $$

where only the Bessel function of the first kind is used in order to avoid singularities along the axis.

The $z$-components of the incident, scattered and internal magnetic fields are, respectively,

$$ H^i_z = 0 \quad (3.10) $$
The transverse (r and \(\phi\)) components of these fields can be found from the following relations derived from Eqs. (1.2.24 - 25)

\[
E_r = (k^2 - \beta^2)^{-1} \left[ - i\beta \frac{\partial E_z}{\partial r} - \frac{\beta k}{r} \left( \frac{1}{\varepsilon} \right) H_j \right], \tag{3.13}
\]

\[
E_\phi = (k^2 - \beta^2)^{-1} \left[ - \frac{\beta k}{r} E_z + ik \left( \frac{1}{\varepsilon} \right) \frac{\partial H_j}{\partial r} \right], \tag{3.14}
\]

where \(k, \varepsilon\) and \(\mu\) are the quantities appropriate to either medium 1 or 2.

It is now necessary to impose boundary conditions on these fields in order to determine the quantities \(a_j, b_j, a_j^S, b_j^S\). The conditions of interest are the continuity of the tangential components of \(E\) and \(H\) across the surface of the cylinder as discussed in Section 1.2.

We need therefore find from Eqs. (3.13, 3.14) only the \(\phi\)-components of \(E\) and \(H\) in the two regions. These are given by

\[
E_\phi^i = \sum_{j=0}^{\infty} - E_0 \frac{\beta k}{Rk_j^2} - i \frac{j}{\lambda} J_j^2 (\lambda R) e^{-i\lambda \phi}, \tag{3.15}
\]

\[
E_\phi^S = \sum_{j=0}^{\infty} \left[ - \frac{\beta k}{\varepsilon^2} \frac{\partial k_j^2}{\lambda} H_j^2 (\lambda R)^2 + i \left( \frac{\mu_j^2}{\varepsilon^2} \right) \frac{\partial k_j^2}{\lambda} + b_j^S H_j^2 (\lambda R) \right] e^{-i\lambda \phi}. \tag{3.16}
\]
2.3.2

\[
E_\phi = \sum_{l=-\infty}^{\infty} \left[ -\frac{j_0 \beta}{R \lambda^2} a_\lambda J_\lambda (\nu) + i \left( \frac{\varepsilon_1}{\mu_1} \right)^{\frac{1}{2}} \rho k_1 \frac{\beta}{U} b_\lambda J_\lambda '(\nu) \right] e^{-il\phi} \quad (3.17)
\]

\[
H_\phi^i = \sum_{l=-\infty}^{\infty} i^{l-1} E_0 \left( \frac{\varepsilon_2}{\mu_2} \right)^{\frac{1}{2}} J_\lambda '(\nu) e^{-il\phi} \quad (3.18)
\]

\[
H_\phi^s = \sum_{l=-\infty}^{\infty} \left[ -\frac{j_0 \beta}{R \lambda^2} b_\lambda J_\lambda (2) (\nu) - i \frac{\rho k_2}{\lambda} \left( \frac{\varepsilon_2}{\mu_2} \right)^{\frac{1}{2}} a_\lambda J_\lambda (2)' (\nu) \right] e^{-il\phi} \quad (3.19)
\]

\[
H_\phi = \sum_{l=-\infty}^{\infty} \left[ -\frac{j_0 \beta}{R \lambda^2} b_\lambda J_\lambda (\nu) - i \left( \frac{\varepsilon_1}{\mu_1} \right)^{\frac{1}{2}} \rho k_1 \frac{\beta}{U} a_\lambda J_\lambda '(\nu) \right] e^{-il\phi} \quad (3.20)
\]

Then, from the boundary conditions at \( r = \rho \ (R = 1) \)

\[
E_z^i + E_z^s = E_z \quad (3.21)
\]

\[
E_\phi^i + E_\phi^s = E_\phi \quad (3.22)
\]

\[
H_z^i + H_z^s = H_z \quad (3.23)
\]

\[
H_\phi^i + H_\phi^s = H_\phi \quad (3.24)
\]

we find

\[
a_\lambda = \left( E_0 \sin \theta i \frac{J_\lambda (\lambda)}{J_\lambda (\nu)} + a_\lambda J_\lambda (2) (\lambda) \right) / J_\lambda (\nu) \quad (3.25)
\]

\[
b_\lambda = b_\lambda J_\lambda (2) (\lambda) / J_\lambda (\nu) \quad (3.26)
\]

\[
a_\lambda^s = -E_0 \sin \theta i \frac{J_\lambda (\lambda)}{H_\lambda (2) (\lambda)} + \frac{2i}{\pi \lambda^2} \frac{N_2 - \frac{\mu_1}{\mu_2} N_1}{H_\lambda (2) (\lambda)} \quad (3.27)
\]

\[
b_\lambda^s = -\frac{2}{\pi} \frac{(\varepsilon_2)}{\mu_2} i^{l-1} \frac{\beta \rho}{\mu \lambda^2} \frac{V}{\rho k_2 U H_\lambda (2) (\lambda)} \quad (3.28)
\]

where

\[
N_1 = J_\lambda (\nu) / (U J_\lambda (\nu)) \quad (3.29)
\]

\[
N_2 = H_\lambda (2) (\lambda) / (\lambda H_\lambda (2) (\lambda)) \quad (3.30)
\]
The solution of the scattering problem for this particular polarization of the incident field is now complete. The solution for the case when the incident magnetic field vector lies in the plane \( \phi = 0 \) (Fig. 2.2) is obtained from the above results by making the following transformations in Eqs. (3.1-28):

\[
\begin{align*}
E &\rightarrow H \\
H &\rightarrow -E \\
\varepsilon &\leftrightarrow \mu
\end{align*}
\]
Thus by varying the angle of incidence $\theta$ of the incoming plane wave we can obtain two sets of "characteristic fields" of the cylinder, one for each polarization. We do not obtain all the modes in each of these sets in this way however. From Eqs. (3.2) and (3.3) we see that the modes obtained are those for which $-\rho k_2 \leq \lambda \leq \rho k_2$ and hence $\beta$ is real. These modes are the propagating radiation modes but there exists, as well as these, an infinite number of evanescent radiation modes (i.e. modes for which $\beta$ is imaginary). The evanescent modes are required for a complete description of the radiation field and, while they do not carry energy away from a source or imperfection, are necessary to describe the field exactly in the vicinity of such a source or imperfection. If we had followed approach (i) listed above, these evanescent modes would have arisen as the below-cutoff modes of the enclosing metallic waveguide. In the present approach, we obtain them simply by extending the formulae derived for $|\lambda| \leq \rho k_2$ beyond this region. While they no longer have the physical interpretation of the propagating modes, these fields nevertheless satisfy the wave equation and boundary conditions so the extension is perfectly valid.

In this way then we have obtained two sets of continuum modes which together form a complete set for the expansion of the radiation field, the "incident transverse magnetic" (ITM) modes, for which $H_z^i = 0$; and the "incident transverse electric" (ITE) modes, for which $E_z^i = 0$.

For later use, we list the components of these modes below (the factor $e^{-il\phi}$ is omitted from each component):
2.3.2 ITM modes

In the core \((R < 1)\),

\[
\psi_{M} = a_{M} J_{\lambda}(UR) \tag{3.33}
\]

\[
\psi_{M} = \frac{\rho}{U} \left[ - i \beta a_{M} J_{\lambda}(UR) - \frac{\mu}{\xi_{2}} k_{1} b_{M} J_{\lambda}(UR) \right] \tag{3.34}
\]

\[
\psi_{M} = \frac{\rho}{U} \left[ - \frac{\beta}{R} a_{M} J_{\lambda}(UR) + i \frac{\mu}{\xi_{1}} k_{1} b_{M} J_{\lambda}(UR) \right] \tag{3.35}
\]

\[
\psi_{M} = b_{M} J_{\lambda}(UR) \tag{3.36}
\]

\[
\psi_{M} = \frac{\rho}{U} \left[ - \frac{\beta}{R} a_{M} J_{\lambda}(UR) - i \beta b_{M} J_{\lambda}(UR) \right] \tag{3.37}
\]

\[
\psi_{M} = \frac{\rho}{U} \left[ - i \frac{\mu}{\xi_{2}} k_{1} a_{M} J_{\lambda}(UR) - \frac{\beta}{R} b_{M} J_{\lambda}(UR) \right] \tag{3.38}
\]

and in the cladding \((R > 1)\),

\[
\psi_{M} = \frac{\lambda}{\rho k_{2}} J_{\lambda}(\lambda R) + a_{M} H_{\lambda}(2)(\lambda R) \tag{3.39}
\]

\[
\psi_{M} = - \frac{i \beta}{k_{2}} J_{\lambda}(\lambda R) - \frac{i \beta \rho}{\lambda} a_{M} H_{\lambda}(2)(\lambda R) - \frac{\mu}{\xi_{2}} \frac{\rho k_{2}}{R \lambda^{2}} b_{M} H_{\lambda}(2)(\lambda R) \tag{3.40}
\]

\[
\psi_{M} = - \frac{\beta \rho}{R \lambda k_{2}} J_{\lambda}(\lambda R) - \frac{\beta \rho}{\lambda} a_{M} H_{\lambda}(2)(\lambda R) + i \frac{\mu}{\xi_{2}} \frac{\rho k_{2}}{\lambda} b_{M} H_{\lambda}(2)(\lambda R) \tag{3.41}
\]

\[
\psi_{M} = b_{M} H_{\lambda}(2)(\lambda R) \tag{3.42}
\]

\[
\psi_{M} = \frac{\xi_{2}}{\mu_{2}} \frac{\rho k_{2}}{R \lambda^{2}} J_{\lambda}(\lambda R) + \frac{\xi_{2}}{\mu_{2}} \frac{\rho k_{2}}{R \lambda^{2}} a_{M} H_{\lambda}(2)(\lambda R) \tag{3.43}
\]

\[
- \frac{i \beta \rho}{\lambda} b_{M} H_{\lambda}(2)'(\lambda R) \]
\[
\psi_{3/2, M} = -i \left( \frac{e_2}{\mu_2} \right)^{1/2} J_{2/1}^i (\lambda R) - i \left( \frac{e_2}{\mu_2} \right)^{1/2} \frac{\rho k_2}{\lambda} a_{\lambda} \cdot H_{\lambda}^M (2)^i (\lambda R) \\
- \frac{\beta \rho}{R \lambda^2} b_{\lambda} \cdot H_{\lambda}^M (2) (\lambda R), \quad (3.44)
\]

where

\[
a_{\lambda} = - \frac{2i}{\pi \rho k_2 \lambda} \times \frac{N_2 - \frac{\mu_1}{\mu_2} N_1}{M J_{\lambda}^i (U) H_{\lambda}^M (2) (\lambda)} \quad (3.45)
\]

\[
b_{\lambda} = \frac{H_{\lambda}^M (2) (\lambda)}{J_{\lambda}^i (U)} \times b_{\lambda} \quad (3.46)
\]

\[
a_{\lambda} = - \frac{\lambda}{\rho k_2} \left[ \frac{J_{\lambda}^i (\lambda)}{H_{\lambda}^M (2) (\lambda)} + \frac{2i \left[ N_2 - \frac{\mu_1}{\mu_2} N_1 \right]}{\pi \lambda^2 M \left[ H_{\lambda}^M (2) (\lambda) \right]^2} \right] \quad (3.47)
\]

\[
b_{\lambda} = - \frac{2i}{\pi \left( \frac{e_2}{\mu_2} \right)^{1/2} \frac{\beta \rho}{\lambda^3 M} \left[ \frac{V}{\rho k_2 U H_{\lambda}^M (2) (\lambda)} \right]^2} \quad (3.48)
\]

and \( \psi_{3/2} \) is a normalization constant to be determined below (Section 2.4).

**ITE modes**

In the core,

\[
\psi_{3/2, E_z} = - b_{\lambda} E_{\lambda}^z (UR) \quad (3.49)
\]

\[
\psi_{3/2, E_r} = \frac{\rho}{U} \left[ i \beta b_{\lambda} E_{\lambda}^r (UR) - \frac{\mu_1}{\mu_2} \frac{1}{e_1} \frac{k_1}{E_1} a_{\lambda} E_{\lambda}^i (UR) \right] \quad (3.50)
\]

\[
\psi_{3/2, E_\psi} = \frac{\rho}{U} \left[ i \beta b_{\lambda} E_{\lambda}^\psi (UR) + \frac{\mu_1}{\mu_2} \frac{1}{e_1} \frac{k_1}{E_1} a_{\lambda} E_{\lambda}^i (UR) \right] \quad (3.51)
\]

\[
\psi_{3/2, E_z} = a_{\lambda} E_{\lambda}^z (UR) \quad (3.52)
\]
2.3.2

\[
\psi_{E}^{\frac{1}{2}} H_{r} = \frac{\rho}{U} \left[ \frac{\epsilon_{1}}{\mu_{1}} \frac{\lambda k_{1}}{R \lambda} b_{\lambda} E_{J_{\lambda}}(UR) - i \beta a_{\lambda} E_{J_{\lambda}}^{*}(UR) \right] \tag{3.53}
\]

\[
\psi_{E}^{\frac{1}{2}} H_{\phi} = \frac{\rho}{U} \left[ \frac{i \epsilon_{1}}{\mu_{1}} k_{1} b_{\lambda} E_{J_{\lambda}}^{'}(UR) - \frac{\beta \rho}{R \lambda} a_{\lambda} E_{J_{\lambda}}^{*}(UR) \right] \tag{3.54}
\]

and in the cladding

\[
\psi_{E}^{\frac{1}{2}} E_{z} = - b_{\lambda} SE_{H_{\lambda}}^{(2)}(\lambda R) \tag{3.55}
\]

\[
\psi_{E}^{\frac{1}{2}} E_{r} = - \left( \frac{\mu_{2}}{\epsilon_{2}} \right)^{\frac{1}{2}} \frac{\rho}{R \lambda} J_{\lambda}(\lambda R) + \frac{i \beta \rho}{\lambda} b_{\lambda} SE_{H_{\lambda}}^{(2)}(\lambda R) \tag{3.56}
\]

\[
\psi_{E}^{\frac{1}{2}} E_{\phi} = i \left( \frac{\mu_{2}}{\epsilon_{2}} \right)^{\frac{1}{2}} J_{\lambda}^{'}(\lambda R) + \frac{\beta \rho}{\lambda} b_{\lambda} SE_{H_{\lambda}}^{(2)}(\lambda R) \tag{3.57}
\]

\[
\psi_{E}^{\frac{1}{2}} H_{z} = \frac{\lambda}{\rho k_{2}} J_{\lambda}(\lambda R) + a_{\lambda} SE_{H_{\lambda}}^{(2)}(\lambda R) \tag{3.58}
\]

\[
\psi_{E}^{\frac{1}{2}} H_{r} = - \frac{i \beta}{k_{2}} J_{\lambda}^{'}(\lambda R) - \left( \frac{\epsilon_{2}}{\mu_{2}} \right)^{\frac{1}{2}} \frac{\lambda \rho k_{2}}{R \lambda^{2}} b_{\lambda} SE_{H_{\lambda}}^{(2)}(\lambda R) \tag{3.59}
\]

\[
\psi_{E}^{\frac{1}{2}} H_{\phi} = - \frac{\lambda \rho}{R \lambda k_{2}} J_{\lambda}(\lambda R) + i \left( \frac{\epsilon_{2}}{\mu_{2}} \right)^{\frac{1}{2}} \frac{\lambda \rho k_{2}}{\lambda} b_{\lambda} SE_{H_{\lambda}}^{(2)}(\lambda R) \tag{3.60}
\]

where

\[
a_{\lambda}^{E} = a_{\lambda}^{M}(\mu \rightarrow \epsilon), \tag{3.61}
\]

\[
b_{\lambda}^{E} = \frac{\mu_{2}}{\epsilon_{2}} b_{\lambda}^{M}, \tag{3.62}
\]

\[
a_{\lambda}^{SE} = a_{\lambda}^{SM}(\mu \rightarrow \epsilon), \tag{3.63}
\]
and $\psi_E$ is another normalization constant (which we shall now calculate). An alternative notation for the transverse components of the ITM and ITE modes is given in Appendix 2.B.

### 2.4 Orthonormality of the Radiation Modes

Because the radiation modes have infinite energy, the orthonormality condition Eq. (2.26) can no longer be expected to apply. In its place we must have a condition of the form

$$\int \int_{S_\infty} \mathcal{E}^{(i)}(\lambda, \ell) \times \mathcal{H}^{(i')}*(\lambda', \ell') \cdot \hat{\mathcal{E}} \, da = \delta_{ii'} \delta_{\ell\ell'} \delta(\lambda - \lambda'),$$

where the superscript $i$ and $i'$ indicate linear combinations of ITM and ITE modes. The values of $\psi_M$ and $\psi_E$ and the appropriate linear combinations can be found from direct calculation using the modal fields found in the previous section (see Appendix 2.A). The algebra involved in these calculations is, however, rather horrendous. Fortunately, there is an alternative (and far more elegant) approach which makes use of the physical interpretation given to the radiation modes in the previous section, namely that they can be considered to be the fields resulting from a scattering process. Mathematically, the orthonormality conditions result from the generalization to the vector wave equation of a theorem used implicitly by Møller and proven only for the scalar equation by Friedman.

Snyder suggested that such a generalization could be used to normalize the radiation modes but did not carry this through to prove the orthogonality of ITM and ITE modes. In fact, he incorrectly stated...
that this orthogonality only held for the azimuthally symmetric radiation modes. This error has subsequently been repeated by others\textsuperscript{23,24} (including the present author).

In this section, we prove the vector generalization of Friedman's scalar result. In so doing, we prove the physically plausible result that the orthonormality conditions which apply between radiation modes of any open, dissipationless waveguide (of arbitrary cross-section) are identical with those applying to the incident plane waves from which they are derived. The generalization is a trivial one but requires a certain amount of notation in order to cast the boundary-value problem in operator terms.

For harmonic fields with time dependence $e^{i\omega t}$, Maxwell's equations are

$$\omega \varepsilon \cdot \mathbf{E} + \nabla \times (i\mu) = 0 \quad (4.2)$$

$$\nabla \times \mathbf{E} + \omega \mu \cdot (i\varepsilon) = 0 \quad (4.3)$$

where $\varepsilon$ and $\mu$ are permittivity and permeability dyadics.

We wish to re-cast these equations in the form of an operator equation involving only the transverse parts of the electric and magnetic fields. So, following Bresler, Joshi and Marcuvitz\textsuperscript{25} we write

$$\varepsilon = \varepsilon_{zt} + \varepsilon_{z} \hat{z} \hat{z} + \varepsilon_{zt} \hat{z} \hat{z} + \varepsilon_{zt} \hat{z} \hat{z}$$

where $\varepsilon_{zt}$ is a transverse dyadic, $\varepsilon_{z}$ a scalar, $\varepsilon_{zt}$ and $\varepsilon_{zt}$ vectors, and a similar breakup is introduced for the $\mu$ dyadic.

We also define the operators
\[ \Gamma_{\Xi z} = \begin{pmatrix} 0 & i\hat{z} \times \mathbf{I}^t \\ i\hat{z} \times \mathbf{I}^t & 0 \end{pmatrix} \], \quad (4.5) \\
\[ \mathbf{I}^t = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I}^t \end{pmatrix} \], \quad (4.6) \\

and

\[ \mathbf{L}_{\Xi} = \begin{pmatrix} \omega \mathbf{I}_{\Xi t} - \frac{1}{\mu_z} \mathbf{E}_{\Xi t} \mathbf{E}_{\Xi z} t - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} + \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} - \frac{1}{\mu_z} \mathbf{E}_{\Xi t} \mathbf{E}_{\Xi z} t - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} \end{pmatrix} \]

\[ \quad - \frac{1}{\mu_z} \mathbf{E}_{\Xi t} \mathbf{E}_{\Xi z} t - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} \mathbf{L}_{\Xi} = \begin{pmatrix} \omega \mathbf{I}_{\Xi t} - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} \mathbf{L}_{\Xi} = \begin{pmatrix} \omega \mathbf{I}_{\Xi t} - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} \end{pmatrix} \]

where \( \mathbf{I} \) is the unit dyadic, \( \mathbf{V}_{\Xi t} \) is the transverse gradient operator and \( \mathbf{I}_{\Xi t} = \mathbf{I} - \frac{1}{\mu_z} \mathbf{E}_{\Xi t} \mathbf{E}_{\Xi z} t \) so that \( \mathbf{I}_{\Xi t} \) is the unit operator in a "transverse four-space". The elements of this four-space are the transverse field vectors, \( \mathbf{V} \), defined by

\[ \mathbf{V} = \mathbf{I}_{\Xi t} \cdot \mathbf{V} \]. \quad (4.8) \\

where

\[ \mathbf{\Phi} = \begin{pmatrix} \mathbf{E} \\ i \mathbf{H} \end{pmatrix} \]. \quad (4.9) \\

Then for fields with a \( z \)-dependence of the form \( e^{-i\beta z} \), Eqs. (4.2, 3) can be combined to give

\[ (\mathbf{L} - \beta \Gamma_{\Xi}) \cdot \mathbf{V} = 0 \]. \quad (4.10) \\

In an isotropic, homogeneous medium, the operator \( \mathbf{L}_{\Xi} \) simplified to

\[ \mathbf{L}_{\Xi} = \begin{pmatrix} \omega \mathbf{I}_{\Xi t} - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} & 0 \\ 0 & \omega \mathbf{I}_{\Xi t} - \frac{1}{\omega} \mathbf{V}_{\Xi t} \times \hat{z} \frac{1}{\mu_z} \hat{z} \times \mathbf{V}_{\Xi t} \end{pmatrix} \]. \quad (4.11)
2.4

So if \( \mathbf{u} \) is a field vector describing the incident free-space fields of Section 2.3, then it satisfies

\[
(L - \beta_{\infty}) \cdot \mathbf{u} = 0
\]

(4.12)

with \( \varepsilon = \varepsilon_2 \) and \( \mu = \mu_2 \) in Eq. (4.11).

On the other hand, the field vector \( \mathbf{v} \) describing the total field satisfies the perturbed equation

\[
(L - \beta_{\infty} + M) \cdot \mathbf{v} = 0,
\]

(4.13)

where \( M \) is an operator which takes account of the discontinuity in \( \varepsilon \) and \( \mu \) across the waveguide surface. For a fibre with arbitrary, radially symmetric variation in \( \varepsilon \) and \( \mu \) in cross-section, we have

\[
M = H(\rho - r)
\]

(4.14)

where

\[
H(\rho - r) = \begin{cases} 1 & r < \rho \\ 0 & r > \rho \end{cases}
\]

(4.15)

Thus we can write

\[
(L - \beta_{\infty}) \cdot (\mathbf{v} - \mathbf{u}) = - M \cdot \mathbf{v}
\]

or

\[
\mathbf{v} = \mathbf{u} - \lim_{\zeta \to 0} (L - \beta_{\infty} - i \zeta)_{\infty}^{-1} \cdot M \cdot \mathbf{v},
\]

(4.16)

where the inverse function is some integral operator. (Note that while \( \Gamma_{\infty} \) is singular in the "six-space" in which the vectors \( \mathbf{v} \) are defined, in "four-space" \( \Gamma_{\infty}^{-1} = \Gamma_{\infty} \).)
Before proceeding, we define an inner product between two field vectors in four-space by

$$\langle \psi_1, \psi_2 \rangle = \int \int \left[ E_{t_1}^* \cdot E_{t_2} + (iH_{t_1})^* \cdot (iH_{t_2}) \right] \, da$$  \hspace{1cm} (4.17)

and the Hermitian adjoint $E^\dagger$ to an operator $E$ by

$$\langle \psi, E^\dagger \psi \rangle = \langle E^\dagger \psi, \psi \rangle$$  \hspace{1cm} (4.18)

for all $\psi$. The subscript $t$ on the electric and magnetic fields indicates the transverse part.

In our case, the adjoint waveguide is identical with the original waveguide since we are dealing with non-dissipative media. The operators $\Gamma_{\approx z}$, $L_{\approx z}$ and $M_{\approx z}$ are, therefore, all self-adjoint (see Ref. 25).

We are now ready to proceed. Consider two sets of field vectors $(u_1, v_1)$ and $(u_2, v_2)$ satisfying Eq. (4.16). We can write

$$v_2 = \lim_{\zeta \to 0} \left( u_1 - \frac{L_{\approx z} - \beta_1 \Gamma_{\approx z} - i\zeta \Gamma_{\approx z}}{M_{\approx z}} v_1 \right)^{-1} M_{\approx z} v_1$$  \hspace{1cm} (4.19)

$$v_2^\dagger = \lim_{\zeta \to 0} \left( u_2^\dagger - \frac{L_{\approx z} - \beta_2 \Gamma_{\approx z} + i\zeta \Gamma_{\approx z}}{M_{\approx z}} v_2^\dagger \right)^{-1} M_{\approx z} v_2^\dagger.$$  \hspace{1cm} (4.20)

Our aim is to relate $(v_1, \Gamma_{\approx z} v_2^\dagger)$ to $(u_2, \Gamma_{\approx z} u_2^\dagger)$ and hence to relate the normalization constant of the waveguide modes to that of the incident free-space fields. So from Eqs. (4.19 and 20) we write

$$\langle v_1, \Gamma_{\approx z} v_2^\dagger \rangle = \langle u_1, \Gamma_{\approx z} u_2^\dagger \rangle - \langle L_{\approx z}^{-1} M_{\approx z} v_1, \Gamma_{\approx z} u_2^\dagger \rangle$$

$$- \langle u_1, \Gamma_{\approx z} L_{\approx z}^{-1} M_{\approx z} v_2^\dagger \rangle + \langle L_{\approx z}^{-1} M_{\approx z} v_1, \Gamma_{\approx z} L_{\approx z}^{-1} M_{\approx z} v_2^\dagger \rangle$$  \hspace{1cm} (4.21)

where

$$L_{\approx z} = L_{\approx z} - \beta_j \Gamma_{\approx z} - i\zeta \Gamma_{\approx z}, \quad j = 1, 2$$  \hspace{1cm} (4.22)

and the limit $\zeta \to 0$ is understood. Now
(L^{-1}_{\lambda 1} M v_1, L^{\dagger -1}_{\lambda 2} M v_2) = (M v_1, L^{\dagger -1}_{\lambda 2} \Gamma u_2) \nonumber \\
= (\beta_2 - \beta_1 + i\zeta)^{-1} (M v_1, u_2^{\dagger}), \quad (4.23) 

(u_1, L^{\dagger -1}_{\lambda 2} M v_2) = (L^{\dagger -1}_{\lambda 2} u_1, M v_2) \nonumber \\
= (\beta_1 - \beta_2 - i\zeta)^{-1} (u_1, M v_2), \quad (4.24) 

(M v_1, u_2^{\dagger}) = (M v_1, v_2^{\dagger}) + (M v_1, L^{\dagger -1}_{\lambda 2} M v_2) \nonumber \\
= (L^{\dagger -1}_{\lambda 2} M v_1, M v_2) \quad \text{and} \quad (u_1, M v_2) = (v_1, M v_2) + (L^{\dagger -1}_{\lambda 1} M v_1, M v_2^{\dagger}) \quad (4.25) 

Combining Eqs. (4.23 - 26) we find 

(L^{-1}_{\lambda 1} M v_1, L^{\dagger -1}_{\lambda 2} M v_2) + (u_1, L^{\dagger -1}_{\lambda 2} M v_2) \nonumber \\
= (\beta_2 - \beta_1 + i\zeta)^{-1} (M v_1, (L^{\dagger -1}_{\lambda 2} - L^{\dagger -1}_{\lambda 1}) M v_2) \nonumber \\
= \frac{\beta_2 - \beta_1 + 2i\zeta}{\beta_2 - \beta_1 + i\zeta} (L^{-1}_{\lambda 1} M v_1, L^{\dagger -1}_{\lambda 2} M v_2) \quad (4.27) 

\text{since} \quad L^{\dagger -1}_{\lambda 2} - L^{\dagger -1}_{\lambda 1} = (\beta_2 - \beta_1 + 2i\zeta) L^{\dagger -1}_{\lambda 1} L^{\dagger -1}_{\lambda 2}. \quad (4.28) 

Substituting Eq. (4.27) in Eq. (4.21) we have then 

(v_1, L^{\dagger -1}_{\lambda 2} v_2) = (u_1, L^{\dagger -1}_{\lambda 2} v_2) - \frac{i\zeta}{\beta_2 - \beta_1 + i\zeta} (L^{-1}_{\lambda 1} M v_1, L^{\dagger -1}_{\lambda 2} M v_2^{\dagger}) \quad (4.29) 

Finally, taking the limit \( \zeta \to 0 \) we find that 

(v_1, L^{\dagger -1}_{\lambda 2} v_2) = (u_1, L^{\dagger -1}_{\lambda 2} v_2) \quad (4.30) 

While it is still thinly disguised, this is the result we want because, if we take \( v \) to stand for either \( u \) or \( v \), then
\[ \langle \mathbf{w}_1 \mid \mathbf{w}_2 \rangle = \int_{S_0} \int \left( \mathbf{E}_{t_2}^\dagger \mathbf{H}_{t_1}^* + \mathbf{E}_{t_1}^* \mathbf{H}_{t_2} \right) \cdot \mathbf{\hat{z}} \, da \]

\[ = 2 \int_{S_0} \int \mathbf{E}_{t_2}^\dagger \mathbf{H}_{t_1}^* \cdot \mathbf{\hat{z}} \, da \tag{4.31} \]

since our waveguide is self-adjoint and has reflection symmetry in the z-direction. Therefore, if \( \mathbf{E}, \mathbf{H} \) are the incident fields (free space) and \( \mathbf{\tilde{E}}, \mathbf{\tilde{H}} \) the modal fields we have from Eqs. (4.30 and 31) that

\[ \int_{S_0} \int \mathbf{\tilde{E}}(\lambda, l) \times \mathbf{\tilde{H}}^*(\lambda', l') \cdot \mathbf{\hat{z}} \, da = \int_{S_0} \int \mathbf{\tilde{E}}(\lambda, l) \times \mathbf{\tilde{H}}^*(\lambda', l') \cdot \mathbf{\hat{z}} \, da \tag{4.32} \]

This equation immediately implies that the orthogonality conditions which apply to the fields resulting from scattering from any, open, dissipationless waveguide are identical with those applying to the incident fields. Now for the ITM modes the incident fields have \( (r, \phi, z) \) components:

\[ E^M = \psi^{-\frac{i}{2}} M e^{-i l} \left\{ -\frac{i \beta}{k} J_{\frac{1}{2}} (\lambda R), -\frac{i \beta}{R \lambda k} J_{\frac{1}{2}} (\lambda R), \frac{\lambda}{R \lambda k} J_{\frac{1}{2}} (\lambda R) \right\} \tag{4.33} \]

\[ H^M = \psi^{-\frac{i}{2}} M e^{-i l} \left[ \left\{ 2 \text{Re} \left( \frac{\mu_2}{\mu_1} \right) J_{\frac{1}{2}} (\lambda R), -i \left( \frac{2}{\mu_2} \right) J_{\frac{1}{2}} (\lambda R), 0 \right\} \right] \tag{4.34} \]

So

\[ \int_{S_0} \int E^M(\lambda, l) \times H^M(\lambda', l') \cdot \mathbf{\hat{z}} \, da = \frac{2 \pi \beta^2 \delta_{\lambda, \lambda'}}{\psi_M \psi_{M'}} \left( \frac{E_2}{\mu_2} \right)^{1/2} \frac{E_2}{\lambda'^2} \frac{\beta}{k} \times \mathbf{I}_{M M}, \tag{4.35} \]

where the \( \phi \) integration has been performed and

\[ I_{M M} = \int_0 \left( J_{\frac{1}{2}}^*(\lambda R) J_{\frac{1}{2}}^*(\lambda'R) + \frac{E_2}{\lambda R^2} \frac{J_{\frac{1}{2}} (\lambda R)}{\lambda'} \right) \, dR \]

\[ \sim \frac{1}{\lambda} \delta(\lambda - \lambda'), \tag{4.36} \]

using Eqs. (A.3 and 5). Thus, combining Eqs. (4.32, 35 and 36)
provided

\[ \psi_M = 2\pi \rho^2 \left( \frac{\epsilon_2}{\mu_2} \right)^{1/2} \frac{\beta}{\lambda k_2} \]  

(4.38)

and, in an analogous way,

\[ \int_{S_{\infty}} \int_{S_{\infty}} \mathcal{E}(\lambda, \lambda) \times \mathcal{H}^*(\lambda', \lambda') \times \mathcal{E} \lambda \lambda \delta \lambda \lambda' \]  

(4.39)

where

\[ \psi_E = 2\pi \rho^2 \left( \frac{\mu_2}{\epsilon_2} \right)^{1/2} \frac{\beta}{\lambda k_2} \]  

(4.40)

If we turn now to the cross-product between an ITM and an ITE mode we see that

\[ \int_{S_{\infty}} \int_{S_{\infty}} \mathcal{E}(\lambda, \lambda) \times \mathcal{H}^*(\lambda', \lambda') \times \mathcal{E} \lambda \lambda \delta \lambda \lambda' \]  

(4.41)

where

\[ I_{\text{EM}} = \int_0^\infty \left\{ J_\lambda \left( \lambda R \right) \frac{J_\lambda' \left( \lambda R \right)}{\lambda'} + \frac{J_\lambda \left( \lambda R \right)}{\lambda} J_\lambda' \left( \lambda R \right) \right\} dR \]

\[ = 0 \quad \text{for all } \lambda, \lambda' \]  

(4.42)

using Eqs. (A.4 and 5).

Thus an ITM mode is always orthogonal to an ITE mode regardless of their eigenvalues. This means that in its simplest form, the orthonormality condition (Eq. (4.1)) requires "linear combinations" which are simply the ITM and ITE modes themselves.
Any linear combinations which preserve orthogonality can, however, be used and in the next chapter the following combinations are chosen for convenience:

\[
\begin{align*}
\begin{pmatrix}
E^{(1)} \\
H^{(1)}
\end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix}
E^M + E^E \\
H^M + H^E
\end{pmatrix} \\
\begin{pmatrix}
E^{(2)} \\
H^{(2)}
\end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix}
E^M - E^E \\
H^M - H^E
\end{pmatrix}.
\end{align*}
\]

These combinations clearly satisfy the orthonormality condition.

2.5 RÉSUMÉ

In this chapter, we have derived the radiation modes of the circular dielectric rod by considering them to be the fields resulting from the scattering of an infinite plane wave by the rod. The advantages of this approach are that

(i) the scattering of a plane wave from a dielectric cylinder had already been widely studied,

(ii) the modal fields have a natural decomposition based on the properties of the incident (or free-space) fields of which they are perturbations, and

(iii) mode orthonormality can be deduced from a general theorem relating the incident and scattered fields on an open waveguide.

While the orthonormality conditions can be derived without invoking this picture of an incident and scattered field, the straightforward, direct calculation (Appendix 2A) is, if not more difficult in
principle, certainly much more tedious. Further, it is restricted to a particular system whereas our more general result (Eq. (4.30)) applies to a large class of waveguide cross-sections.

The more general approach has been suggested before but it has not been properly exploited in the case of the circular dielectric cylinder. The proof given in Section 2.4 is intended to fill this gap.
APPENDIX 2A

DIRECT CALCULATION OF NORMALIZATION AND ORTHOGONALITY

We wish to calculate the integral \( \int_S \int G \, da \) where

\[
G = (\psi^*_M \psi^*_M)^{1/2} \left[ \frac{e^-}{r} (\lambda, \ell) \times \frac{h^*_\ell}{r} (\lambda', \ell') \cdot \hat{z} \right] .
\]

(A.1)

Since the integral is infinite and the integrand is well-behaved for \( R < 1 \), the only significant contribution comes from the fields in the region \( R > 1 \). In this region,

\[
G = (\psi^*_M \psi^*_M)^{1/2} \left[ e^M_\ell (\lambda) h^{M*}_\ell (\lambda') - e^M_\ell (\lambda) h^{M*}_\ell (\lambda') \right]
\]

\[
= \frac{1}{2} \left[ e^M_\ell (\lambda) h^{M*}_\ell (\lambda') \right] \left[ J_\ell -1 (\lambda' R) J_{\ell' -1} (\lambda' R) + J_\ell +1 (\lambda' R) J_{\ell' +1} (\lambda' R) \right]
\]

\[
+ \left[ e^M_\ell (\lambda) h^{M*}_\ell (\lambda') \right] \left[ \beta \left( \frac{\varepsilon_2}{\mu_2} \right) a_\ell a_{\ell'} + \beta' \left( \frac{\mu_2}{\varepsilon_2} \right) b_\ell b_{\ell'} \right]
\]

\[
\times \left[ H_{\ell -1} (\lambda' R) H_{\ell' -1} (\lambda R) + H_{\ell +1} (\lambda' R) H_{\ell' +1} (\lambda R) \right]
\]

\[
+ \left[ e^M_\ell (\lambda) h^{M*}_\ell (\lambda') \right] \left[ a_\ell a_{\ell'} \left[ H_{\ell -1} (\lambda' R) J_{\ell' -1} (\lambda R) + H_{\ell +1} (\lambda' R) J_{\ell' +1} (\lambda R) \right] \right]
\]

\[
+ \left[ e^M_\ell (\lambda) h^{M*}_\ell (\lambda') \right] \left[ a_{\ell'} \left[ J_{\ell -1} (\lambda R) H_{\ell' -1} (\lambda' R) + J_{\ell +1} (\lambda R) H_{\ell' +1} (\lambda' R) \right] \right]
\]

\[
- \frac{1}{2} \left[ \frac{\varepsilon_2}{\mu_2} \beta \beta^* \right] \left[ a_\ell a_{\ell'} \left[ H_{\ell -1} (\lambda' R) J_{\ell' -1} (\lambda R) + H_{\ell +1} (\lambda' R) J_{\ell' +1} (\lambda R) \right] \right]
\]

\[
+ \left[ \frac{1}{2} \frac{\varepsilon_2}{\mu_2} \beta \beta^* \right] \left[ J_{\ell -1} (\lambda R) H_{\ell' -1} (\lambda' R) - J_{\ell +1} (\lambda R) H_{\ell' +1} (\lambda' R) \right]
\]
\[ + \left[ \frac{i e^{2}}{\lambda} \right] \left( \beta \beta' a_{\lambda}^{*} b_{\lambda'}^{*} - k_{2}^{2} a_{\lambda}^{*} b_{\lambda'}^{*} \right) \]

\[ \times \left\{ H_{\lambda-1}^{(2)}(\lambda R) H_{\lambda'1}^{(1)}(\lambda' R) - H_{\lambda+1}^{(2)}(\lambda R) H_{\lambda'+1}^{(1)}(\lambda' R) \right\} e^{-i(\lambda-\lambda')\phi}, \quad (A.2) \]

where we have substituted the field components given by Eqs. (3.39 - 48) and have used the Cylinder Function relations:\textsuperscript{26}

\[ z_{\lambda}^{(1)}(x) z_{\lambda'}^{(2)}(y) + \frac{\lambda \lambda'}{xy} z_{\lambda}^{(1)}(x) z_{\lambda'}^{(2)}(y) \]

\[ = \frac{1}{k} \left[ z_{\lambda-1}^{(1)}(x) z_{\lambda-1}^{(2)}(y) + z_{\lambda+1}^{(1)}(x) z_{\lambda+1}^{(2)}(y) \right] \quad (A.3) \]

and

\[ \frac{\lambda'}{y} z_{\lambda}^{(1)}(x) z_{\lambda'}^{(2)}(y) + \frac{\lambda}{x} z_{\lambda}^{(1)}(x) z_{\lambda'}^{(2)}(y) \]

\[ = \frac{1}{k} \left[ z_{\lambda-1}^{(1)}(x) z_{\lambda-1}^{(2)}(y) - z_{\lambda+1}^{(1)}(x) z_{\lambda+1}^{(2)}(y) \right] , \quad (A.4) \]

\[ z_{\lambda}^{(1)} \text{ and } z_{\lambda'}^{(2)} \text{ being any two Cylinder Functions.} \]

If we now perform the \( \phi \) integration over Eq. (A.1) we see that only terms with \( \lambda = \lambda' \) are non-zero. So we are left to consider integrals of the form

\[ \int_{1}^{\infty} z_{n}^{(1)}(\lambda R) z_{n}^{(2)}(\lambda' R) R dR . \]

Because the major contribution to each of these integrals comes from the region \( R >> 1 \), we can use the asymptotic forms of the Cylinder Functions\textsuperscript{26} to obtain

\[ \int_{1}^{\infty} J_{n}(\lambda R) J_{n}(\lambda' R) R dR = \frac{(\lambda \lambda')^{-\frac{1}{2}}}{2\pi} \int_{1}^{\infty} \left\{ e^{i[\lambda R^{2}-(n+\frac{1}{2})\pi/2]} + e^{-i[\lambda R^{2}-(n+\frac{1}{2})\pi/2]} \right\} e^{-i[\lambda R-(n+\frac{1}{2})\pi/2]} dR \]

\[ \sim \frac{1}{\lambda} \delta(\lambda - \lambda') + F_{1} \quad (A.5) \]
\[
\int_{1}^{\infty} J_{n}(\lambda R) H_{n}^{(1)}(\lambda' R) R \, dR = \frac{2}{\pi} (\lambda \lambda')^{-\frac{1}{2}} \int_{1}^{\infty} e^{i[\lambda R-(n+\frac{1}{2})\pi/2]} + e^{-i[\lambda R-(n+\frac{1}{2})\pi/2]} \, dR \\
\approx \frac{1}{\lambda} \delta(\lambda - \lambda') + F_{2} \tag{A.6}
\]

\[
\int_{1}^{\infty} H_{n}^{(2)}(\lambda R) H_{n}^{(1)}(\lambda' R) R \, dR = \frac{2}{\pi} (\lambda \lambda')^{-\frac{1}{2}} \int_{1}^{\infty} e^{-i(\lambda - \lambda') R} \, dR \\
\approx \frac{2}{\pi} \delta(\lambda - \lambda') + F_{3} \tag{A.7}
\]

where \(F_{1}, F_{2}\) and \(F_{3}\) are bounded functions and we have used the result \(27\)

\[
\int_{C>0}^{\infty} e^{i(\lambda - \lambda') R} \, dR = \pi \delta(\lambda - \lambda') + F \tag{A.8}
\]

where \(F\) is a bounded function.

So returning to Eq. (A.2) we now have that

\[
\int_{S_{\infty}} G \, da = \delta_{\lambda, \lambda'} \delta(\lambda - \lambda') \times \frac{2\pi \rho^{2} \beta}{k_{2} \lambda} \left[ \varepsilon_{2} \right]^{\frac{1}{2}} \left[ 1 + \frac{\rho k_{2}}{\lambda} \right] \left[ a_{SM}^{*} + a_{SM} \right] \\
+ \left( \frac{\rho k_{2}}{\lambda} \right)^{2} \left[ a_{SM}^{*} \left| a_{SM} \right|^{2} + \frac{\mu_{2}}{\varepsilon_{2}} \left| b_{SM} \right|^{2} \right] \tag{A.9}
\]

But

\[
\left[ a_{SM}^{*} + a_{SM} + \frac{2\rho k_{2}}{\lambda} \left[ a_{SM}^{*} + a_{SM} \right] \right] \left[ \left| a_{SM} \right|^{2} + \frac{\mu_{2}}{\varepsilon_{2}} \left| b_{SM} \right|^{2} \right] \frac{\rho k_{2}}{\lambda} \\
= - \left( \frac{J}{H} + \frac{2iN_{\mu}}{\pi^{2} M^{2} H^{2}} \right) + \left( \frac{J}{H^{*}} \right) - \frac{2iN_{\mu}^{*}}{\pi^{2} M^{*} H^{*2}} \\
+ 2 \left[ \frac{J}{H} \right]^{2} + \frac{2iN_{\mu}}{\pi^{2} M^{2} H^{2}} \left( \frac{V}{(\rho k_{2})} \right) \left( \frac{\rho k_{2}/\lambda}{2} \right)^{2} \\
= \frac{2i}{\pi^{2} M^{2} H^{2}} \left( N_{\mu} M^{*} - N_{\mu}^{*} M \right) + \frac{8}{(\pi^{2} M^{2} H^{2})^{2}} \left[ N_{\mu} N_{\mu}^{*} + \frac{V}{(\rho k_{2})} \left( \frac{\rho k_{2}}{\lambda} \right)^{2} \right], \tag{A.10}
\]

where
\[ J = J_{\lambda}^{(1)}(\lambda), \quad (A.11) \]
\[ H = H_{\lambda}^{(2)}(\lambda), \quad (A.12) \]
\[ N_{\mu} = N_{2} - \frac{\mu_1}{\mu_2} N_{1}, \quad (A.13) \]

and \( M \) is given by Eq. (3.31). Now since \( N_{1} \) is real, the Wronskian determinant for \( H, H^{*} \) gives
\[ N_{\mu}^{*} - N_{\mu} = \frac{4i}{\pi \lambda^2} |H|^{-2}, \quad (A.14) \]

and so expression (A.10) becomes
\[ \frac{N_{\mu}^{*} - N_{\mu}}{2|\mu|^2} \left\{ \left[ N_{\mu} M^{*} - N_{\mu} M \right] - \left[ N_{\mu}^{*} - N_{\mu} \right] \left[ N_{\mu} M^{*} + \frac{\sqrt{\frac{\mu_1}{\mu_2}}}{k_2} \right]^2 \right\}. \quad (A.15) \]

Using the definition of \( M \) and rearranging further it is then found that this expression is identically zero.

So we have the result that
\[ \int_{S_{\infty}} \int e^{M}(\lambda, \mu) \times h^{M}(\lambda', \mu') \cdot \hat{z} \, da = \psi_{M}^{-1} \left[ 2\pi \rho^2 \frac{\varepsilon_2}{\mu_2} \frac{1}{\lambda k_2} \right] \delta_{\mu}, \delta(\lambda - \lambda') \quad (A.16) \]

A similar calculation for the ITE modes shows that
\[ \int_{S_{\infty}} \int e^{E}(\lambda, \mu) \times h^{E}(\lambda', \mu') \cdot \hat{z} \, da = \psi_{E}^{-1} \left[ 2\pi \rho^2 \frac{\mu_2}{\varepsilon_2} \frac{1}{\lambda k_2} \right] \delta_{\mu}, \delta(\lambda - \lambda') \quad (A.17) \]

So we verify that the normalization conditions (4.37-40) are obeyed by the ITM and ITE modes.

ORTHOGONALITY

We proceed now to show that an ITM(\( \lambda, \mu \)) mode is orthogonal to an ITE(\( \lambda', \mu' \)) mode even if \( \lambda = \lambda' \) and \( \mu = \mu' \). This requires a calculation
very similar to that used to find $\psi_M$, so we will omit most of the
detail.

We wish to calculate the integral

$$I = \int_{S_\infty} \int \mathcal{E}(\lambda, \lambda') \times H^*(\lambda', \lambda') \cdot \hat{\mathcal{E}} \, d\lambda.'$$  \hspace{1cm} (A.18)

On substituting the field components for $R > 1$ given by Eqs. (3.39 - 64)
and using the integrals Eqs. (A.5 - 7) it is found that

$$I = C \delta_{\lambda \lambda'} \delta(\lambda - \lambda'),$$  \hspace{1cm} (A.19)

where

$$C = \left[ \frac{\mathcal{E}_1}{k_1} \right]^2 \frac{\partial k_1}{\lambda} \left[ b^* - b \right] \left[ a^* - a \right].$$  \hspace{1cm} (A.20)

Inserting the explicit forms for $a^* \mathcal{E}, b^* \mathcal{E}, a^* \mathcal{E}$ from Eqs. (3.63, 64,
47, 48) gives

$$C = A \left\{ \frac{1}{M^* H^*} - \frac{1}{M H^2} \right\} - 2 \left\{ \frac{1}{M^* H^*} \left( \frac{J}{H} + \frac{2iN_\mu}{\pi \lambda^2 M^* H^*} \right) - \frac{1}{M H^2} \left( \frac{J}{H^*} - \frac{2iN_\mu^*}{\pi \lambda^2 M^* H^*} \right) \right\},$$  \hspace{1cm} (A.21)

where

$$A = - \frac{2k_1}{\pi} \frac{\delta}{k_2} \left( \frac{V}{U \lambda} \right)^2,$$  \hspace{1cm} (A.22)

$$N_\varepsilon = N_2 - \frac{\varepsilon_1}{\varepsilon_2} N_1$$  \hspace{1cm} (A.23)

and $N_1, N_2, M, J, H, N_\mu$ are as defined in Eqs. (3.29 - 31, A.11 - 13).

On rearranging this expression and using the Wronskian
determinant between $H, H^*$ we find that

$$C = \frac{A}{|M H|^2} \left[ (M^* - M) - (N_\mu^* - N_\mu)(N_\mu^* + N) \right].$$  \hspace{1cm} (A.24)

But since $M^* - M = N_\mu^* N - N_\mu N$ and N is real, it follows that $C = 0$,
i.e.
\[
\int_{S_\infty} \int \mathcal{E}^E(\lambda, \kappa) \times \mathcal{M}^*(\lambda', \kappa') \cdot \hat{\mathcal{H}} \, da = 0 \tag{A.25}
\]

for all \( \lambda, \lambda', \kappa, \kappa' \).
APPENDIX 2B

ALTERNATIVE FORM FOR TRANSVERSE COMPONENTS
OF RADIATION MODES

**ITM MODES**

In the core \((R<1)\),

\[
\psi_{M e f} = \frac{p}{2U} \left\{ - i \chi_M J_{\ell-1}^{(UR)} + i \zeta_M J_{\ell+1}^{(UR)} \right\} \quad (B.1)
\]

\[
\psi_{M e f} = \frac{p}{2U} \left\{ - \chi_M J_{\ell-1}^{(UR)} - \zeta_M J_{\ell+1}^{(UR)} \right\} \quad (B.2)
\]

\[
\psi_{M e f} = \frac{p}{2U} \left( \gamma * J_{\ell-1}^{(UR)} + \sigma * J_{\ell+1}^{(UR)} \right) \quad (B.3)
\]

\[
\psi_{M e f} = \frac{p}{2U} \left\{ - i \gamma * J_{\ell-1}^{(UR)} + i \sigma * J_{\ell+1}^{(UR)} \right\} \quad (B.4)
\]

and, in the cladding \((R>1)\),

\[
\psi_{M e f} = - \frac{i \beta}{2k_2} \left\{ J_{\ell-1}^{(\lambda R)} - J_{\ell+1}^{(\lambda R)} \right\}
\]

\[+ \frac{p}{2\lambda} \left\{ - i \chi_{SM} H_{\ell-1}^{(2)}(\lambda R) + i \zeta_{SM} H_{\ell+1}^{(2)}(\lambda R) \right\} \quad (B.5)
\]

\[
\psi_{M e f} = - \frac{\beta}{2k_2} \left\{ J_{\ell-1}^{(\lambda R)} + J_{\ell+1}^{(\lambda R)} \right\}
\]

\[+ \frac{p}{2\lambda} \left\{ - \chi_{SM} H_{\ell-1}^{(2)}(\lambda R) - \zeta_{SM} H_{\ell+1}^{(2)}(\lambda R) \right\} \quad (B.6)
\]

\[
\psi_{M e f} = \frac{1}{2} \left( \frac{\epsilon_2}{\mu_2} \right)^{\frac{1}{2}} \left\{ J_{\ell-1}^{(\lambda R)} + J_{\ell+1}^{(\lambda R)} \right\}
\]

\[+ \frac{p}{2\lambda} \left( \gamma * H_{\ell-1}^{(2)}(\lambda R) + \sigma * H_{\ell+1}^{(2)}(\lambda R) \right) \quad (B.7)
\]
\[
\psi_{M}^{1/2} h_{\phi}^{M} = -\frac{i}{2} \left( \frac{\epsilon_2}{\mu_2} \right)^{1/2} \left[ J_{l-1}^{(\lambda R)} - J_{l+1}^{(\lambda R)} \right] \\
+ \frac{\rho}{2\lambda} \left[ -i Y_{SM}^{*} H_{l-1}^{(2)}(\lambda R) + i \sigma_{SM}^{*} H_{l+1}^{(2)}(\lambda R) \right]. \tag{B.8}
\]

ITE MODES

In the core, all field components are obtained by simply changing \( M \) to \( E \) in Eqs. (B.1-4). In the cladding,

\[
\begin{align*}
\psi_{E}^{1/2} \begin{pmatrix} e_{r} \\ e_{\phi} \end{pmatrix} & = \psi_{M}^{1/2} \begin{pmatrix} e_{r} \\ e_{\phi} \end{pmatrix} \left[ \frac{B}{k_2} + i \left( \frac{\mu_2}{\epsilon_2} \right)^{1/2}, \pm J_{l+1}, M \rightarrow E \right] \tag{B.9} \\
\psi_{E}^{1/2} \begin{pmatrix} h_{r} \\ h_{\phi} \end{pmatrix} & = \psi_{M}^{1/2} \begin{pmatrix} h_{r} \\ h_{\phi} \end{pmatrix} \left[ \left( \frac{\epsilon_2}{\mu_2} \right)^{1/2} \rightarrow -i \frac{B}{k_2}, \pm J_{l+1}, M \rightarrow E \right]. \tag{B.10}
\end{align*}
\]

The quantities \( \chi, \xi, \eta \) and \( \sigma \) are defined by Eqs. (3.B.1-4) and (3.2.18-25).
REFERENCES


3.1 INTRODUCTION

The modes found in Chapter 2 can now be used to study energy propagation within the dielectric waveguide. We can expand an arbitrary field propagating along the waveguide in the form

\[
\begin{bmatrix}
E \\
H
\end{bmatrix} = \sum_p \sum a_p \begin{bmatrix}
E_p \\
H_p
\end{bmatrix} e^{i(\omega t - \beta_p z)} + \sum_{i=1}^{2} \sum_{\ell=-\infty}^{\infty} \int_{0}^{\infty} A_{\ell}^{(i)}(\lambda) \begin{bmatrix}
E_{\ell}^{(i)}(\lambda, \ell) \\
H_{\ell}^{(i)}(\lambda, \ell)
\end{bmatrix} e^{i(\omega t - \beta \ell z)} d\lambda, \quad (1.1)
\]

where the first term on the right-hand side is a sum over bound modes and the second a sum over the unbound mode combinations Eqs. (2.4.43-44). The coefficients \(a_p\) and \(A_{\ell}^{(i)}(\lambda)\) depend on the nature of the illuminating field and are found using the orthonormality conditions (Eqs. 2.2.26, 2.4.37, 2.4.39) and the orthogonality of bound and radiation fields either at the aperture of the fibre or at some discontinuity.

As before we assume that the axis of the waveguide coincides with the z-axis of a cylindrical coordinate system and the aperture is in the plane \(z=0\). In general, an electromagnetic field incident at this aperture will produce both a reflected and a transmitted field.
which can be found by expanding the fields on each side of the interface in terms of the appropriate modes and imposing continuity conditions on the transverse components of the fields at \( z = 0 \). Formally,

\[
E_i + \sum_n a_n e_n = \sum_m a_e e_m \quad (1.2)
\]

\[
H_i - \sum_n a_n h_n = \sum_m a_h h_m \quad (1.3)
\]

where \( E_i, H_i \) are the transverse incident fields, \( e_n, h_n \) are the transverse fields of the bound and unbound modes of the fibre and \( e_n', h_n' \) the transverse fields of modes appropriate to the region \( z < 0 \) (either free space or some feed mechanism). The summation symbols implicitly include integration over continuum modes.

If we now cross-multiply Eq. (1.2) by \( h_m^* \) and Eq. (1.3) by \( e_m^* \) and use the orthonormality conditions Eqs. (2.2.26, 2.4.37 and 2.4.39), we find

\[
\sum_n a_n' f_n + A_m = a_m \quad (1.4)
\]

\[
- \sum_n a_n' f_n + B_m = a_m \quad (1.5)
\]

where

\[
f_{nm} = \int_{S_\infty} \int_{S_\infty} e_m^* h_n^* Z \, da \quad (1.6)
\]

\[
f_n' = \int_{S_\infty} \int_{S_\infty} e_n' h_m^* Z \, da \quad (1.7)
\]

\[
A_m = \int_{S_\infty} \int_{S_\infty} E_i h_m^* Z \, da \quad (1.8)
\]

and

\[
B_m = \int_{S_\infty} \int_{S_\infty} e_m^* H_i^* Z \, da . \quad (1.9)
\]
Subtracting Eqs. (1.5) and (1.4) gives a set of integral equations for the modal coefficients of the reflected modes (and hence of the transmitted modes),

\[ \sum_{n} a_{nm}(f_{nm} + f_{nm}') = B_{m} - A_{m}, \quad m = 1, 2, \ldots \]  

These equations are most unlikely to be solvable without a great deal of computational effort. Heyke¹ and Cardama and Kornhauser² circumvent this difficulty by placing the dielectric waveguide inside a large, concentric, perfectly conducting waveguide, thus eliminating the continuous spectrum and transforming Eqs. (1.10) into a system of linear equations. Even so, the sum on the left-hand side is an infinite one and numerical solutions must be obtained by truncating the sum and continuing to add further terms until a stable solution is reached. (Cardama and Kornhauser² have also used this method to treat the "inverse" problem of radiation into a uniform medium from the termination of a fibre — a problem with a long history in the literature of dielectric-rod antennas.³)

In cases where the discontinuity in refractive index at the waveguide boundary is so large that there will be a significant amount of reflection and diffraction this full modal treatment is clearly necessary; but in a very large number of situations of practical interest, the change in refractive index is actually very small (and can be made negligibly so using index matching materials).

The earlier papers of Snyder⁴,⁵ and Marcuse⁶ (and more recently Clarricoats and Chan⁷ for the finitely-clad fibre) take advantage of this fact in order to use the Born approximation⁸ to determine the modal coefficients of the bound transmitted field. In
other words, they ignore the reflected field and assume that the fields in the fibre aperture are identical with those of the incident wave.

The effect of this approximation is to replace Eqs. (1.4 - 5) by the equations

\[ a_m = A_m \quad (1.11) \]
\[ a_m = B_m \quad (1.12) \]

These equations are mutually inconsistent because we have demanded continuity of both the tangential electric and magnetic fields, neither of which is exact.\(^9\)

Snyder\(^5\) argues that this inconsistency is not significant when the dielectric difference, \(\delta\), is small provided the angle of incidence of the illuminating field is not too large. Under such circumstances, he ignores Eq. (1.12) and uses Eq. (1.11) to study the excitation of the bound modes of the dielectric cylinder by a truncated plane wave. Essentially the same argument is also used by Kapany, Burke and Sawatari in their work.\(^10\)

On the other hand, Marcuse\(^6\) retains boths Eqs. (1.11) and (1.12) and takes the geometric mean of these to find the excitation coefficient for the HE\(_{11}\) mode under excitation by a Gaussian beam.

That both of these approximations are justified is confirmed by Cardama and Kornhauser\(^2\) who show that for excitation by a normally incident plane wave, Snyder's results are in error by 1.25% while for the case of excitation by a Gaussian beam, Marcuse's result is within 0.2% of the full modal solution when the beam is incident from a medium matched to the cladding and 4% when incident from a vacuum. Moreover, the results of Heyke\(^1\) show that a very accurate estimate of the
proportion of incident power which is reflected at the aperture can be found by simply considering plane wave reflection from a dielectric half-space of refractive index equal to that of the core (under the assumption of relatively small changes in index with which we are concerned). In this way the restriction of small angles of incidence can be removed by using Fresnel's laws to correct for reflections as in Ref. 11.

Since we are explicitly interested in the continuous spectrum, and very little information is lost in using the Born approximation, in this chapter we shall use Eq. (1.11) to determine the excitation of unbound modes under illumination by (i) a truncated plane wave (Section 3.2) and (ii) a quasi-monochromatic incoherent source (Section 3.3). From these results the excitation by a more general source can also be deduced as we shall suggest briefly in Section 3.4.

3.2 PLANE WAVE EXCITATION

3.2.1 Amplitude Coefficients

Consider the system shown in Fig. 3.1 in which a plane wave with its electric field in the plane of incidence illuminates a disc of radius d concentric with the dielectric cylinder. The plane of incidence is assumed to coincide with the plane \( \phi = 0 \) and the incident wave vector, \( \mathbf{k}_i \), makes an angle \( \theta \) with the z-axis.

The incident electric field is, therefore,

\[
E_i = E_i \hat{\mathbf{e}}_z ,
\]

where

\[
E_i = E_0 \exp\left[i(k_z r \sin \theta \cos \phi + \omega t - \beta_z z)\right] , \quad r \leq d
\]

\[
= 0 , \quad r > d .
\]
$E_0$ is an arbitrary amplitude constant, $k_2 \sin \theta \cos \phi$ is the transverse component of the wave vector in the outer medium and $\beta' = k_2 \cos \theta$ is the longitudinal component. $\hat{\mathbf{z}}$ is a unit vector in the direction of the

![Diagram](image_url)

Fig. 3.1: Semi-infinite dielectric rod of radius $\rho$, illuminated by a truncated, obliquely incident plane wave over a disc of radius $d$.

The magnetic field vector, $\mathbf{H}$, is in the direction of the $y$ axis.

incident electric field — for this polarization, $\cos \theta \hat{x} + \sin \theta \hat{z}$ (where $\hat{x}$ and $\hat{z}$ are unit vectors along the $x$ and $z$ axes).

We can now expand the transverse field in Eq. (2.1) at $z = 0$ in the form given by Eq. (1.1). The coefficients, $a_p$, for the bound modes in the expansion are given in an approximate form in Ref. 5 and fully in Appendix 3A. To find the excitation coefficients for the unbound modes, we cross-multiply by $\mathbf{h}^{(i)*}$ in Eq. (1.1) and use the orthonormality relations to find
\[ A_\lambda^{(i)}(\lambda) = \int_{-\infty}^{\infty} \int_0^{2\pi} E_x x h_\lambda^{(i)}(\lambda, \ell) \cdot \hat{\ell} \, d\alpha \]  

\[ = \int_0^{2\pi} \int_0^d E_x h_\lambda^{(i)}(\lambda, \ell) \cos \theta \, r \, dr \, d\phi, \]  

(2.3) 

where the subscript \( y \) on the modal magnetic field indicates the \( y \) component. Since, for an arbitrary vector \( \mathbf{v} \), 

\[ \mathbf{v} = v_x \mathbf{e}_x - v_y \mathbf{e}_y + v_\phi \mathbf{e}_\phi, \]  

(2.4)

Eq. (2.4) can be rewritten as 

\[ A_\lambda^{(i)}(\lambda) = E_0 \cos \theta \int_0^{2\pi} e^{ik_2 r \sin \theta} \cos \phi \left[ h_\lambda^{(i)} \sin \phi \right. \right. 

\left. + h_\phi^{(i)} \cos \phi \right] d\phi \, r \, dr, \]  

(2.5)

where we have substituted Eq. (2.2) but omitted the factor \( e^{i\omega t} \). Now making the substitutions

\[ \Delta = \rho k_2 \sin \theta \]  

(2.6) 

\[ R = \frac{r}{\rho} \]  

(2.7) 

\[ D = \frac{d}{\rho}, \]  

(2.8) 

and using the addition theorem Eq. (2.3.4), we obtain

\[ A_\lambda^{(i)}(\lambda) = \rho^2 E_0 \cos \theta \int_0^{2\pi} \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} i^n J_n(\Delta R) e^{i(\ell-n)\phi} \]  

\[ \times \left[ h_\lambda^{(i)} \sin \phi + h_\phi^{(i)} \cos \phi \right] d\phi \, R \, dR, \]  

(2.9)

where the \( \phi \)-dependent part of the modal fields has been extracted. Thus since
Before proceeding further, we must introduce some new notation to simplify future algebra:

\[ I_M(D) = \int_0^D \left[ P_- J_{\ell-1}(\Delta R) \left( h_x M^* - i h_\phi M^* \right) + P_+ J_{\ell+1}(\Delta R) \left( h_x M^* + i h_\phi M^* \right) \right] R dR \quad (2.15) \]

where \( P_- \) and \( P_+ \) are polarization factors, both equal to \( \cos \theta \) in the present case. If \( D > 1 \), the integral must be split into two parts, one over the core region and another over the cladding. So we write

\[ I_M(D) = I_M^<(D) \quad D \leq 1 \]
\[ = I_M^< (1) + I_M^> (D) , \quad D > 1 . \quad (2.16) \]

Similarly,

\[ I_E = I_M (M+E) . \quad (2.17) \]

We also define

\[ \gamma_M = k_1 \left( \frac{\epsilon_1}{\mu_1} \right)^{1/2} a_\ell M^* + i \beta_d M^* \quad (2.18) \]
\[ \sigma_M = \gamma_M (i + i) \quad (2.19) \]
\[ \gamma_E = i \beta_a E^* - k_1 \left( \frac{\epsilon_1}{\mu_1} \right)^{1/2} d_\ell E^* \quad (2.20) \]
\[ \sigma_E = \gamma_E (i \rightarrow -i) \quad (2.21) \]
\[ \gamma_{SM} = k_2 \left( \frac{\epsilon_2}{\mu_2} \right)^{1/2} a_\lambda^{SM*} + i \beta b_\lambda^{SM*} \quad (2.22) \]
\[ \sigma_{SM} = \gamma_{SM} (i \rightarrow -i) \quad (2.23) \]
\[ \gamma_{SE} = i \beta a^{SE*} - k_2 \left( \frac{\epsilon_2}{\mu_2} \right)^{1/2} b_\lambda^{SE*} \quad (2.24) \]
\[ \sigma_{SE} = \gamma_{SE} (i \rightarrow -i) \quad (2.25) \]

and finally

\[ I(Z_a, Z_b, n; x, y) = \int_x^y Z_n (aR) \tilde{Z}_n (bR) R \, dR \]
\[ = \frac{R}{a^2 - b^2} \left[ a \tilde{Z}_n (bR) Z_{n+1} (aR) - b Z_n (aR) \tilde{Z}_{n+1} (bR) \right] \bigg|_R^Y , \quad a \neq b \]
\[ = \frac{R^2}{4} \left[ 2 Z_n (aR) \tilde{Z}_n (aR) - Z_{n-1} (aR) \tilde{Z}_{n+1} (aR) \right] \bigg|_R^Y , \quad a = b , \quad (2.26) \]

where \( Z_n, \tilde{Z}_n \) are any two cylinder functions and the integral has been evaluated using Lommel's formulae.\(^{13}\)

Returning then to Eq. (2.14) we have

\[ A^{(1)}_\lambda (\lambda) = E_0 \, \pi^2 m^2 \lambda^2 \times \frac{1}{\sqrt{2}} \left[ I_M + I_E \right] \quad (2.27) \]

and

\[ A^{(2)}_\lambda (\lambda) = E_0 \, \pi^2 m^2 \lambda^2 \times \frac{1}{\sqrt{2}} \left[ I_M - I_E \right] , \quad (2.28) \]

where \( I_M \) and \( I_E \) are given by Eqs. (2.16) and (2.17) and...
Equations (2.27–32) completely determine the excitation of the unbound modes under the Born approximation for the case where the electric field vector of the incident plane wave lies in the plane of
incidence. For the alternative case where the electric field is perpendicular to this plane (Fig. 3.2), the unit vector \( \hat{z} \) is \( \hat{y} \), the unit vector along the y axis. The effect of substituting this polarization vector for the original one in Eq. (2.3) can be found by changing the factors \( p_- \) and \( p_+ \) in Eqs. (2.29-32) from \( \cos \theta \) to \( i \) and \( -i \) respectively.

![Fig. 3.2: As for Fig. 3.1 but now the electric field vector is in the direction of the negative y axis.](image)

The analysis leading to these results is identical with that presented above.

3.2.2 POWER PROPAGATION

To calculate the power propagating down the waveguide we must now find the time-averaged Poynting vector, \( \mathcal{S} \), given by

\[
\mathcal{S} = \frac{1}{2} R(\mathbf{E} \times \mathbf{H}^*) ,
\]

where \( R \) indicates the real part and (if we include only propagating
modes — those with $\lambda \leq \rho k_2$ — $E \times H^*$ is found from the equation

$$E \times H^* = \Sigma \Sigma a_{p,q}^* e^{ipq} e_{p,q} H^*$$

$$+ \Sigma \Sigma \rho k_2 d\lambda \int_0^{\rho k_2} d\lambda' A_{\ell}^{(i)}(\lambda) A_{\ell'}^{(i')}(\lambda') \times e^{i(\beta - \beta' - \beta)z} \mathcal{E}^{(i)}(\lambda, \ell) \times \mathbf{h}^{(i')*(\lambda', \ell')}$$

$$+ \Sigma \Sigma \rho k_2 A_{\ell}^{(i)}(\lambda) a_{p}^* e\mathcal{E}^{(i)}(\lambda, \ell) \times \mathcal{h}^{(i)(\lambda, \ell)} + \Sigma \Sigma \rho k_2 A_{\ell}^{(i)}(\lambda) a_{\ell'}^* e\mathcal{E}^{(i)}(\lambda, \ell) \times \mathcal{h}^{(i')*(\lambda, \ell')}$$

where the unbound mode coefficients, $A_{\ell}^{(i)}(\lambda)$, are given by Eqs. (2.27–28) and the bound mode coefficients, $a_{\ell}$, are listed in Appendix 3A.

To find the total $z$-directed power flow we now integrate the $z$ component of the Poynting vector over the area $S_\infty$. That is

$$P_{tot} = \int_{S_\infty} \int \mathbf{S} \cdot \hat{z} \, da$$

$$= \Sigma |a_p|^2 + \Sigma \Sigma \rho k_2 \lambda = \infty i = 1 \int_0^{\rho k_2} A_{\ell}^{(i)}(\lambda) |^2 \, d\lambda$$

$$= \Sigma |a_p|^2 + \Sigma \Sigma \rho k_2 \lambda = 0 i = 1 \int_0^{\rho k_2} A_{\ell}^{(i)}(\lambda) |^2 \, d\lambda ,$$

where we have used mode orthogonality and the fact that

$$A_{-\ell}^{(1)} = (-1)^\ell A_{\ell}^{(2)} .$$

The prime on the summation over $\ell$ indicates that the $\ell = 0$ term must be halved.
The second term in Eq. (2.36) gives the total power radiated from a waveguide of infinite length and is plotted in Fig. 3.3 as a function of \( \nu \), for the case of on-axis illumination with \( D = 1 \). The evanescent modes make no contribution to the total power radiated.

Fig. 3.3: Total power launched into the radiation field by a truncated, normally incident plane wave of radius equal to the waveguide-core radius.

The quantity of greater interest in a large number of applications is the power propagating within the waveguide, \( P_{\text{in}} \), as a function of axial distance from the source.
\[ P_{in} = \int_{S_F} S^* \frac{\partial}{\partial z} \, da \]

\[ = \pi \sqrt{2} R \sum_{p,q} \sum_{a} a_p^* a_q e^{i(\beta_p - \beta_q)z} e^{i\beta_p z} c_{pq} \]

\[ + \sum_{p} \sum_{i} \sum_{i'} \int_{-\infty}^{\infty} \frac{d\lambda}{\lambda} \int_{0}^{\rho k_2} \, d\rho \lambda \, A_{q}^{*}(i)(\lambda) A_{q'}^{*}(i')^{*}(\lambda') \times e^{i(\beta_{i'} - \beta_{i})z} C_{q}(i,\lambda; i', \lambda') \]

\[ + \sum_{p} \sum_{i} \sum_{i'} \int_{0}^{\rho k_2} \, d\rho \lambda \, a_p^* e^{i(\beta_p - \beta_{i'})z} C_{q}(p; i, \lambda) \]

\[ + A_{q}^{*}(i)(\lambda) a_p^* e^{i(\beta_p - \beta_{i'})z} C_{q}(i, \lambda; p) \]  

\[ (2.37) \]

where \( S_F \) is the waveguide cross-section perpendicular to the axis,

\[ C_{pq} = \int_{0}^{1} e_p^* h_{q}^* \cdot \hat{z} \, dR \]  

\[ C_{q}(i,\lambda; i', \lambda') = \int_{0}^{1} e^*_{i}(\lambda, \lambda) \times h_{i'}^* \cdot \hat{z} \, dR \]  

\[ C_{q}(p; i, \lambda) = \int_{0}^{1} e_{p}^* h_{q}^* \cdot \hat{z} \, dR \]  

\[ C_{q}(i, \lambda; p) = \int_{0}^{1} e^*_{i}(\lambda, \lambda) \times h_{p}^* \cdot \hat{z} \, dR \]  

\[ (2.38)-(2.41) \]

and the \( \phi \) integration has been performed.

Evidently, the fact that we are now integrating over a finite cross-section and cannot, therefore, use mode orthogonality makes the calculation of \( P_{in} \) considerably more difficult than that of \( P_{tot}' \), even though the integrals Eqs. (2.38-41) can be calculated analytically using the Lommel formulae (see Appendix 3B).

If we were interested in the bound-mode power, matters could be simplified a great deal because the bound modes are "approximately
orthogonal" over the cylinder cross-section, i.e.

\[ C_{pq} = \eta_p \delta_{pq} \]  \hspace{1cm} (2.42)

where \( \eta_p \) is the fraction of the power of mode \( p \) within the cylinder.

Thus, to a very good approximation, the bound mode power within the waveguide can be written as

\[ P_{in}^B = \frac{1}{2} \sum_p |a_p|^2 \eta_p. \]  \hspace{1cm} (2.43)

However, because the energy distribution of an unbound mode is concentrated in the core of the waveguide to a far smaller degree than that of a bound mode, the unbound modes cannot be expected to be even approximately orthogonal (either to each other or to the bound modes) over the cylinder cross-section. It is therefore necessary to retain in full the second and third terms in Eq. (2.37).

So the problem of calculating the power propagated within the waveguide is now reduced to a numerical one. The first step is to determine which bound modes can propagate in a particular waveguide — that is, which modes have their cutoff below the \( V \)-value of the waveguide. The first and third terms, \( P_{in}^B \) and \( P_{in}^{\text{cross}} \) can then be evaluated immediately, the latter term requiring a straightforward one-dimensional integration over \( \lambda \).

Unfortunately, the two-dimensional integral over \( \lambda, \lambda' \) involved in the unbound mode contribution, \( P_{in}^{UB} \), presents considerably more difficulty. The source of this difficulty is that there does not seem to exist an acceptably reliable cubature procedure for an oscillatory function of two variables — and the integrand in this case is highly oscillatory.
The most useful procedure would clearly be one which sampled a denser mesh of points near peaks in the integrand than in smooth regions. Such an adaptive routine, based on Simpson's rule was, in fact used. However, as the frequency of oscillations increased (i.e. as \( z \) or \( V \) became large), it did not produce consistently reliable results. For this reason, the results to be presented below were also re-calculated using two other routines — one a two-dimensional Simpson's rule and the other a two-dimensional Romberg rule — and the results were compared to remove a number of spurious discontinuities in the curves for \( P_\infty \) against \( z \). No one routine was accurate over the whole range, each having difficulty converging in a different region.

This is obviously not a practical procedure to adopt generally, particularly since all these difficulties were encountered at values of \( V \leq 10 \), at very small values of \( z \) and at small angles of incidence of the source field. As the angle of incidence increases (so that higher values of \( l \) become significant) and \( V \) and \( z \) increase (so that the integrand oscillates more rapidly) the difficulties and time involved will also increase. There are, of course, more sophisticated cubature rules than those mentioned above, but these generally require special types of integrand. For the more general integrand, relatively little of immediate practical use seems to have been achieved in the century since Maxwell's contribution.

Despite these limitations, results have been obtained for a range of values of \( V \) (\( < 10 \)) appropriate to single-mode fibres and visual photoreceptors (see Chapter 6). The influence of changes in angle of incidence and radius of illumination of the truncated plane wave have also been examined. All results below are for the case \( \mu_1 = \mu_2 \).
Figures 3.4 and 3.5 show $P^\text{UB}$ and $P^\text{cross}$ as functions of the normalized axial distance from the source, $z/p$, for $\delta = 0.05$ and values of $V$ between 1 and 2. The source is assumed to have unit power over the fibre aperture, to be normally incident ($\theta = 0$) and to have radius equal to the radius of the fibre ($D = 1$). Only one bound mode, the HE$_{11}$ mode, is involved in the cross-term as all other modes are cut off at these low values of $V$.

Comparison of these results with those found from Eq. (2.43) for the bound mode contribution, $P^B_{in}$, (Table 3.1) shows that for the lower values of $V$, the unbound mode contribution is initially the larger. However, the energy in the unbound modes is radiated from the fibre very quickly and it is only in the shorter visual photoreceptors that the unbound modes play a significant role.

<table>
<thead>
<tr>
<th>$V$</th>
<th>$P^B_{in}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.029</td>
</tr>
<tr>
<td>1.2</td>
<td>0.111</td>
</tr>
<tr>
<td>1.5</td>
<td>0.285</td>
</tr>
<tr>
<td>2</td>
<td>0.528</td>
</tr>
<tr>
<td>3.5</td>
<td>0.786</td>
</tr>
<tr>
<td>8</td>
<td>0.917</td>
</tr>
</tbody>
</table>

Table 3.1
Contribution of Bound Modes to Power in Fibre

Even at $z = 0$, the contributions from $P^\text{UB}_{un}$, $P^\text{cross}_{in}$ and $P^B_{in}$ do not sum to equal the source power because we have neglected the evanescent modes and their cross-terms with other modes. While these
Fig. 3.4: Propagating unbound mode power within the core as a function of distance from the aperture. The illuminating field is identical with that in Fig. 3.3.

Fig. 3.5: Power within the core due to cross-terms between bound and unbound modes, again using the same source as in Fig. 3.3.
evanescent modes do not contribute to net power propagation across an infinite cross-section, they may contribute in any subdivision of this cross-section. They will also contribute to power absorption in visual photoreceptors and we will discuss this aspect further in Chapter 6.

Figures 3.6 and 3.7 show that under the same illumination as for Figs. 3.4 and 3.5, increasing V causes a marked decrease in the power carried within the fibre by the unbound modes. Figure 3.7 also shows that while the initial power carried by the cross-term is smaller than that due entirely to unbound modes, the former decays more slowly because of the presence of the bound mode component. Again only the HE\textsubscript{11} mode is included in the calculation of $P_\text{in}^{\text{cross}}$ at V=3.5 but the corresponding calculation at V=8 would also have to include the HE\textsubscript{21} and HE\textsubscript{31} modes. Table 3.1 gives the bound mode contributions at V=3.5 and 8.

The effect of moving the incident plane wave off-axis is illustrated in Fig. 3.8 for the case $\delta = 0.01$ and $D=1$ (with polarization as illustrated in Fig. 3.1). The angle $\theta_c$ in the diagram is the complement of the critical angle given by Snell's law, namely $\sin^{-1}(\delta^2)$. The curves illustrate the expected result that the effect of increasing $\theta$ is to increase the proportion of power launched into the unbound modes. They further illustrate that as $V$ increases, the result expected from geometric optics is approached — i.e. for $\theta < \theta_c$, bound modes account for the majority of the power propagated while for $\theta > \theta_c$, the unbound modes dominate.

To carry the ray analogy a little further, it would be expected that as $\theta$ increases, more and more skew rays would be excited by the source field. In the present calculations this is evidenced by
Fig. 3.6: As in Fig. 3.4 but with $V$ increased.

Fig. 3.7: Comparison of $P_{in}^{UB}$ and $P_{in}^{cross}$ when $V = 3.5$. 
Fig. 3.8: Variation of power launched into radiation field with angle of incidence of source field.
Fig. 3.9: Total power launched into the radiation field as a function of \( V \) for various angles of incidence of the source field. The peaks correspond to bound mode cutoff frequencies.

The fact that progressively higher order modes (in \( \ell \)) must be included as \( \theta \) increases.

If the results of Fig. 3.8 are plotted with the variable \( V \) as abscissa rather than \( \theta \), one obtains the interesting curves shown in Fig. 3.9. The peaks in these curves occur at the cutoff frequencies of the various bound modes, matching precisely the dips in the
corresponding bound mode curves given by Pask and Snyder. Moreover each of the matching pairs of curves sums to unity, thus providing a check on the unbound mode calculations.

\[ P_{\text{in}}^{\text{UB}}(D) \]

![Graph](image)

**Fig. 3.10:** Propagating unbound-mode power within the core as a function of axial distance from the aperture for various values of the source radius \( d \) in Fig. 3.1. \( D = d/p \) and, in this example, \( \theta = 0 \) and \( V = 4 \).

The remaining parameter to consider is the radius, \( d \), of the illuminating disc. Figure 3.10 shows that light falling outside the aperture of the fibre can, nevertheless, contribute significantly to the power carried by unbound modes within the cylinder. The fact that \( P_{\text{in}}^{\text{UB}} \) initially increases with \( z \) indicates that this is achieved by "beating" or constructive interference between the fields in the core and cladding resulting in a transfer of power into the fibre core. The source was again an on-axis plane wave with unit power over the fibre aperture (so that the total incident power increases with \( D \)). An application of this effect in the case of bound modes is given in Ref. 23.
3.2.3 Comparison with Diffraction at a Circular Aperture

Since in all the calculations presented above the difference in dielectric constant between the fibre core and cladding is very small, it is of some interest to investigate how the radiation field is altered relative to what would be obtained by simply passing a plane wave through a circular aperture in an otherwise opaque screen.

To do this, we consider a plane wave (wave vector \( \mathbf{k}_2 \)) incident at angle \( \theta \) on a thin, opaque screen containing a circular hole of radius \( \rho \). The polarization vector of the incident wave lies in the plane of incidence, assumed to be the xz plane (Fig. 3.11) and the medium on either side of the screen has dielectric constant \( \varepsilon_2 \) (equal to that of the cladding in Fig. 3.1).

In the vector Kirchhoff approximation, the time-averaged diffracted power per unit solid angle is\(^\text{24}\)

\[
\frac{dP}{d\Omega_1} = P_i \cos \theta \left( \frac{(\rho k_2)^2}{4\pi} \right) \left( \cos^2 \phi + \cos^2 \theta_1 \sin^2 \phi \right) \left| \frac{2J_1(\rho k_2 \xi)}{\rho k_2 \xi} \right|^2,
\]

(2.44)

where \( P_i \) is the total normally incident power (put equal to 1 in the numerical calculations), \( \theta_1 \) and \( \phi \) are respectively the polar and azimuthal angles of \( \mathbf{k}_2 \) the diffracted wave vector as shown in Fig. 3.11 and

\[
\xi = (\sin^2 \theta_1 + \sin^2 \theta - 2 \sin \theta \sin \theta_1 \cos \phi)^{\frac{1}{2}}.
\]

(2.45)

Integrating Eq. (2.44) over the azimuthal angle \( \phi \) gives the distribution of power as a function of polar angle \( \theta_1 \):

\[
dP = P_i \cos \theta \left( \frac{(\rho k_2)^2}{4\pi} \right) \left[ 2\int_0^{2\pi} \left( \cos^2 \phi + \cos^2 \theta_1 \sin^2 \phi \right) \left| \frac{2J_1(\rho k_2 \xi)}{\rho k_2 \xi} \right|^2 d\phi \right] \sin \theta_1 d\theta_1
\]

(2.46)

\[
= \frac{P_i}{\sin \theta_1} (1 + \cos^2 \theta_1) |J_1(\rho k_2 \sin \theta_1)|^2 d\theta_1
\]

(2.47)

In the case of normal incidence.
Now when a plane wave strikes the circular aperture of a dielectric cylinder, the field is divided into two parts — one bound and the other radiated. In order to compare the intensity distribution of the radiated field with that of the field diffracted at a circular aperture (Eq. (2.46) or (2.47)) we must therefore also divide the latter into two corresponding parts. This can be done approximately by invoking Snell's law and assuming that the part of the diffracted power within a cone bounded by the critical angle \( \theta_c \) \( = \cos^{-1}(\frac{n_2}{n_1}) \) corresponds to the bound mode field while power diffracted outside this cone should be compared with the unbound mode power. A polar plot of \( \frac{dP}{(\sin\theta_1 \, d\theta_1)} \) illustrates this separation (Fig. 3.12).

Since we are interested in the range \( \theta_c \leq \theta_1 \leq \frac{\pi}{2} \), in Eqs. (2.46 and 47), we make the change of variable

\[
\cos\theta_2 = \frac{n_1}{n_2} \cos\theta_1 , 
\]

where the range of interest is \( 0 \leq \theta_2 \leq \frac{\pi}{2} \).

Then

\[
dP = \frac{p_1 \cos\theta \, n_2}{n_1} \frac{(pk_2)^2}{4\pi} \sin\theta_2 \left[ \int_0^{2\pi} \left( \cos^2\phi + \left( \frac{n_2}{n_1} \cos\theta_2 \right)^2 \sin^2\phi \right) \right. \\
x \left. \left[ 2J_1(pk_2 \xi') \right]^2 \right] \, d\theta_2 ,
\]

where

\[
\xi' = \left[ 1 - \left( \frac{n_2}{n_1} \cos\theta_2 \right)^2 + \sin^2\theta - 2 \sin\theta \cos\phi \left[ 1 - \left( \frac{n_2}{n_1} \cos\theta_2 \right)^2 \right] \right]^{\frac{1}{2}}
\]

and when \( \theta = 0 \),

\[
dP = \frac{p_1 \, n_2}{n_1} \sin\theta_2 \left[ \frac{n_1^2 + n_2^2 \cos^2\theta_2}{n_1^2 - n_2^2 \cos^2\theta_2} \right] \\
x \left[ J_1 \left( pk_2 \left[ 1 - \left( \frac{n_2}{n_1} \cos\theta_2 \right)^2 \right] \right) \right]^2 \, d\theta_2 .
\]
Fig. 3.11: Diffraction of a plane wave at a circular aperture in an opaque screen. The polarization vector of the incident wave lies in the plane of incidence (the xz plane) and the wave is incident from below with wave vector $k_2$.

Fig. 3.12: Polar diagram of diffraction pattern produced by a plane wave normally incident at a circular aperture ($\theta = 0$ in Fig. 3.11). The sidelobes are magnified by a factor of 10 relative to the central lobe to make them visible.
This intensity distribution is to be compared with

\[
\frac{dP_{UB}}{d\Omega} = \sum_{\ell=0}^{\infty} \sum_{i=1}^{2} |A_{\ell}^{(i)}(\lambda)|^2 d\lambda
\]

\[
= \rho k_2 \sum_{\ell=0}^{\infty} \sum_{i=1}^{2} |A_{\ell}^{(i)}(\rho k_2 \sin \theta_2)|^2 \cos \theta_2 d\theta_2, \quad \theta \neq 0 \quad (2.52)
\]

\[
= \rho k_2 \sum_{i=1}^{2} |A_1^{(i)}(\rho k_2 \sin \theta_2)|^2 \cos \theta_2 d\theta_2, \quad \theta = 0 \quad (2.53)
\]

from Eq. (2.36).

Figures 3.13(a–d) compare the intensity distributions given by Eqs. (2.51 and 53) for \( V = 3.5 \) and \( \delta \) varying between 0.01 and 0.2. As one would expect, the correlation is very good for \( \delta \) very small but gradually worsens. Thus for small \( \delta \) the waveguide only slightly perturbs the free-space fields — the major effect coming from diffraction at the aperture.

The discrepancy between Eqs. (2.51 and 53) at small values of \( \theta_2 \) (as well as the anomalous result in Fig. 3.13a that the intensity is greater than one for small \( \theta_2 \)) is due to the approximation involved in using Snell's law to separate the diffracted field into two mode categories. The identification of rays and modes is only approximate and particularly so at such small values of \( V \).

One can use the intensity distribution given by Eq. (2.51) to approximate \( P_{UB}^{tot} \) by integrating over \( \theta_2 \). Figure 3.14 shows that this approximation coincides very closely with the exact result previously illustrated in Fig. 3.2.
3.2.4 "Slab" Approximation to Total Unbound Mode Power

Another interesting approximation to the total power radiated from a dielectric waveguide can be found by explicitly making use of the fact that, particularly at small values of \( V \), most of the unbound mode power does leave the waveguide very quickly. This suggests that we
consider the approximation \( \lambda >> 1 \) which corresponds to a large transverse component of the unbound-modal wave vector.

![Graph showing comparison of total power in unbound modes with total power found by integrating the diffraction pattern of a circular aperture and by integrating the "slab" approximation of Section 3.2.4.]

Fig. 3.14: Comparison of total power in the unbound modes as shown in Fig. 3.1 (-----) with total power found by integrating the diffraction pattern of a circular aperture (---) and by integrating the "slab" approximation of Section 3.2.4 (· · · · · ·).

Then using the asymptotic expressions \(^8\)

\[
J_1(U) \sim \left( \frac{2}{\pi U} \right)^{1/2} \cos \left( U - \frac{3\pi}{4} \right) \quad (2.54)
\]

and

\[
H_1^{(2)}(\lambda) \sim \left( \frac{2}{\pi \lambda} \right)^{1/2} e^{-i\left( \lambda - \frac{3\pi}{4} \right)} , \quad (2.55)
\]

we find that in the case \( D = 1, \theta = 0 \)

\[
\sum_{i=1}^{2} |A_1^{(i)}(\lambda)|^2 \sim \frac{2k_1\lambda^2 \left[ 1 + \frac{\beta^2}{K_2^2} \right]}{\pi\beta U^2 \left( U^2 + \lambda^2 \cot^2 \left( U - \frac{\pi}{4} \right) \right)} . \quad (2.56)
\]
Apart from the phase shift of $\frac{\pi}{4}$, this expression is quite similar to the corresponding expressions for the amplitude coefficients of radiation modes on a slab waveguide.\textsuperscript{25} The result on integrating this approximation over $\lambda$ is shown in Fig. 3.14 and once again we have very close agreement with the exact result.

3.3 EXCITATION BY A QUASI-MONOCROMATIC INCOHERENT SOURCE\textsuperscript{12}

The plane-wave source used in the preceding section is a perfectly coherent one — that is, the relationship between the phase and polarization of light emitted from any two points in the source is constant. In this section we consider the opposite extreme of excitation by a totally incoherent source placed at the fibre aperture. Specifically, we consider a uniform source of radius $d$ in which all points are statistically independent in the sense that the phase and polarization of light emitted from any two points are totally uncorrelated. We do, however, assume the source to be quasi-monochromatic.\textsuperscript{26,27}

Following Snyder and Pask,\textsuperscript{28} we write the transverse electric field of the uniform source as

$$E_i = E_0 \left[ \mathcal{R} \cos \psi + \mathcal{I} \sin \psi \right] e^{i(\omega t - \xi)},$$

(3.1)

where $E_0$ is an amplitude constant and $\psi$ and $\xi$ are random functions of $x, y$ (or $r, \phi$) and time. Thus $E_i$ has random phase and polarization. The quasi-monochromaticity of the source restricts $\psi$ and $\phi$, however, to be slowly varying functions of time.\textsuperscript{26,27}

As in the plane-wave case, this source field can be expanded in terms of bound and unbound modes (Eq. (1.1)) at $z = 0$. Then using
Eqs. (2.4.37-42) the expansion coefficients for the unbound modes are given by

\[ A_{\lambda}^{(i)}(\lambda) = E_0 \int_S \int_S \left( \hat{\alpha} \cos \psi + \hat{\beta} \sin \psi \right) \times \hat{B}^{(i)*}(\lambda, \xi) e^{-i \xi \cdot \hat{z}} \, d\alpha , \quad (3.2) \]

where \( S \) is the disc of radius \( d \), centred on the waveguide axis, on which the source has non-zero intensity. The corresponding coefficients for the bound modes are found by replacing \( A_{\lambda}^{(i)}(\lambda) \) and \( h^{(i)*}(\lambda, \xi) \) by \( a_p \) and \( h_p^* \).

The instantaneous power flow carried by an electromagnetic field with complex field vectors given by Eq. (1.1) across a surface \( S \) perpendicular to the waveguide axis is

\[ P = \int_S \int_S \left[ E_T \times \tilde{H}_T + E_T^* \times \tilde{H}_T^* + E_T^* \times \tilde{H}_T + E_T \times \tilde{H}_T^* \right] \cdot \hat{z} \, d\alpha , \quad (3.3) \]

where \( E_T, H_T \) are the real, physical fields given by

\[ E_T = \frac{1}{2} (E + E^*) \]

and

\[ H_T = \frac{1}{2} (H + H^*) . \quad (3.4) \]

The time-averaged power flow, \( P \) is then

\[ P = \langle P \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} P \, dt \]

\[ = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left[ E \times H + E \times H^* + E^* \times H + E^* \times H^* \right] \cdot \hat{z} \, d\alpha . \quad (3.5) \]

In the case of the plane wave studied in Section 3.2, the only time-dependence of the field is in the factor \( e^{i \omega t} \), where \( \omega \) is a fixed frequency. The above time-average therefore reduces immediately to

\[ \frac{1}{2} R(E \times H^*) \]

since the terms in \( E \times H \) and \( E^* \times H^* \) involve the factor \( e^{\pm 2i \omega t} \).
which averages to zero over a time-interval large compared with its period.

In the present case, the first and fourth terms in the integrand of Eq. (3.5) will also average to zero provided \( \psi \) and \( \xi \) are slowly varying functions of time compared with the exponential term.

(The polychromatic case requires more careful consideration.\(^\text{27} \)) So with this proviso,

\[
P = \frac{1}{2} \mathcal{R} \int \int \left\{ \sum_{p, q} \left[ a_p a_q^* \right] e^{i(\beta - \beta)z} e_p^* e_q^* \right\} \\
+ \sum_{i, i' = 1}^2 \sum_{l = -\infty}^\infty \int_0^{\rho k_2} d\lambda \int_0^{\rho k_2} d\lambda' \left\{ A_{\ell} (i) (\lambda) A_{\ell} (i') (\lambda') \right\} \\
\times e^{i(\beta' - \beta)z} \frac{\xi (i)}{(\lambda, \ell)} \times \frac{\xi (i')^*}{(\lambda', \ell)}
\]

\[
+ \sum_{p} \sum_{i = 1}^2 \sum_{l = -\infty}^\infty \int_0^{\rho k_2} d\lambda \left\{ \left( a_p \right) A_{\ell} (i)^* (\lambda) e^{i(\beta - \beta)z} e_p^* \frac{\xi (i)}{(\lambda, \ell)} \times \frac{\xi (i')^*}{(\lambda', \ell)} \right\} \\
+ \left\{ A_{\ell} (i) (\lambda) a_p^* e^{i(\beta - \beta)z} e_p^* \frac{\xi (i)}{(\lambda, \ell)} \times \frac{\xi (i')^*}{(\lambda', \ell)} \right\} \right\} \cdot \hat{\xi} \, da ,
\]

where again \( \{ \} \) denotes the time average. Now,

\[
A_{\ell} (i) (\lambda) A_{\ell} (i')^* (\lambda') = E_0^2 \int \int \int \int \left\{ \left( \hat{\xi} \cos \psi + \hat{\xi} \sin \psi \right) \\
\times h (i)^* (\lambda, \ell) \cdot \hat{\xi} \right\} \times \left\{ \left( \hat{\xi} \cos \psi' + \hat{\xi} \sin \psi' \right) \\
\times h (i')^* (\lambda', \ell) \cdot \hat{\xi} \right\} \right\} e^{i(\xi' - \xi)} \, da \, da' ,
\]

Using the fact that unless \( x = x' \), \( y = y' \), the time-average of the random phase term \( e^{i(\xi' - \xi)} \) is zero\(^{27} \) and integrating over the azimuthal angle \( \phi \) we obtain
\[
(A^*_\ell (i)'(\lambda)A^*_{\ell'}(i')(\lambda')) = \lim_{T \to \infty} \frac{\pi \rho^2 E_0^2}{T} \int_{-T}^{T} dt \int_D \mathcal{D} R dR \left\{ h_y^{(i)}(\lambda, \bar{\lambda}) \right\}
\]
\[
x h_y^{(i)}(\lambda', \bar{\lambda}) \cos^2 \psi + h_x^{(i)}(\lambda, \bar{\lambda}) \n\]
\[
x h_x^{(i')}(\lambda', \bar{\lambda}) \sin^2 \psi - \left[ h_y^{(i)}(\lambda, \bar{\lambda}) h_x^{(i')}(\lambda', \bar{\lambda}) \right.
\]
\[
+ h_x^{(i)}(\lambda, \bar{\lambda}) h_y^{(i')}(\lambda', \bar{\lambda}) \left\{ \sin \psi \cos \psi \right\} .
\]
Thus, since as \( T \to \infty \)
\[
\frac{1}{2T} \int_{-T}^{T} \cos^2 \psi \, dt = \frac{1}{2T} \int_{-T}^{T} \sin^2 \psi \, dt = \frac{1}{2}
\]
and
\[
\frac{1}{2T} \int_{-T}^{T} \cos \psi \sin \psi \, dt = 0 ,
\]
\[
A^*_\ell (i) (\lambda)A^*_{\ell'}(i')(\lambda') = \pi \rho^2 E_0^2 \int_D \left[ h_x^{(i)}(\lambda, \bar{\lambda}) h_x^{(i')}(\lambda', \bar{\lambda}) \right.
\]
\[
+ h_y^{(i)}(\lambda, \bar{\lambda}) h_y^{(i')}(\lambda', \bar{\lambda}) \right\} \mathcal{D} R dR . \tag{3.7}
\]
Similarly
\[
\langle a_p a_q^* \rangle = \pi \rho^2 E_0^2 \int_D \left[ h_{px}^* h_{qx} + h_{py}^* h_{qy} \right] \mathcal{D} R dR \tag{3.8}
\]
\[
\langle a_p A^*_{\ell}(\lambda) \rangle = \pi \rho^2 E_0^2 \int_D \left[ h_{px}^* h_{x}^{(i)}(\lambda, \bar{\lambda}) + h_{py}^* h_{y}^{(i)}(\lambda, \bar{\lambda}) \right] \mathcal{D} R dR \tag{3.9}
\]
and
\[
\langle A^*_\ell (i)(\lambda) a_p^* \rangle = \langle a_p A^*_{\ell}(\lambda) \rangle^* . \tag{3.10}
\]
The integrals in Eqs. (3.7-10) can again be performed using the Lommel formulae. The results are given in Appendix 3C.

Finally, substituting Eqs. (3.7-10) in Eq. (3.6) we find
\[ P_{\text{tot}} = \frac{1}{2} \sum_{p} \left\{ \Sigma \left( |a_p|^2 \right) + \sum_{i=1}^{2} \sum_{l=-\infty}^{\infty} \int_{0}^{0k_2} \langle |A_{\chi}(i)(\lambda)|^2 \rangle \, d\lambda \right\} \]  

(3.11)

and

\[ p_{\text{in}} = \pi p^2 R \left\{ \Sigma \left( a_p a_{p}^{*} \right) e^{i(\beta - \beta_{p})z} \right. \left. \Sigma_{p,q} \right\} \left( a_{p} a_{q}^{*} \right) e^{i(\beta - \beta_{q})z} \left. \Sigma_{p} \right\} \left( a_{p} a_{q}^{*} \right) e^{i(\beta - \beta_{p})z} \]  

\[ + \sum_{p} \sum_{i=1}^{2} \sum_{l=-\infty}^{\infty} \int_{0}^{0k_2} \int_{0}^{0k_2} d\lambda \int_{0}^{0k_2} d\lambda' \langle A_{\chi}(i)(\lambda) \, A_{\chi}(i')^{*}(\lambda') \rangle \right. \]  

\[ \times e^{i(\beta' - \beta)z} C_{\chi}(i,\lambda; i', \lambda') \]  

\[ + \sum_{p} \sum_{i=1}^{2} \sum_{l=-\infty}^{\infty} \int_{0}^{0k_2} \int_{0}^{0k_2} d\lambda \left\{ \langle a_{p} A_{\chi}^{*}(i) \rangle e^{i(\beta - \beta_{p})z} \right. \]  

\[ + \langle A_{\chi}(i)(\lambda) a_{p}^{*} \rangle e^{i(\beta - \beta_{p})z} C_{\chi}(i,\lambda; p, \lambda) \left( a_{p} a_{q}^{*} \right) e^{i(\beta - \beta_{p})z} C_{\chi}(i,\lambda; p, \lambda) \} \} \right. \]  

(3.12)

where \( C_{pq}, C_{\chi}(p; i, \lambda), C_{\chi}(i, \lambda; p) \) and \( C_{\chi}(i, \lambda; i', \lambda') \) are given by Eqs. (2.38-41).

Snyder and Pask\(^{28}\) have shown that only a small proportion of the energy leaving the incoherent source is carried by bound modes when \( \delta \) is small. Specifically, they find that for small \( \delta \) the time-averaged power radiated from a source of radius \( \rho \) (i.e. covering the fibre aperture) can be approximated by

\[ P_{s} = \frac{2 \nu^2}{3} \delta \]  

(3.13)

and that the proportion of power launched into the unbound modes is

\[ \frac{P_{\text{tot}}^{\text{UB}}}{P_{s}} \approx 1 - \frac{3}{4} \delta \left( 1 - \frac{1}{2V} \right). \]  

(3.14)

Figure 3.15 shows some of the first ten terms in the sum over \( l \) for \( P_{\text{tot}}^{\text{UB}} \) (normalized by \( P_{s} \)) when \( \delta = 0.05 \). At \( V = 2 \), these sum to 0.95 whereas Eq. (3.14) gives 0.97. Clearly, at larger values of \( V \),
Fig. 3.15: Some of the first ten terms in the sum over $\ell$ for $P_{UB}^{tot}$ (Eq. 3.11).

progressively larger values of $\ell$ would have to be included to adequately estimate $P_{UB}^{tot}$. Also, as $V$ increases, Fig. 3.15 shows that the condition of "all modes equally excited" is being approached.

The fact that so many $\ell$ values do contribute significantly makes the numerical calculations with the incoherent source even more difficult than for the plane-wave since many more integrals must be performed. The reason that this occurs is that an incoherent source can
be considered to be composed of a spectrum of plane waves incident over a range of angles. Therefore, as the discussion concerning Fig. 3.8 would lead us to expect, modes with a range of \( \ell \)-values will be excited. This also means that in the calculation of the cross-term in Eq. (3.12), all bound modes which are not cutoff must be included — unlike the plane wave, the incoherent source is not selective in its excitation.

Figure 3.16 shows that while a large amount of power is launched into the unbound modes, when \( V = 3.7 \) and \( \delta = 0.05 \), this power radiates from the cylinder quite quickly although it does so more slowly than the corresponding plane-wave excited field.
As we shall see later, (Chapter 5), at larger values of $V$ this radiation occurs much more gradually. But for the reasons discussed in Section 3.2, the methods presented in this chapter are impractical at such large values of $V$.

3.4 GENERALIZATION AND CONCLUSIONS

In this chapter, we have used the complete set of spectral modes introduced in Chapter 2 to investigate the propagation of light (and particularly the radiation field) along a dielectric waveguide under illumination by two specific source fields. An "Airy disc" source will also be discussed briefly in Chapter 6 in connection with visual photoreceptors. But it is clear that Eq. (1.1) and the mode orthonormality relations can be used to carry through the same procedure for any source for which an expression for the electric (or magnetic) field can be written down — such as the commonly discussed Gaussian beam, for example.

However, while the above comments are true in principle, this approach has several practical limitations even in the cases where the expansion coefficients in Eq. (1.1) can be found analytically. The approach is certainly adequate (and perhaps necessary) for very small fibres such as those found in visual photoreceptors but for larger $V$ or longer fibres, some alternative to the tedious numerical calculations described in Sections 3.2 and 3.3 is clearly desirable. The following chapters examine such an alternative based on an asymptotic approximation to the integral over the unbound modes.
APPENDIX 3A

PLANE-WAVE EXCITATION COEFFICIENTS FOR BOUND MODES

Using the Born approximation, the coefficients $a_p$ in Eq. (1.1) for bound modes excited by a plane-wave source are found by precisely the same method as was used in Section 3.2.1 to find the coefficients for the unbound modes. That is, using orthogonality, we write

$$ a_p = \int_S \int E_1 \times h_p^* \cdot \mathbf{Z} \, da $$

(A.1)

$$ = E_0 \pi \rho^2 \lambda I, $$

(A.2)

where

$$ I = \int_0^D \left[ p_u J_{\ell-1}(\Delta R) \left( h_{pr}^* - h_{\phi p}^* \right) + p_p J_{\ell+1}(\Delta R) \left( h_{pr}^* + i h_{\phi p}^* \right) \right] R \, dR $$

(A.3)

$$ = I_p^<(D), \quad D \leq 1 $$

(A.4)

$$ = I_p^<(1) + I_p^>(D), \quad D > 1 $$

and

$$ I_p^<(D) = \frac{2}{U_p} \left[ p_u Y_3^* \left( J_{\triangle u}, J_{\ell-1; 0, D} \right) + p_p Y_4^* \left( J_{\triangle u}, J_{\ell+1; 0, D} \right) \right] $$

(A.5)

$$ I_p^>(D) = \frac{2n_3}{W_p} \left[ p_u Y_3^* \left( J_{\triangle u}, J_{\ell-1; 1, D} \right) + p_p Y_4^* \left( J_{\triangle u}, J_{\ell+1; 1, D} \right) \right] $$

(A.6)

$F$ is the function defined by $^{13}$
\[ F(Z_{a,b,n+1};x,y) = \int_{x}^{y} R \left( \frac{R}{a^2 + b^2} \right)^{\frac{1}{2}} \frac{a}{n+1} (aR) K_{n+1} (bR) dR \]

\[ = \left[ \frac{R}{a^2 + b^2} \left( \frac{a}{n+1} (bR) Z_n (aR) - b K_n (bR) Z_{n+1} (aR) \right) \right]_{R=x}^{y} , \quad (A.7) \]

where \( Z_n \) is a cylinder function.

If we were to use the simplified modes given by Eqs. (2.2.29-35), the excitation coefficients would also simplify to

\[ a_p = \frac{E_0}{\varepsilon_{\infty}} \left( \frac{\mu_{\infty}}{\mu} \right)^{\frac{1}{2}} \frac{\varepsilon_1}{\mu_1}^{\frac{1}{2}} i^{\ell-1} , \quad (A.8) \]

where \( \xi \) is defined by Eq. (2.2.35),

\[ N = N^<(D) , \quad D \leq 1 \]
\[ = N^<(1) + N^>(D) , \quad D > 1 \quad (A.9) \]

and

\[ N^<(D) = I \left( J_{\Delta}, J_{U_p}, d_{\ell+1} ; 0, D \right) / J_{d_{\ell+1}} (U_p) , \quad (A.10) \]
\[ N^>(D) = F \left( J_{\Delta}, K_{W_p}, d_{\ell+1} ; D \right) / K_{d_{\ell+1}} (W_p) . \quad (A.11) \]

Here the upper sign applies to HE modes and the lower sign to EH modes.
APPENDIX 3B
CROSS-INTEGRALS FOR CALCULATION OF
POWER IN WAVEGUIDE

We define the quantities

\( \chi_M = \beta a_M^M - i \left( \frac{\mu_1}{\epsilon_1} \right)^{1/2} k_1 b_M^M \), \hfill (B.1)

\( \zeta_M = \chi_M (-i \rightarrow i) \), \hfill (B.2)

\( \chi_E = - i k_1 \left( \frac{\mu_1}{\epsilon_1} \right)^{1/2} \alpha_E^E - \beta b_E^E \), \hfill (B.3)

\( \zeta_E = \chi_E (-i \rightarrow i) \), \hfill (B.4)

\( \varepsilon_{j j'}^{(1)} = \begin{cases} 1 & j = j' \\ -1 & j \neq j' \end{cases} \), \hfill (B.5)

and

\( \varepsilon_{j j'}^{(2)} = \begin{cases} 1 & j = j' = 1; j = 1, j' = 2 \\ -1 & j = j' = 2; j = 2, j' = 1 \end{cases} \). \hfill (B.6)

Then the integrals in Eqs. (2.38 - 41) can be found from the following:

\[
C_{pq} = \frac{2i}{U P U Q} \left[ \gamma_1 P \gamma_3 ^* I \left( J_U, J_U, \ell; 0, 1 \right) \right.
\]

\[
- \gamma_2 P \gamma_4 ^* I \left( J_U, J_U, \ell+1; 0, 1 \right) \left. \right] \hfill (B.7)
\]
\[ C_{\lambda}(p;M,\lambda) = \frac{i\rho}{\psi_{M}^2} \left[ Y_{1,\lambda} Y_{M} I \left( J_{p}^u, J_{p}^\ell - 1, 0, 1 \right) \right. \]
\[ \left. - \gamma_{2,\lambda}^0 Y_{M} I \left( J_{p}^u, J_{p}^\ell + 1, 0, 1 \right) \right] \] (B.8)

\[ C_{\lambda}(p;E,\lambda) = C_{\lambda}(p;M,\lambda) \quad [M \rightarrow E] \] (B.9)

\[ C_{\lambda}(M,\lambda;p) = \frac{i\rho}{\psi_{M}^2} \left[ X_{M} Y_{\lambda}^* I \left( J_{p}^u, J_{p}^\ell - 1, 0, 1 \right) \right. \]
\[ \left. + \zeta_{M} Y_{\lambda}^* I \left( J_{p}^u, J_{p}^\ell + 1, 0, 1 \right) \right] \] (B.10)

\[ C_{\lambda}(E,\lambda;p) = C_{\lambda}(M,\lambda;p) \quad [M \rightarrow E] \] (B.11)

\[ C_{\lambda}(M,\lambda;M,\lambda') = \frac{\rho^2 (\psi_{M} \psi_{M'})^{\frac{1}{2}}}{2\nu^1} \left[ X_{M} Y_{\lambda'} I \left( J_{U}^u, J_{U}^\ell - 1, 0, 1 \right) \right. \]
\[ \left. + \zeta_{M} X_{\lambda'} I \left( J_{U}^u, J_{U}^\ell + 1, 0, 1 \right) \right] \] (B.12)

\[ C_{\lambda}(E,\lambda;E,\lambda') = C_{\lambda}(M,\lambda;M,\lambda') \quad [M \rightarrow E] \] (B.13)

\[ C_{\lambda}(E,\lambda;M,\lambda') = C_{\lambda}(M,\lambda;M,\lambda') \quad [\psi_{M} + \psi_{E}, \chi_{M} + \chi_{E}, \zeta_{M} + \zeta_{E}] \] (B.14)

\[ C_{\lambda}(M,\lambda;E,\lambda') = C_{\lambda}(E,\lambda;M,\lambda') \quad [M \leftrightarrow E] \] (B.15)

and

\[ C_{\lambda}(p;\lambda,\lambda) = \frac{1}{\sqrt{2}} \left\{ C_{\lambda}(p;M,\lambda) + E_{i1}^{(2)} C_{\lambda}(p;E,\lambda) \right\}, \] (B.16)

\[ C_{\lambda}(i,\lambda;p) = \frac{1}{\sqrt{2}} \left\{ C_{\lambda}(M,\lambda;p) + E_{i1}^{(2)} C_{\lambda}(E,\lambda;p) \right\}, \] (B.17)

\[ C_{\lambda}(i,\lambda;i',\lambda') = \frac{i\rho}{\psi_{M}^2} \left\{ C_{\lambda}(M,\lambda;M,\lambda') + E_{i1}^{(1)} C_{\lambda}(E,\lambda;E,\lambda') \right. \]
\[ \left. + E_{i1}^{(2)} \left( C_{\lambda}(E,\lambda;M,\lambda') + E_{i1}^{(1)} C_{\lambda}(M,\lambda;E,\lambda') \right) \right\}, \] (B.18)
where $\gamma_1$, $\gamma_2$, $\gamma_3$ and $\gamma_4$ are the quantities defined in Eqs. (2.2.13-16) with $n=1$ and $\gamma_M', \sigma_M', \gamma_E$ and $\sigma_E$ are defined in Eqs. (2.18-21). The function $I$ is defined in Eq. (2.26). The intricate notation is regrettable but necessary to avoid multiple listing of very similar equations.
APPENDIX 3C

COEFFICIENTS FOR INCOHERENT SOURCE

The integrals required in Eqs. (3.7-10) are

\[ D_{pq} = \int_0^D \left( h_p^* h_q - h_{pq}^* h_{pq} \right) R dR \]  \hspace{1cm} (C.1)

\[ D_{i}(p; i, \lambda) = \int_0^D \left( h_p^* h(i, \lambda) - h_{pz}^* h_{pz}(i, \lambda) \right) R dR \]  \hspace{1cm} (C.2)

and

\[ D_{i}(i, \lambda; i', \lambda') = \int_0^D \left( h_p^* h(i, \lambda) h_{pz}^* h(i', \lambda') \right) \] 

\[ - h_z^* h(i, \lambda) h_z(i', \lambda') \] 

\[ R dR . \]  \hspace{1cm} (C.3)

These can be calculated as follows:

\[ D_{pq} = D_{pq}^< , \quad D \leq 1 \]

\[ = D_{pq}^< (1) + D_{pq}^> (D) , \quad D > 1 \]  \hspace{1cm} (C.4)

where

\[ D_{pq}^< = \frac{2}{U_p U_q} \left[ \gamma_3^p \gamma_3^q I \left[ j_{p} j_{q} , \ell - 1 ; 0 , D \right] \right. \]

\[ + \gamma_4^p \gamma_4^q I \left[ j_{p} j_{q} , \ell + 1 ; 0 , D \right] \]  \hspace{1cm} (C.5)

\[ D_{pq}^> = \frac{2n_p n_q}{W_p W_q} \left[ \gamma_3^p \gamma_3^q G \left[ k_w , k_w , \ell - 1 , 1 , D \right] \right. \]

\[ + \gamma_4^p \gamma_4^q G \left[ k_w , k_w , \ell + 1 , 1 , D \right] \]  \hspace{1cm} (C.6)

and
where

\[ D_\lambda^<(p;M,\lambda) = \frac{\rho_{\psi} M^{-1/2}}{U_p} \left[ \gamma_3 \gamma_M^* I \left( J_U J_{U_p} \ell-1;0,D \right) \right. \]
\[ \left. + \gamma_4 \gamma_M^* I \left( J_U J_{U_p} \ell+1;0,D \right) \right] \]

\[ D_\lambda^>(p;M,\lambda) = \frac{\eta_3 \psi M^{-1/2}}{W_p} \left[ \left( \frac{\epsilon_2}{U_2} \right)^{1/2} \gamma_3 \gamma P \left( J_{\lambda}, K_{W_p} \ell-1,1,D \right) \right. \]
\[ \left. - \gamma_4 \gamma P \left( J_{\lambda}, K_{W_p} \ell+1,1,D \right) \right] \]
\[ + \frac{\rho}{\lambda} \left[ \gamma_4 \gamma_{SM}^* F \left( H_{\lambda}^{(2)}, K_{W_p} \ell-1,1,D \right) \right. \]
\[ \left. - \gamma_4 \gamma_{SM}^* F \left( H_{\lambda}^{(2)}, K_{W_p} \ell+1,1,D \right) \right] \]

\[ D_\lambda^<(p;E,\lambda) = D_\lambda^<(p;M,\lambda) \left[ M \leftrightarrow E \right] \]

\[ D_\lambda^>(p;E,\lambda) = D_\lambda^>(p;M,\lambda) \left[ \left( \frac{\epsilon_2}{U_2} \right)^{1/2} + \frac{i B}{k_2}, -\gamma_4^* + \gamma_4^*, M \leftrightarrow E \right] \]
\[ D_2(i, \lambda; i', \lambda') = \frac{1}{2} \left\{ D_2(M, \lambda; M, \lambda') + E^{(1)}_{ii'} D_2(E, \lambda; E, \lambda') \right. \\
+ E^{(2)}_{ii'} \left\{ D_2(E, \lambda; M, \lambda') + E^{(1)}_{ii'} D_2(M, \lambda; E, \lambda') \right\} \right\} (C.13) \]

where \( E^{(1)}_{ii'} \) and \( E^{(2)}_{ii'} \) are defined by Eqs. (B.5-6) and

\[ D_2^< (M, \lambda; M, \lambda') = \frac{\rho^2 (\Psi_M^* \Psi_M^{'})^{-\frac{1}{2}}}{2U^{'}} \left[ \begin{array}{c} \gamma^*_{M} \gamma^{*'}_{M} I \left[ J_{U}, J_{U'}, \lambda, 1; 0, D \right] \\
+ \sigma_{M} \sigma^{*'}_{M} I \left[ J_{U}, J_{U'}, \lambda, 1; 0, D \right] \end{array} \right], (C.14) \]

\[ D_2^< (E, \lambda; E, \lambda') = D_2^< (M, \lambda; M, \lambda') [M \rightarrow E], (C.15) \]

\[ D_2^< (E, \lambda; M, \lambda') = D_2^< (M, \lambda; M, \lambda') [\Psi_M^* \Psi_E^* \gamma_M^* \gamma_E^* \sigma_M^* \sigma_E], (C.16) \]

\[ D_2^< (M, \lambda; E, \lambda') = D_2^< (E, \lambda; M, \lambda') [M \leftrightarrow E], (C.17) \]

\[ D_2^> (M, \lambda; M, \lambda') = \frac{1}{2} (\Psi_M^* \Psi_M^{'})^{-\frac{1}{2}} \left\{ \begin{array}{c} \epsilon_2 \\
\frac{1}{U^{'}} \end{array} \right\} \left[ \begin{array}{c} I_{\lambda-1}^{(1)} + I_{\lambda+1}^{(2)} \\
I_{\lambda-1}^{(2)} + \sigma_{SM} I_{\lambda+1}^{(2)} \\
+ \frac{1}{\lambda'} \left[ \gamma^*_{SM} I_{\lambda-1}^{(3)} + \sigma^*_{SM} I_{\lambda+1}^{(3)} \right] \\
+ \frac{\rho^2}{\lambda \lambda'} \left[ \gamma_{SM} \gamma^*_{SM} I_{\lambda-1}^{(4)} + \sigma_{SM} \sigma^*_{SM} I_{\lambda+1}^{(4)} \right] \end{array} \right], (C.18) \]

\[ D_2^> (E, \lambda; E, \lambda') = \frac{1}{2} (\Psi_E^* \Psi_E^{'})^{-\frac{1}{2}} \left\{ \begin{array}{c} \epsilon_2 \\
\frac{1}{k_2^2} \end{array} \right\} \left[ \begin{array}{c} I_{\lambda-1}^{(1)} + I_{\lambda+1}^{(2)} \\
I_{\lambda-1}^{(2)} - \sigma_{SE} I_{\lambda+1}^{(2)} \\
- \frac{\beta}{\lambda'} \left[ \gamma_{SE} I_{\lambda-1}^{(3)} - \sigma_{SE} I_{\lambda+1}^{(3)} \right] \\
+ \frac{\rho^2}{\lambda \lambda'} \left[ \gamma_{SE} \gamma^*_{SE} I_{\lambda-1}^{(4)} + \sigma_{SE} \sigma^*_{SE} I_{\lambda+1}^{(4)} \right] \end{array} \right], (C.19) \]
\[ D^\lambda_\mathcal{E} (E, \lambda; M, \lambda') = \frac{1}{2} (\psi_E^* \psi_M') - \frac{1}{2} \left\{ \beta \left( \frac{\varepsilon_2}{\mu_2} \right) \frac{1}{k_2} \left\{ I_\lambda^{(1)} - I_\lambda^{(2)} \right\} \right. \\
\left. + \left[ \frac{\varepsilon_2}{\mu_2} \right]^2 \rho \lambda \left( \gamma_{SE} I_\lambda^{(2)} + \sigma_{SE} I_\lambda^{(2)} \right) \right. \\
\left. + \frac{i \beta}{k_2} \frac{\rho}{\lambda} \left[ \gamma_{SM} I_\lambda^{(3)} - \sigma_{SM} I_\lambda^{(3)} \right] \right\} \]  
\[ + \frac{\rho^2}{\lambda \lambda'} \left\{ \gamma_{SM} I_\lambda^{(4)} + \sigma_{SM} I_\lambda^{(4)} \right\} \]  
(\text{C.20})

\[ D^\lambda_\mathcal{M} (M, \lambda; E, \lambda') = \frac{1}{2} (\psi_M^* \psi_E') - \frac{1}{2} \left\{ \beta \left( \frac{\varepsilon_2}{\mu_2} \right) \frac{1}{k_2} \left\{ I_\lambda^{(1)} - I_\lambda^{(2)} \right\} \right. \\
\left. - \left[ \frac{\varepsilon_2}{\mu_2} \right]^2 \rho \lambda \left( \gamma_{SM} I_\lambda^{(2)} - \sigma_{SM} I_\lambda^{(2)} \right) \right. \\
\left. - \frac{i \beta}{k_2} \frac{\rho}{\lambda} \left[ \gamma_{SM} I_\lambda^{(3)} + \sigma_{SM} I_\lambda^{(3)} \right] \right\} \]  
\[ + \frac{\rho^2}{\lambda \lambda'} \left\{ \gamma_{SM} I_\lambda^{(4)} + \sigma_{SM} I_\lambda^{(4)} \right\} \]  
(\text{C.21})

The functions \( I_n^{(1)} \) are defined by

\[ I_n^{(1)} = I \left[ J_\lambda, J_\lambda', n; 1, D \right] \]  
(\text{C.22})

\[ I_n^{(2)} = I \left[ H_\lambda^{(1)}, J_\lambda', n; 1, D \right] \]  
(\text{C.23})

\[ I_n^{(3)} = I \left[ J_\lambda', H_\lambda^{(2)}, n; 1, D \right] \]  
(\text{C.24})

\[ I_n^{(4)} = I \left[ H_\lambda^{(1)}, H_\lambda^{(2)}, n; 1, D \right] \]  
(\text{C.25})

The bound mode coefficients are given in an approximate form in Ref. 28.
REFERENCES


13. N.W. McLachlan, Bessel Functions for Engineers (Oxford University Press, London, 1941).


16. A. Hart, Indiana University (A.N.U. Program library routine "SQADAP").


CHAPTER 4

ASYMPTOTIC EXPANSION OF THE RADIATION FIELD WITHIN
A CIRCULAR DIELECTRIC WAVEGUIDE – LEAKY MODES

4.1 INTRODUCTION

In the preceding two chapters, we have developed an expansion of the fields carried by a dielectric waveguide of the form

\[ E = E_B + E_{UB}, \]  

(1.1)

where \( E_B \) is a discrete sum over bound modes and \( E_{UB} \) is a superposition of unbound modes given by

\[ E_{UB} = \sum_{\ell=0}^{\infty} \int_{0}^{\infty} E_{\ell}(\lambda) e^{-i\beta z} d\lambda, \]  

(1.2)

where

\[ E_{\ell}(\lambda) = 2 \sum_{i=1}^{2} \sum_{n=-\ell,\ell} A_n^{(i)}(\lambda) \xi^{(i)}(\lambda,n). \]  

(1.3)

The factor \( e^{i\omega t} \) has again been excluded, \( A_n^{(i)} \) is an amplitude coefficient, \( \xi^{(i)}(\lambda,n) \) is a radiation mode defined by Eqs. (2.4.43-44) and the second summation in Eq. (1.3) is simply the \( n = 0 \) term when \( \ell = 0 \) and the sum of \( n = -\ell \) and \( n = \ell \) terms otherwise. A formally identical expansion exists for the magnetic field.

As we have seen in Chapter 3, this "transverse spectral representation" (so called because Eq. (1.1) involves a summation over \( \lambda \), the transverse component of the modal wave vector) can be rather slowly convergent. An alternative representation more commonly used in planar
systems is the longitudinal spectral representation in which the summation is carried out with respect to the longitudinal component, $\beta$. While it has no particular virtue in itself, the latter representation is more directly amenable to asymptotic expansion at large axial distances from the source (i.e. large $z$).

The method widely used in the derivation of asymptotic expansions of the fields radiated from antennae or reflected from plane interfaces is the steepest descent approximation. In this method, the contour of integration in the $\beta$-surface is deformed into one on which the integrand has constant phase and is sharply peaked at a saddle point. In performing this deformation, a number of poles of the integrand are traversed and contributions from the residues at these poles, together with the saddle-point contribution, constitute the asymptotic approximation to the waveguide field.

The poles in question correspond to resonant frequencies, and hence discrete modes, of the system. We shall see in Section 4.3 that the sum over real resonant frequencies is identical with the proper, bound mode contribution in Eq. (1.1). However, even in a non-dissipative system, some of the resonant frequencies encountered may be complex, representing damped resonances.

The complex ("leaky") modes decay in the direction of propagation but are "improper" in the sense that while they are solutions of the source-free field equations, their fields become infinite at infinitely large radial distances from the waveguide (see Section 4.3). This means that while the complex modes have the same mathematical form as the bound modes (extended below cutoff) and their propagation characteristics are obtained from a transverse resonance
calculation which is formally identical with the solution of the bound-
mode eigenvalue equation, they cannot be orthonormalized in the same way
as the proper modes. But this does not prevent their being used as a
rapidly convergent representation of the radiation field within a
restricted spatial region.  

The usefulness of such a leaky mode expansion appears to have
been suggested originally by Marcuvitz and the method has been widely
adopted in the treatment of problems involving a source and plane,
homogeneous interfaces, with particular emphasis on antenna
applications. Tamir and Oliner have made a particularly
detailed study of leaky modes at a plane interface.

Some similar work on the circular dielectric waveguide has
been done by Andersen who calculated the far-field radiation from a
dielectric rod excited by a magnetic ring source. As in the studies of
planar systems, the emphasis here was placed on the "space-wave"
(saddle-point) contribution as this decays algebraically along the
interface while the leaky mode decays exponentially and is therefore
negligible after a sufficiently large distance.

However, while in the planar systems the distance involved is
actually quite small, in an optical fibre a "sufficiently large
distance" may be of the order of kilometres — by which time the space
wave itself will have decayed beyond importance. This possibility was
indicated as early as 1961 by Snitzer's observation of a mode pattern
below the theoretical cutoff frequency, suggesting that the radiation
field indeed behaved very much like a single leaky mode. But it was not
this observation which sparked off the renewed interest in leaky modes
for optical fibre systems. It was rather the theoretical observation
that a number of rays which would have been totally internally reflected from a plane boundary in fact leak energy away from the curved boundary of a cylindrical waveguide - but that they do so quite slowly.\textsuperscript{27-30}

Such rays give rise to a class of weakly leaky modes on the circular cylindrical waveguide which do not exist in the slab. These are the "tunnelling" leaky modes - so called because their fields pass from a radially propagating region in the core, through an evanescent region and then propagate out to infinity.\textsuperscript{31}

The remainder of the leaky modes are more strongly leaky and have been labelled "refracting"\textsuperscript{31} because they correspond to rays which would not be totally reflected, even at a plane interface. All the leaky modes on a homogeneous slab waveguide are, obviously, refracting.

The general behaviour of both the tunnelling and refracting modes has recently been described from a largely heuristic point of view\textsuperscript{31-33} and a great deal of attention is also being given to the ray-optical approximations to these modes.\textsuperscript{31,34-39} The purpose of this chapter is to derive the leaky modes as part of an asymptotic approximation to the radiation field, to estimate their range of importance and to obtain orthogonality conditions by which their amplitudes can be determined without using the continuous spectrum. Chapter 5 will then determine the detailed characteristics of the leaky modes and use these to give some criterion for the region of validity of the leaky-ray approximation.
4.2 ALTERNATIVE REPRESENTATIONS OF THE WAVEGUIDE FIELD — STEEPEST DESCENT APPROXIMATION

Before transforming from the transverse to the longitudinal spectral representation, it is necessary to study the analytic behaviour of $E_\phi(\lambda)$ in some detail. This will be done explicitly for the case of plane wave excitation discussed in Section 3.2 but the principle will be the same in general.

We begin by factoring $E_\phi(\lambda)$ in the form

$$E_\phi(\lambda) = \frac{E_\phi(\lambda)}{[M(\lambda) M^*(\lambda)]}, \quad (2.1)$$

where $M(\lambda)$ is given by Eq. (2.3.31). This quantity can, in turn, be factored as

$$M(\lambda) = \left[N - \frac{\mu_1}{\mu_2} N_1\right] \left[N - \frac{\varepsilon_1}{\varepsilon_2} N_1\right] - \frac{V}{U\lambda} \left[\frac{\beta}{k_2}\right]^2.$$

$$= G_+ G_-^* \quad (2.2)$$

where

$$G_\pm = N_2 \pm \frac{1}{2} \left[\frac{\mu_1}{\mu_2} + \frac{\varepsilon_1}{\varepsilon_2}\right] N_1 \pm \frac{1}{4} \left[\frac{\mu_1}{\mu_2} - \frac{\varepsilon_1}{\varepsilon_2}\right]^2 N_1^2 \pm \left[\frac{V}{U\lambda}\right]^2 \left[\frac{\beta}{k_2}\right]^2 \sqrt{2}. \quad (2.3)$$

The quantities in Eqs. (2.2-3) are as defined in Section 2.3.

On making the change of variable $\lambda = -iW$, it is clear that $M(\lambda) = 0$ is precisely the eigenvalue equation given in Eq. (2.2.21) and, in fact, $G_+ = 0$ corresponds to HE modes while $G_- = 0$ gives the EH modes. $E_\phi(\lambda)$ therefore has a simple pole at each bound mode, as well as at any other zero of $M(\lambda)$ or $M^*(\lambda)$.

---

Throughout Sections 4.2 and 4.3, complex conjugations are to be interpreted as $[g(\lambda)]^* = g^*(\lambda)$ and not $g^*(\lambda^*)$ so that $|g(\lambda)|^2 = g^*(\lambda) g(\lambda)$ for all functions $g$. This ensures that the analytic continuation from the real $\lambda$ axis into the complex $\lambda$ plane is correctly carried out.
4.2

There may also be branch-point singularities at

\[ U = (V^2 + \lambda^2)^{1/2} = 0 \] and/or \( \beta = \left( k_2^2 - \frac{\lambda^2}{\rho^2} \right)^{1/2} = 0 \). To investigate this possibility, we consider the \( z \)-component of \( E_{kz} \) in the core,

\[
E_{kz} = C \left( I_M a_k \psi_M^{-1/2} I_E b_k \psi_E^{-1/2} \right) J_k(\text{UR}) \cos \phi, \tag{2.4}
\]

where

\[
C = 2\pi \rho^2 E_0 \sqrt{2} \tag{2.5}
\]

and, assuming \( D \leq 1 \),

\[
I_M = \frac{\rho}{\psi_M^{1/2}} \left[ p_- \gamma_M I \left( J_{U, \Delta}^{\lambda-1;0,D} \right) + p_+ \sigma_M I \left( J_{U, \Delta}^{\lambda+1;0,D} \right) \right] \tag{2.6}
\]

\[
I_E = I_M \left[ M + E \right]. \tag{2.7}
\]

All other quantities are as previously defined.

Under the change of variable \( U \rightarrow -U \), the right hand side of Eq. (2.4) remains unchanged so that the branch of the square root \( (k_1^2 - \beta^2)^{1/2} \) can be chosen arbitrarily. \( U = 0 \) is thus not a branch point of the integrand. On the other hand, under the change of variable \( \beta \rightarrow -\beta \), \( E_{kz} \) changes from a forward- to a backward-travelling mode, and so the points \( \lambda = \pm \rho k_2 \) are branch points.

The reason for the difference in singularity for \( U \) and \( \beta \) is connected with the fact that there is a radiation condition at infinity necessitating a proper choice of sign for \( \beta \) whereas \( U \) is not similarly restricted.

The branch cuts from \( \lambda = \pm \rho k_2 \) are chosen as shown in Fig. 4.1 so that on the top sheet of the \( \lambda \)-surface, \( \text{Im}(\beta) = \beta_1 < 0 \) everywhere while on the bottom sheet, \( \beta_1 > 0 \). Thus forward modes on the top sheet are
Fig. 4.1: The $\lambda$-surface showing branch cuts from $\lambda = \pm \rho k_2$, forward bound mode poles on the upper sheet of the fourth quadrant (x) and forward leaky mode poles on the lower sheet of the first quadrant (*). The location of the poles is purely schematic. Small losses have been assumed to displace the branch cuts and poles from the axes.

either bound or leaky modes while those on the bottom sheet grow in the direction of propagation. The branch lines are defined by the equation

$$\lambda_r \lambda_i = \rho^2 k_2 r_{2r} k_{2i},$$  \hspace{1cm} (2.8)

where subscripts $r$ and $i$ stand for real and imaginary parts respectively and we have introduced small losses into the cladding medium for the purposes of this diagram to separate the branch lines from the real and imaginary $\lambda$ axes.

In a lossless medium, the bound mode poles lie on the negative imaginary axis. Forward leaky modes lie in the first quadrant of the
top sheet. Any other poles on the top sheet would correspond to incoming waves growing in either the axial (2nd quadrant) or radial (3rd quadrant) direction of propagation or to radially outgoing backward modes, again growing in the direction of propagation (4th quadrant).

Returning to the function $E_y(\lambda)$, if we now make the change of variable $\lambda \rightarrow -\lambda$ in Eqs. (2.4-7), we find that $E_y$ is an odd function of $\lambda$. This causes some difficulty because the transformation to the longitudinal representation is most conveniently performed if the range of integration in Eq. (1.2) can be extended from $(0, \infty)$ to $(-\infty, \infty)$. This cannot be done directly but is achieved as follows:

We note that for real $\lambda$,

$$G_\pm^*(-\lambda) = G_\pm(\lambda) \quad (2.9)$$

and

$$G_\pm^*(\lambda) - G_\pm(\lambda) = N_2^* - N_2$$

$$= \frac{4i}{\pi \lambda^2} |H_2^{(2)}(\lambda)|^{-2}, \quad (2.10)$$

where the Wronskian determinant for the Hankel functions has been used in obtaining the last line.

Then we can write

$$\frac{1}{|M(\lambda)|^2} = \frac{1}{G_+^*(\lambda)G_-^*(\lambda) - G_+^*(\lambda)G_-^*(\lambda)}$$

$$\times \left[ G_+^*(\lambda)G_-^*(\lambda) - G_+^*(\lambda)G_-^*(\lambda) \right]^{-1}$$

$$= \frac{\pi \lambda^2 |H_2^{(2)}(\lambda)|^2}{4i(N_2 + N_2^* - 2N_1)} \left[ \frac{1}{G_+(\lambda)G_-^*(\lambda) - G_+^*(\lambda)G_-^*(\lambda)} \right]. \quad (2.11)$$

\[+\] In the case of the slab waveguide, $E_y$ is an even function of $\lambda$ but the procedure described here is still necessary to obtain a longitudinal representation in which there are no pole contributions.\[40\]
But since
\[ G_+^*(\lambda) + G_-^*(\lambda) = N_2 + N_2^* - 2N_1 = G_-(\lambda) + G_+^*(\lambda), \]
we can also rewrite Eq. (2.11) as
\[ \frac{1}{|M(\lambda)|^2} = \frac{\pi \lambda^2 |\mathcal{H}_Y^{(2)}(\lambda)|^2}{G_+(\lambda) + G_-^*(\lambda)} \left( \frac{1}{G_+(\lambda)G_-(\lambda)} - \frac{1}{G_+^*(\lambda)G_-^*(\lambda)} \right). \] (2.12)

Then substituting Eqs. (2.1) and (2.12) in Eq. (1.2) and making the change of variable \( \lambda \rightarrow -\lambda \) in the second half of the integrand, we find
\[ E_{\text{UB}} = \sum_{\ell=0}^{\infty} \int_{-\infty}^{\infty} \tilde{f}_\ell(\lambda) e^{-i\beta z} d\lambda, \] (2.13)
where
\[ \tilde{f}_\ell(\lambda) = \frac{\pi \lambda^2 |\mathcal{H}_Y^{(2)}(\lambda)|^2}{[G_+(\lambda) + G_-^*(\lambda)]} \cdot \frac{E_{\ell,\lambda}}{G_+^*(\lambda)G_-^*(\lambda)}. \] (2.14)

The choice of upper or lower signs in the denominator of Eq. (2.14) is arbitrary. The integration path \( \Gamma \) is shown in Fig. 4.1.

If we now form the closed contour \( \Gamma + \Gamma_S - \Gamma_B \) as illustrated in Fig. 4.1, then we find
\[ \int_{\Gamma} \tilde{f}_\ell(\lambda) e^{-i\beta z} d\lambda = \int_{\Gamma_B} \tilde{f}_\ell(\lambda) e^{-i\beta z} d\lambda - 2\pi i \sum_{\text{poles}} R(\tilde{f}_\ell,\lambda), \] (2.15)
where \( R(\tilde{f}_\ell,\lambda) \) is the residue of \( \tilde{f}_\ell \) at a bound mode pole (the only poles crossed by this contour deformation). The integral along the semi-circular contour \( \Gamma_S \) makes no contribution (by Jordan's lemma).

We shall see in Section 4.3 that the sum over poles in Eq. (2.15) (when summed over \( \ell \)) is identical with the bound-mode field in Eq. (1.1). Thus the integral around the branch cut on the right hand
side of Eq. (2.15) represents the total waveguide field. The required longitudinal representation is now found by transforming that integral into the $\beta$-plane.

Putting

$$\beta = \left( k_z^2 - \frac{\lambda^2}{\rho^2} \right)^{1/2},$$

$$d\beta = -\frac{\lambda}{\rho^2 \beta} d\lambda \quad (2.16)$$

we find

$$E = \sum_{\ell=0}^{\infty} \int_{-\infty}^{\infty} f_\ell(\lambda) \frac{\rho^2 \beta}{\lambda} e^{-i\beta z} d\beta,$$ \quad (2.17)

where the integration is carried out along the path $P$ in Fig. 4.2. The integrand has branch points in the $\beta$-plane at $\beta = \pm k$ ($\lambda = 0$) because of the presence of the Hankel functions. Physically, these branch points are again associated with the fact that we are dealing with an open structure for which a radiation condition must be satisfied. By choosing the branch cuts as shown, we ensure that on the top sheet of the $\beta$-surface, $\lambda_i < 0$ everywhere while on the bottom sheet $\lambda_i > 0$. Thus the top sheet corresponds to modes which decay in the radial direction and the bottom sheet to modes which grow as $r$ increases. The positions of the bound and leaky modes are shown in Fig. 4.2. Once again, in drawing this figure, small losses have been assumed so that the branch lines are defined by

$$\beta_r \beta_i = k_{2r} k_{2i}.$$ \quad (2.18)

An asymptotic expansion of Eq. (2.17) valid at large distances from the source is now found using the method of steepest descents. We begin by making yet another change of variable to remove the multiple-valuedness of the integrand in the $\beta$-surface.
Fig. 4.2: The β-surface showing branch cuts at $\beta = \pm k_2^*$, forward bound mode poles (x) on the upper sheet and leaky mode poles (*) on the lower sheet of the fourth quadrant. Again, the location of the poles is schematic and small losses have been assumed.

Putting

$$\beta = k_2 \sin \psi ,$$

and

$$\lambda = \rho k_2 \cos \psi ,$$

where

$$\psi = \xi + i \eta ,$$

the two-sheeted β-surface maps onto the ψ-plane as shown in Fig. 4.3 where $T_n$ is the nth quadrant of the top sheet ($\lambda_1 > 0$) and $B_n$ the corresponding quadrant of the bottom sheet. There are no branch cuts in the ψ-plane since both sheets of the β-surface are mapped into a simply connected plane.
Fig. 4.3: The \( \psi \)-plane. Region \( T_n (B_n) \) corresponds to the \( n \)th quadrant of the top (bottom) sheet of the \( \beta \)-surface in Fig. 4.2. The contour \( P \) in Fig. 4.2 transforms to \( P' \) as shown and the bound and leaky mode poles now lie in the positions indicated. The contour labelled SDP is the steepest descent path in the case where the field inside the waveguide core is being considered.

With this transformation, the pole singularities on the \( \beta \)-surface map into regions \( T_n \) (proper modes) and \( B_n \) (improper modes) and the contour of integration \( P \) (i.e. the line \( \beta_1 = 0 \)) maps into \( P' \) as shown in Fig. 4.3.

In this new coordinate system, Eq. (2.17) becomes

\[
E = \sum_{\ell=0} \int_{P'} f_\ell (\lambda) e^{-ikz} \sin \psi z \rho k_2 \sin \psi \, d\psi.
\]  

(2.22)
Defining spherical coordinates in space by

\[ r = R \rho = \zeta \cos \theta \]  
(2.23)

and

\[ z = \zeta \sin \theta , \]  
(2.24)

and using the asymptotic form of \( H_{2}^{(2)}(\xi) \) for large argument, we see that for large \( R \), the integrand in Eq. (2.22) contains a factor

\[ \exp \left[ -ik_{2} \zeta (\cos \psi \cos \theta + \sin \psi \sin \theta) \right] = \exp \left[ -ik_{2} \zeta \cos (\psi - \theta) \right] . \]

The integrand therefore has a stationary phase point at \( \psi = \theta \) and the path of steepest descent is one passing through this point.

The equation of the steepest descent path is found by requiring that the phase, the real part of \( k_{2} \zeta \cos (\psi - \theta) \), remain constant. This gives

\[ \cos (\xi - \theta) \cosh \eta = 1 . \]  
(2.25)

When the path of integration \( P' \) is deformed into the steepest descent path, SDP, any poles traversed will contribute to the expansion for the waveguide field — whether they lie on the proper or improper sheets of the \( \beta \)-surface. This is how the leaky modes enter the expansion.

For the field inside the waveguide, \( \theta = \frac{\pi}{2} \) so the steepest descent path is the curve

\[ \sin \xi \cosh \eta = 1 \]  
(2.26)

as illustrated in Fig. 4.3. On the \( \beta \)-surface, this curve corresponds to the line \( \beta_{r} = k_{2} \). Thus the path of integration traverses all the leaky mode poles corresponding to outward-going waves. Any poles which lie to the left of the original contour correspond to incoming waves \( \lambda_{r} < 0 \) and are therefore not crossed during the deformation of the integration path.
4.3

So, finally, the waveguide electric field, Eq. (1.1) can be rewritten in the form

\[ E = \sum_{\ell=0}^{\infty} \left\{ 2\pi i \sum_{\text{poles}} R(\ell, \lambda) e^{-i\beta z} + \int_{\text{SDP}} f_\ell(\lambda) e^{-ik z} \sin \psi \rho k^2 \sin \psi \, d\psi \right\}, \quad (2.27) \]

where the sum over poles now includes not only bound modes (\( \lambda_r = 0, \lambda_i < 0 \)) but also leaky modes (\( \lambda_r > 0, \lambda_i > 0 \)). Any other poles correspond to waves growing in the positive z-direction and so are not encountered as they lie in the first and second quadrants of the \( \beta \)-surface.

In the following section we evaluate the pole contributions and examine in a little more detail the behaviour of the leaky modes. The saddle-point contribution is considered in Section 4.4.

4.3 POLE CONTRIBUTIONS

Simple pole singularities of \( f_\ell \) occur at zeros, \( \lambda_p \), of the function \( M \) given by Eq. (2.2). Zeros of \( G_+ \) correspond to HE modes (either above or below cutoff) and zeros of \( G_- \) to EH modes. From Eq. (2.14) the residue of \( f_\ell \) at such a pole is given by

\[ R[f_\ell, \lambda_p] = \lim_{\lambda \to \lambda_p} \left[ \frac{\lambda - \lambda_p}{G_\pm(\lambda_p)} \right] \frac{\pi \lambda_p^2}{4i} \left| H_\ell(\lambda_p) \right|^2 \frac{f_\ell(\lambda_p)}{\left| G_\mp(\lambda_p) \right|^2} \]

\[ = \left[ \left\{ \frac{dG_\pm}{d\lambda} \right\}_{\lambda = \lambda_p} \right]^{-1} \frac{\pi \lambda_p^2}{4i} \left| H_\ell(\lambda_p) \right|^2 \frac{f_\ell(\lambda_p)}{\left| G_\mp(\lambda_p) \right|^2}, \quad (3.1) \]

where the upper sign corresponds to HE modes and the lower sign to EH modes. The fact that \( G_\pm(\lambda_p) = 0 \) has been used to simplify the denominator of Eq. (2.14).
As an example, we shall evaluate these residues in the case
where \( \mu_1 = \mu_2, \epsilon_1 = \epsilon_2 \) and \( d \ll \rho \) to simplify algebra slightly. In this
approximation,

\[
G_\pm = N_2 - N_1 \pm \left( \frac{V}{U \lambda} \right)^2 \frac{\ell \beta}{k}
\]

\[
= \frac{H^{(2)}_{\ell-1}(\lambda)}{\lambda H^{(2)}_{\ell}(\lambda)} - \frac{J_{\ell-1}(U)}{U J_{\ell}(U)} + \xi \left( \frac{V}{U \lambda} \right)^2 \left( -1 \mp \frac{\beta}{k} \right),
\]

(3.2)

where \( k \) is written for either \( k_1 \) or \( k_2 \).

Differentiating with respect to \( \lambda \),

\[
\frac{1}{\lambda} \frac{dG_\pm}{d\lambda} = - H^{(2)}_{\ell} \left( 1 - \frac{H^{(2)}_{\ell-1}}{H^{(2)}_{\ell}} \right) + J^{(2)}_{\ell} \left( 1 - \frac{J_{\ell-1}}{J_{\ell}} \right)
\]

\[
- 2kV^2 \left( -1 \pm \frac{\beta}{k} \right) \frac{\lambda^2 + \lambda^2}{(U \lambda)^4},
\]

(3.3)

where

\[
H_n = \frac{H^{(2)}_{n-1}(\lambda)}{\lambda H^{(2)}_{n}(\lambda)} \quad \text{and} \quad J_n = \frac{J_{n-1}(U)}{U J_{n}(U)}.
\]

(3.4)

On substituting \( \lambda = \lambda_p \) and using \( G_\pm(\lambda_p) = 0 \), we find

\[
\left( \frac{dG_\pm}{d\lambda} \right)_{\lambda = \lambda_p} = \frac{J_{\ell-2}(U \lambda_p)}{J_{\ell}(U \lambda_p)} \frac{V^2}{U \lambda_p^2} - \frac{\ell \beta^2}{U \lambda_p^2} \left( 1 \mp \frac{\beta}{k} \right)
\]

\[
\times \left[ 2J_{\ell} - \frac{V}{U \lambda_p} \right]^2 \left( 2\lambda_p^2 - \xi \left( 1 \mp \frac{\beta}{k} \right) \right).
\]

(3.5)

We also have that

\[
|G_\pm(\lambda_p)|^2 = 4 \left( \frac{V}{U \lambda_p} \right)^2 \left( \frac{\ell \beta}{k} \right)^2.
\]

(3.6)

It remains now to evaluate \( E_\ell(\lambda_p) \). We begin with the
z-component inside the waveguide core. From Eq. (2.4),
\[ E_{xz} = C |M|^2 \left( a_\lambda \psi_{M}^{-1/2} - b_\lambda \psi_{E}^{-1/2} \right) J_0(UR) \cos \phi \]

\[ = \frac{4 \rho C J_0(UR) \cos \phi}{(\pi \rho k_\lambda)^2 U |H_\lambda(2)(\lambda)|^2} (S_+ + S_-) \]  

where

\[ S_\mp = p_\mp I^{(\mp)} \left( k_1 \left[ \frac{\epsilon_1}{\mu_1} \right]^{1/2} \left| N_2 - \frac{\mu_1}{\mu_2} N_1 \right|^2 + \left( \frac{\epsilon_2}{\mu_2} \right)^{1/2} \left[ \left( N_2 - \frac{\mu_1}{\mu_2} N_1 \right) + \left( N_2 - \frac{\epsilon_1}{\epsilon_2} N_1 \right)^* \right] \right) \]  

and for brevity we have put

\[ I^{(\mp)} = I \left( J_{\lambda}^{0}, J_{\Delta}^{\pm \mp 1, 0, 0} \right). \]  

Now making the approximation \( \epsilon_1 \approx \epsilon_2 \) (with \( \mu_1 = \mu_2 \)) and using Eq. (3.6), we find

\[ E_{xz} = \frac{8 \rho C J_0(UR) \cos \phi}{U |H_\lambda(2)(\lambda)|^2 (\pi \rho k_\lambda)^2} \left( \frac{\epsilon}{\mu} \right)^{1/2} \left( \frac{\epsilon_1}{k} \right)^{1/2} \left( \frac{V}{U \lambda} \right)^{2} \times \left( (k \mp \beta) p_- I^{(-)} + (k \mp \beta) p_+ I^{(+) \mp 1, 0} \right) \]  

where, again, the upper sign holds for HE modes and the lower sign for EH modes.

The residue of \( f_{xz} \) at \( \lambda_p \) is then found by combining Eqs. (3.1, 5, 6, 10). Thus,

\[ 2\pi i \left( f_{xz}, \lambda_p \right) = \frac{E_0 i J_0(UR) \cos \phi}{\rho U p k_\beta} \frac{(k \pm \beta) p_- I^{(-)} + (k \mp \beta) p_+ I^{(+) \mp 1, 0}}{dG_{\pm}(\lambda_p)/d\lambda} \]  

In the case where \( \beta \approx k \), only the first term of \( \frac{dG_{\pm}(\lambda_p)}{d\lambda} \) need be retained and Eq. (3.11) simplifies to
where $\theta_p$ is given by Eq. (2.2.28). For bound modes, $\lambda_p$ is pure imaginary ($\lambda_p = -iW$) and Eq. (3.12) is identical with the result obtained by Snyder for the $\beta = k$ bound modes. But Eqs. (3.11) and (3.12) are also valid when $\lambda_p$ is complex and give the $z$-component of the leaky mode contribution to the field as simply the analytic continuation of the corresponding bound mode term below cutoff, with a decay factor $e^{iz}$ ($\beta_i < 0$).

Carrying through exactly the same procedure with the transverse components of $E_{z\nu}$ we find, after a considerable amount of algebra, that in the core

$$2\pi R \left[ f_{z\nu}, \lambda_p \right] = \frac{E_0 i^{\ell+1} \lambda_p}{2W_2 \beta k J_{\lambda_p}^2 (U_p) \frac{dG_{\pm}}{d\lambda_p} (\lambda_p)}$$

$$\times \left\{ \left[ (\beta \pm k)^2 p^- I_{\nu}^{(-)} + (k^2 - \beta^2) p^+ I_{\nu}^{(+)} \right] J_{\lambda-1}^2 (U_R) \cos(\ell-1)\phi 
+ \frac{1}{i} \left[ (k^2 - \beta^2) p^- I_{\nu}^{(-)} + (\beta \pm k)^2 p^+ I_{\nu}^{(+)} \right] J_{\lambda+1}^2 (U_R) \cos(\ell+1)\phi \right\}$$

and

$$2\pi R \left[ f_{y\nu}, \lambda_p \right] = \frac{E_0 i^{\ell+1} \lambda_p}{2W_2 \beta k J_{\lambda_p}^2 (U_p) \frac{dG_{\pm}}{d\lambda_p} (\lambda_p)}$$

$$\times \left\{ \pm \left[ (\beta \pm k)^2 p^- I_{\nu}^{(-)} + (k^2 - \beta^2) p^+ I_{\nu}^{(+)} \right] J_{\lambda-1}^2 (U_R) \sin(\ell-1)\phi 
+ \frac{1}{i} \left[ (k^2 - \beta^2) p^- I_{\nu}^{(-)} + (\beta \pm k)^2 p^+ I_{\nu}^{(+)} \right] J_{\lambda+1}^2 (U_R) \sin(\ell+1)\phi \right\}.$$

Again, in the limit $\beta \approx k$, these reduce to the approximate bound mode
expressions given in Eqs. (2.2.29 - 2.2.35) and (3.A.8 - 11), namely

\[ 2\pi i R \left\{ f_{\ell x}, \lambda_p \right\} = \frac{2\lambda^2}{\nu^2} \frac{E_0}{J_{\ell,0}(\nu \rho)} J_{\ell-2}(\nu \rho) J_{\ell-1}(\nu \rho) \cos(\ell-1)\phi \]  

(3.15)

and

\[ 2\pi i R \left\{ f_{\ell y}, \lambda_p \right\} = -\tan(\ell-1)\phi \left\{ 2\pi i R \left\{ f_{\ell x}, \lambda_p \right\} \right\} . \]  

(3.16)

The electric field components in the cladding are found from Eqs. (3.11 - 16) by making the transformation

\[ \frac{J_n(\nu \rho)}{J_n(\nu \rho)} \rightarrow \frac{H_n^{(2)}(\nu \rho)}{H_n^{(2)}(\nu \rho)}, \]  

(3.17)

where \( n = \ell \) for the longitudinal component and \( n = \ell \pm 1 \) for the transverse components.

As we have mentioned above, for bound modes \( \lambda_p = -i\omega \) where \( \omega \) is real so that

\[ \frac{H_n^{(2)}(\nu \rho)}{H_n^{(2)}(\nu \rho)} = \frac{K_n(\omega \rho)}{K_n(\omega \rho)} \]  

(3.18)

and the cladding field has the required exponentially decaying behaviour for large \( \rho \). However, for leaky modes, \( \lambda_p \) is complex with positive real and imaginary parts. The principal asymptotic form

\[ H_n^{(2)}(\nu \rho) \sim \left( \frac{2}{\pi i \nu \rho} \right)^{1/2} \exp\left[-i(\nu \rho - \frac{1}{2}\pi i\ell - \frac{i}{2}\pi)\right] \]  

(3.19)

then indicates that for large \( \rho \), the leaky mode represents an exponentially growing, outgoing cylindrical wave. If we look at the more accurate asymptotic forms
\[ H_n^{(2)}(x) \sim \left( \frac{2}{\pi} \right)^{1/2} \exp \left[ -i \left( (x^2 - n^2)^{1/2} - n \cos^{-1} \left( \frac{n}{x} \right) \right) \right] \left( \frac{x^2 - n^2}{4} \right)^{1/4}, \]

\[
\text{Re}(x) > n, \ |x| >> 1 \quad \text{and} \quad |n - x| >> n^{1/3},
\]

\[ H_n^{(2)}(x) \sim i \left( \frac{2}{\pi} \right)^{1/2} \exp \left[ n \cosh^{-1} \left( \frac{n}{x/2} \right) - (n^2 - x^2)^{1/2} \right] \left( n^2 - x^2 \right)^{1/4}, \]

\[
\text{for } n > \text{Re}(x), \ n >> 1 \quad \text{and} \quad |n - x| >> n^{1/3},
\]

\[
\sim i \left( \frac{2}{\pi} \right)^{1/2} \exp \left[ \frac{n}{3} \left( 1 - \frac{x^2}{n^2} \right)^{3/2} + \frac{2n}{5} \left( 1 - \frac{x^2}{n^2} \right)^{5/2} + \ldots \right] \left( n^2 - x^2 \right)^{1/4},
\]

\[
\text{for } n \geq \text{Re}(x) \quad (3.22)
\]

then we see that for high-order modes, \((\lambda >> 1)\), the cladding field can be divided into two regions. (Re stands for the real part.)

When \(\text{Re}(\lambda_p)R < \lambda\), the electric field exponentially decays from the surface of the waveguide (i.e. the field is evanescent) while in the region \(\text{Re}(\lambda_p)R > \lambda\), the field is oscillatory with an exponential growth. The turning point between these two regions,

\[
 r_{tp} = \frac{R \rho}{\text{Re}(\lambda_p)} \quad (3.23)
\]

is the radius at which the phase velocity of the radially evanescent field equals that of an infinite plane wave in the cladding medium and is therefore the radius at which "leakage" from the mode appears to originate.\(^{32}\) Clearly bound modes have \(r_{tp} = \infty\) and \(r_{tp}\) decreases as the mode becomes more strongly leaky. All tunnelling modes have \(r_{tp} > \rho\) and refracting modes have \(r_{tp} = \rho\).

The physical reason for the growth of the field in the radiation region is indicated in Fig. 4.4. The radiation field at \(r_1\)
is greater than at \( r_2 \) because the leaky mode has lost energy in travelling from \( z_1 \) to \( z_2 \). Since the modes are derived from an infinite waveguide, the growth of the radiation field continues out to infinity. However on any finite waveguide, it is clear from Fig. 4.4 that the field intensity will increase with \( r \) up to a maximum intensity related to the location of the source and then be identically zero.\(^{10}\)

Mathematically, this "cut-out" occurs when the steepest-descent path in Fig. 4.3 lies to the left of the relevant leaky-mode pole so that its residue does not contribute to the field. The apparently unphysical behaviour of the leaky modes is therefore no obstacle to their use in expansions of the waveguide field.
We now turn our attention to the second term in Eq. (2.27), the "space wave".

4.4 SADDLE-POINT CONTRIBUTION

We must consider the integrals

\[ I_\ell(\zeta, \theta) = \int_{SDP} F_\ell(\psi) e^{q(\psi)\zeta} d\psi. \]  

(4.1)

which appear in Eq. (2.27) with

\[ F_\ell(\psi) = \rho k_2 \sin \psi f_\ell(\lambda) \]  

(4.2)

and

\[ q(\psi) = -ik_2 \cos(\psi - \theta). \]  

(4.3)

The quantities \( \zeta \) and \( \theta \) are defined by Eqs. (2.23 - 24) and the integration path SDP by Eq. (2.25).

The factor \( e^{q(\psi)\zeta} \) has a maximum value at the saddle point \( \psi = \theta \) and decreases exponentially away from this point. So provided \( F_\ell(\psi) \) is regular and slowly varying in the vicinity of \( \theta \), an approximation to the integral can be obtained by expanding \( F_\ell(\psi) \) in a power series about \( \theta \) and then integrating term by term.

The result obtained is

\[ I_\ell(\zeta, \theta) \sim \left( \frac{2\pi i}{k_2 \zeta} \right)^{1/2} e^{-ik_2\zeta} \left[ F_\ell(\theta) + \frac{i}{2k_2\zeta} \right. \]

\[ \times \left. \left[ \frac{1}{4} F''_\ell(\theta) \right] + O\left( (k_2 \zeta)^{-5/2} \right) \right]. \]  

(4.4)

The analytic calculation of \( F_\ell(\theta) \) and its derivatives is rather difficult but we can estimate the region in which the leaky modes dominate the saddle-point contribution to the field within the waveguide.
without actually performing these calculations. The approach is one employed by Tamir and Oliner\textsuperscript{14} for the case of the plane interface.

Ignoring the relative excitation of the leaky mode and space wave, we can consider the variation of this intensity ratio with axial distance from the origin. This is given by the function

\[
H_1(\gamma, k_2 z) = (k_2 z) e^{-\gamma k_2 z}
\]

(or \(H_2(\gamma, k_2 z) = (k_2 z)^3 e^{-\gamma k_2 z}\) if \(\xi L_2 = 0\)),

where

\[
\gamma = \frac{\alpha}{\rho k_2},
\]

and

\[
\alpha = -2\rho_1.
\]

\(\alpha\) being the normalized modal attenuation coefficient.

It is clear that eventually \(H_1\) (and \(H_2\)) will approach zero indicating that the space wave dominates the radiation field. But whereas in the case of the plane interface the distance involved is quite small, in a multimode, circular optical fibre, the distance at which \(H_1\) falls below ten is very often so large that the space wave itself is by then negligible. So except at very short distances from the source, any significant radiation field within the optical fibre will be dominated by leaky modes.

For example, any leaky mode which has a value of \(\gamma\) less than \(4.6 \times 10^{-3}\) will give \(H_1 > 10\) even after the space wave has fallen to less than 0.1% of its original value. On a typical multimode optical fibre with \(\rho k_2 > 100\), this means that any mode with an attenuation coefficient less than 0.5 will completely dominate the space wave. Even the
Fig. 4.5: Plot of $F_1 = \log_{10} H_1$ and $F_2 = \log_{10} H_2$ against $\log_{10}(kz)$ where $H_1, H_2$ are defined by Eqs. (4.5). The figure between the curves for $F_1$ and $F_2$ in each pair gives the value of $n$ where $y = 10^{-n}$.

relatively strongly leaky, low-order modes have losses below this value well below cutoff.

In Fig. 4.5 we plot, for illustrative purposes, $F_1 = \log_{10}(H_1)$ and $F_2 = \log_{10}(H_2)$ against $\log_{10}(kz)$ for various values of $y = 10^{-n}$, $n = 3, 5, \ldots, 13$. On a fibre with $V = 50$, $\delta = 0.1$ and $\rho = 100 \mu$, this corresponds to a range of mode attenuations between $5 \times 10^{-4}$ and $5 \times 10^6$ dB/km so that the distance scale is rather exaggerated. But as we shall see in the following chapter, even at $V = 20$, there exists modes with attenuations approaching the lower limit here so that the message
conveyed by Fig. 4.5 is that for all intents and purposes, many leaky modes are indistinguishable from bound modes in practical situations.

The situation becomes a little more complicated if there is a pole, either proper or improper, at $\psi_p$ near the saddle point because then Eq. (4.4) must be modified to take into account the fact that $F_p(\psi)$ is no longer slowly varying in the vicinity of $\theta$. The modification is carried out by adding and subtracting the singularity from $F_p(\psi)$ and then writing $I_p$ as two integrals, one of which has its integrand analytic at $\psi_p$ while the other contains the effect of the singularity. The former is expanded in an ordinary saddle-point expansion as in Eq. (4.4) and the latter is evaluated in closed form in terms of the complementary error function.

Putting

$$a = \lim_{\psi \to \psi_p} \left[ (\psi - \psi_p) F_p(\psi) \right]$$

(4.7)

and

$$b = \text{residue of } F_p(\psi) \text{ at } \psi_p$$

one finds

$$I_p(\zeta, \theta) \sim e^{-ik_2\zeta} \left\{ \pi a \exp[ik_2(1 - \cos(\theta - \psi_p))] \text{erfc}(-ib\zeta) \right.$$  

$$+ \left[ \frac{\pi}{i} \right]^{1/2} \left[ F_p(\theta) \left\{ \frac{2i}{k_2} \right\}^{1/2} + \frac{a}{b} \right]$$  

$$+ \frac{i}{2} \left[ \frac{\pi}{i} \right]^{3/2} \left[ \left\{ F_p''(\theta) + F_p''(\theta) \right\} \left[ \frac{2i}{k_2} \right]^{1/2} + \frac{2a}{b^3} \right] \right\} + \ldots ,$$

(4.9)

where

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} \, dt .$$

(4.10)

Examination of Eq. (4.9) indicates that a pole is "near" the saddle point, in the sense that Eq. (4.4) must be modified, if...
\[ |b\sqrt{c}| < \sqrt{10}. \] (4.11)

For if condition (4.11) is not satisfied, one can use the asymptotic expression

\[
\text{erfc}(x) \sim \frac{e^{-x^2}}{x\sqrt{\pi}} \left[ 1 - \frac{1}{2x^2} + \frac{1.3}{(2x^2)^2} - \frac{1.3.5}{(2x^2)^3} + \ldots \right] \quad (4.12)
\]

valid when \(-\frac{3\pi}{4} < \text{arg} x < \frac{3\pi}{4}\) to reduce Eq. (4.9) identically to Eq. (4.4).

The magnitude of the correction given by Eq. (4.9) to the simple saddle-point calculation when condition (4.11) does hold has been investigated for the plane interface by Bernard and Ishimaru. In Section 5.4 we show by numerical example that in practical situations it is, in fact, quite small. Even in this case therefore, after a relatively short distance, the saddle-point contribution can be justly ignored and Eq. (2.27) can be simplified to

\[
E \approx \sum_p a_p e^{i\beta_p z} \quad (4.13)
\]

where the sum extends over all modes, bound or leaky, which are crossed in transforming to the steepest-descent path. For the field inside the waveguide, this means all forward bound and leaky modes. \(\beta_p\) is real for bound modes and has negative imaginary part for leaky modes.

4.5 ORTHOGONALITY AND EXCITATION OF LEAKY MODES

The advantages of using a leaky-mode approximation to the integral over the continuous spectrum might be questioned if it were always necessary to evaluate the leaky-mode amplitudes by performing a lengthy residue calculation.
In the case of the bound modes, such a calculation is avoided by using the power orthonormality condition, Eq. (2.2.26), to evaluate the modal amplitudes directly. Because leaky modes do not satisfy the usual radiation condition (and indeed diverge) at infinity, they cannot obey the orthonormality condition — the integrals diverge. The leaky mode amplitude is not, therefore, a straight-forward extension of Eq. (3.A.1). This difficulty can be circumvented in one of two ways.

4.5.1 Truncated Leaky Modes and Coupling

A simple, heuristic approach is to define a new set of modes which have the same form as the original leaky modes within a restricted region of space and which are approximately power orthogonal. The small amount of mode coupling caused by the non-orthogonality can then be determined by an analysis identical to that used for proper modes in an imperfect waveguide.46,47

The usefulness of this method depends on the assumption that weakly leaky modes approximate the radiation field within the waveguide. Far enough away from the source when this assumption is valid, the weakly leaky modes are treated as simply an extension to the bound mode set.

The new modes we choose take advantage of the fact that we are only interested in the "guided" portion of the field. Thus we define truncated leaky modes \( \tilde{e}_p \) by

\[
\tilde{e}_p = e_p \quad r \leq r_{tp}^p \\
= 0 \quad r > r_{tp}^p ,
\]

(5.1)

where \( e_p \) is the analytic continuation of the \( z \)-independent part of a
bound mode field below its cutoff frequency and \( r_{tp}^P \) is the turning point for mode \( p \). \( \tilde{h}_p \) is defined similarly. For weakly leaky modes, recall that \( r_{tp}^P \gg \rho \).

At a sufficiently large distance along the fibre we now rewrite Eq. (4.13) in the form

\[
E(x,y,z) \approx \sum_{p} a_p(z) \tilde{e}_p(x,y),
\]

where the sum extends over both bound and weakly leaky modes.

Because the leaky modes in Eq. (5.1) are not really modes of the system, each having been arbitrarily truncated at a different point, there will be some power coupling between them (and between bound and truncated leaky modes). The basis of this approach, however, is that the weakly leaky modes behave essentially like proper modes on a slightly imperfect waveguide where the power exchange between modes is very small and can, to a very good approximation be neglected. Corrections to the amplitude coefficients \( a_p(z) \) can then be found by standard coupled-mode theory.

To establish these results, we define

\[
\tilde{c}_{pq} = \int_{S_{tp}} \tilde{e}_p \times \tilde{h}_q \cdot \hat{z} \, da \tag{5.3}
\]

and

\[
L_{pq} = \oint_{\partial S_{tp}} \tilde{e}_p \times \tilde{h}_q \cdot \hat{z} \, dl \tag{5.4}
\]

where \( S_{tp} \) is the cross-sectional area defined by \( r < r_{min}^P \) (\( r_{min}^P \) being the smaller of \( r_{tp}^P \) and \( r_{tp}^Q \)) and \( \partial S_{tp} \) is the boundary of \( S_{tp} \). Without loss of generality, we will assume \( r_{tp}^{min} = r_{tp}^P \) throughout the remainder of this section.
4.5.1

An analysis following that of Ref. 46 shows that the expansion coefficients $a_q(z)$ satisfy

$$
\sum_q \left( \tilde{c}_{pq} + \tilde{c}_{qp}^* \right) \left( \frac{da_q^*}{dz} - i\beta a_q^* \right) = -\sum_q' \left( L_{pq'} + L_{qp'}^* \right) a_{q'}^* .
$$

(5.5)

Dividing both sides of this equation by $\tilde{c}_{pp}$ and rearranging terms we arrive at the matrix equation

$$
\Lambda \frac{da^*}{dz} = \mathbf{B} a^* ,
$$

(5.6)

where $\Lambda, \mathbf{B}$ represent matrices whose elements $\Lambda_{pq}, \mathbf{B}_{pq}$ are given by

$$
\Lambda_{pq} = \frac{\left( \tilde{c}_{pq} + \tilde{c}_{qp}^* \right)}{2\tilde{c}_{pp}}
$$

(5.7)

and

$$
\mathbf{B}_{pq} = -\left( L_{pq} + L_{qp}^* - i\beta \left( \tilde{c}_{pq} + \tilde{c}_{qp}^* \right) \right) \frac{1}{2\tilde{c}_{pp}} .
$$

(5.8)

The system of equations, Eq. (5.6), is in general difficult to solve. However, more detailed consideration of the coefficients $\tilde{c}_{pq}$ shows that some simplifying assumptions can be made in the case of the circular optical fibre.

If the electric field in Eq. (5.2) is taken to be

$$
\tilde{E}_q(x,y,z) = \tilde{E}_q(x,y) e^{-i\beta z} ,
$$

(5.9)

then Eq. (5.5) reduces to

$$
\left( \beta_p^* - \beta_q \right) \left( \tilde{c}_{pq} + \tilde{c}_{qp}^* \right) = i \left( L_{pq} + L_{qp}^* \right) .
$$

(5.10)

Similarly if we replace $E_q$ by $\sim q$, a backward travelling mode, we find

$$
\left( \beta_p + \beta_q^* \right) \left( \tilde{c}_{pq} - \tilde{c}_{qp}^* \right) = i \left( L_{-p,-q} + L_{-p,-q}^* \right) ,
$$

(5.11)

where we have used the relationships

41
\[ \beta_{-q} = -\beta_q, \]

\[ e^{-qt} = e^{qt}, \quad e^{-qz} = -e^{qz}, \]

\[ h^{-qt} = -h^{qt}, \quad h^{-qz} = h^{qz}. \]  

(5.12)

Combining Eqs. (5.10-11),

\[ \frac{2c}{pq} = \frac{i}{\beta_q - \beta_p} \left[ \frac{L_{pq} + L^*_{qp}}{\beta_q - \beta_p} \right] \frac{i}{\beta_q - \beta_p} \left[ \frac{L_{pq} - qL^*_{qp}}{\beta_q + \beta_p} \right]. \]  

(5.13)

So far the calculation of \( \tilde{c}_{pq} \) has been completely general. If we now restrict ourselves to the circular dielectric rod and, in particular, to the small \( \delta, \beta \approx k \) modes to obtain an estimate of \( c_{pq} \), then the field components required (for \( R > 1 \)) are given by (Section 2.2)

\[ e_{sz}(R) = \left( \frac{\mu}{c} \right)^{\frac{1}{2}} i U_s \lambda_s \frac{H_{\lambda+2}(\lambda R)}{H_{\lambda}(\lambda s)} \left( \frac{H_{\lambda+2}(\lambda s)}{H_{\lambda}(\lambda s)} \right)^{\frac{1}{2}}, \]  

(5.14)

\[ e_{sr}(R) = \left( \frac{\mu}{c} \right)^{\frac{1}{2}} \frac{U_s}{V(\mu \rho^2)} \frac{H_{\lambda+1}(\lambda R)}{H_{\lambda}(\lambda s)} \left( \frac{H_{\lambda+1}(\lambda s)}{H_{\lambda}(\lambda s)} \right)^{\frac{1}{2}}, \]  

(5.15)

and

\[ h_{ks} = \pm \left( \frac{\mu}{c} \right)^{\frac{1}{2}} \zeta_s \times \zeta_s. \]  

(5.16)

where \( s \) stands for either \( p \) or \( q \) and all other quantities are as previously defined.

Invoking Eq. (5.16) simplifies Eq. (5.13) to

\[ \tilde{c}_{pq} = \frac{i\beta_p \left[ L_{pq} + L^*_{qp} \right]}{\beta_q^2 - \beta_p^2} \]

\[ = \frac{2\pi i \beta_p}{\beta_q^2 - \beta_p^2} \left[ \left( \tilde{e}_p \times \tilde{h}_{p} + \tilde{e}_q \times \tilde{h}_{q} \right) \cdot \tilde{h}_r \right] \cdot \tilde{r} = \beta_{tp}. \]  

(5.17)

Substituting Eqs. (5.14-16) into Eq. (5.17) we find
\[
\frac{\partial q}{\partial p} - \frac{\partial p}{\partial q} = \left\{ \frac{\beta_p^2 - \beta_q^2}{\beta_q^2 - \beta_p^2} \right\} \frac{U_p H^{(1)}_\ell (\lambda_p, \lambda_q)}{U_q H^{(1)}_\ell (\lambda_q, \lambda_p)} \left( \frac{H^{(1)}_\ell (\lambda_p) H^{(1)}_\ell (\lambda_q)}{H^{(2)}_\ell (\lambda_p) H^{(2)}_\ell (\lambda_q)} \right)^{\ell^2}
\times \left\{ \frac{\lambda_p H^{(2)}_\ell (\lambda_p, \lambda_q)}{\lambda_q H^{(2)}_\ell (\lambda_q, \lambda_p)} + \lambda_q H^{(1)}_\ell (\lambda_q, \lambda_p) \right\}
\right\} \quad , \quad (5.18)
\]

where \( R_{tp} = \frac{r_{tp}}{\rho}. \)

It can now be shown, after some cumbersome algebra involving the Debye forms of the Hankel functions that, provided \( \ell \gg 1, \)
\( \ell - \lambda_p \gg \ell^{1/3} \text{ and } \ell^{2/3} \left[ 1 - \lambda_q / \lambda_p \right] \gg 1 \) (with \( \lambda_p, \lambda_q \) almost real),
\[
\left| \frac{\partial q}{\partial p} \right| \sim \left| \frac{U_p (\beta_p^2 - \beta_q^2)}{U_q (\beta_q^2 - \beta_p^2)} \right| \cdot \frac{3}{2a(\pi)} \cdot \frac{1}{\ell^{1/3}} e^{\frac{-\ell^2}{\lambda_q}} \left[ 1 + \left( 1 - \frac{\lambda_q^2}{\ell^2} \right) \right] \quad , \quad (5.19)
\]

where
\[
a = \frac{\Gamma(1/3)}{2^{2/3} \cdot 3^{1/6}} .
\]

Thus, provided both modes \( p \) and \( q \) are weakly leaky and have different turning points, \( \left| \frac{\partial q}{\partial p} \right| \) behaves asymptotically as \( e^{-\ell^2} \) for large \( \ell. \)

It can similarly be shown that, even if \( \ell \) is no longer large but \( \lambda_p \) (and hence \( \lambda_q \)) approaches zero with \( \lambda_q (\ell - 1) << 1, \) then
\[
\left| \frac{\partial q}{\partial p} \right| \sim \left| \frac{U_p (\beta_p^2 - \beta_q^2)}{U_q (\beta_q^2 - \beta_p^2)} \right| \cdot \frac{1}{(\ell - 1)^{1/3}} \left| \frac{H^{(2)}_\ell (\lambda_p)}{H^{(2)}_\ell (\lambda_q)} \right| \left( \frac{1}{\ell - 1} \right) \left( \frac{H^{(2)}_\ell (\lambda_p) H^{(2)}_\ell (\lambda_q)}{H^{(2)}_\ell (\lambda_p) H^{(2)}_\ell (\lambda_q)} \right)^{\ell^2} \quad , \quad (5.20)
\]
where $A(l)$ is a well-behaved function of $l$. But as $\lambda_p$ approaches zero, the imaginary part of $\beta_p$ also approaches zero so that again

$$\left| \frac{\tilde{C}_{pq}}{\tilde{C}_{pp}} \right| \rightarrow 0 \quad \text{when} \quad p \neq q .$$

The fact that $|\tilde{C}_{pq}| \ll |\tilde{C}_{pp}|$ for weakly leaky modes now enables us to introduce a simplification into Eq. (5.6). We write

$$A \approx I + \Xi ,$$

(5.21)

where $I$ is the unit matrix and $\Xi$ is a matrix whose elements are much less than 1. Then $A^{-1} \approx I - \Xi$ so that Eq. (5.6) can be rewritten as

$$\frac{d a^*}{dz} = (I - \Xi) B a^* .$$

(5.22)

With the approximate initial condition

$$a_p(0) = \frac{\int_{S_{\infty}} E \times \tilde{h}^*_p \cdot \tilde{Z} \, da}{\int_{S_{\infty}} \tilde{C}_p \times \tilde{h}^*_p \cdot \tilde{Z} \, da} ,$$

(5.23)

the problem is now in the standard coupled mode formulation and can be dealt with using the methods described in Refs. 46 and 47 for slightly absorbing waveguides.

### 4.5.2 An Exact Orthogonality Condition

An alternative to the heuristic approach presented above is to retain the fully leaky mode fields but to replace the power orthogonality relations by a more general form of orthogonality.

This is the approach followed by Shevchenko$^5$ in dealing with the scalar fields of the slab waveguide and by Budden$^4$ in his treatment of radio wave transmission. This generalized orthogonality has the
advantage of being mathematically exact but suffers from the
disadvantage that the physically useful idea of power orthogonality
between unlike modes is lost.

Assuming an $e^{i\omega t}$ time dependence, the fields of a cylindrical
waveguide behave, for large $r$, as

$$r^{-\frac{1}{2}} \exp[-i(\beta z + \lambda \phi + \gamma r)] ,$$

(5.24)

where for bound modes $\lambda = -i\omega$ ($\omega > 0$) and for improper modes $\lambda = \lambda' + i\lambda''$ ($\lambda'' > 0$). It is clear, therefore, that the improper modes are not
square-integrable with respect to the real variable $r$. The approach
taken by Shevchenko and Budden to overcome this problem for the slab
waveguide (and which will presently be extended to the general
cylindrical waveguide) is to deform the contour of integration into the
complex $r$ plane in such a way that for large $|r|$, the decay due to the
imaginary part or $r$ compensates for the growth caused by the positive
imaginary part of $\lambda$.

Consider the integral

$$I(k,m;\lambda,n) = \int_{\Lambda_0} \int e_{km} \times h_{\lambda n} \cdot \mathbf{\tilde{z}} \, da ,$$

(5.25)

where the fields (evaluated at $z = 0$) belong to either trapped or leaky
modes, $k, \lambda$ are mode numbers associated with azimuthal variation, $m, n$ are
associated with radial variation and $\Lambda_0$ is a surface to be determined
below. We write

$$e_{km} \times h_{\lambda n} \cdot \mathbf{\tilde{z}} = f^{mn}_{k\lambda}(r) e^{-i(k+\lambda)\phi} .$$

(5.26)
Then

\[ I(k,m;\ell,n) = \int_{\Gamma} \int_{0}^{2\pi} f_{\ell,\ell}(r) e^{-i(k+\ell)\phi} r \, dr \, d\phi \]

\[ = 2\pi I(m,n) \delta_{k,-\ell}, \quad (5.27) \]

where \( \Gamma \) is a contour in the \( r \)-plane and

\[ I(m,n) = \int_{\Gamma} f_{-\ell,\ell}(r) r \, dr. \quad (5.28) \]

The contour \( \Gamma \) can now be chosen. Put

\[ I(m,n) = \int_{d}^{d} f_{-\ell,\ell}(r) r \, dr + \int_{L_0} f_{-\ell,\ell}(r) r \, dr, \quad (5.29) \]

where \( d > \rho \) and \( L_0 \) is chosen to be a line in the complex \( r \)-plane from \( r = d \) to \( r = d + \rho e^{i\gamma} \) (see Fig. 4.6) such that

\[ (\lambda_m^n + \lambda_m^n) + (\lambda_m^n + \lambda_m^n) \tan \gamma < 0, \]

i.e.

\[ -\frac{\pi}{2} < \gamma < -\tan^{-1}\left(\frac{\lambda_m^n + \lambda_m^n}{\lambda_m^n + \lambda_m^n}\right). \quad (5.30) \]
Then because \( f_{\ell,m}^{{\text{mn}}}(r) \) behaves asymptotically as \( r^{-1} \exp(-i(\lambda_m + \lambda_n)r) \),
the contribution to the integral from the upper limit approaches zero as \( \sigma \to \infty \).

The orthonormality condition

\[
I(m,n) = \frac{1}{2\pi} \delta_{m,n} \quad (5.31)
\]

then follows as for the proper modes.

Thus, provided \( A_0 \) is interpreted as above, putting

\[ A_\infty = A_0(\sigma \to \infty) \]

gives the orthonormality condition

\[
\int_{A_\infty} \int e_{k_m} \times h_{k_n} \cdot \hat{z} \; da = \delta_{k_m,k_n} \delta_{m,n} \quad (5.32)
\]

for both bound and leaky modes.

With \( d \) chosen so that the source field is contained within the region \( r \leq d \), the amplitude coefficients for both bound and leaky modes in Eq. (4.13) are given by

\[
a_p = \frac{\int_{A_\infty} \int E \times h_{k_m} \cdot \hat{z} \; da}{\int_{A_\infty} \int e_{k_m} \times h_{k_m} \cdot \hat{z} \; da} \quad (5.33)
\]

where the label \( p \) stands for the pair of mode numbers \((\ell,m)\).

This method has thus provided an orthogonality condition which is satisfied exactly and in which bound and leaky modes have formally identical expansion coefficients. It is also a condition applicable to a general cylindrical waveguide, independent of the specific refractive index profile. However the derivation of the expansion coefficients loses the simple interpretation of power orthogonality between dissimilar modes. And while the heuristic method seems more complicated,
it must be emphasized that the leaky mode expansion is only of practical use when weakly leaky modes dominate the unbound field in which case the truncated leaky modes are, to a good approximation, power orthogonal so that coupling only provides a small correction.

4.6 RÉSUMÉ

In this chapter, we have found an asymptotic representation of the dielectric waveguide field in which a major role is played by improper, leaky modes — modes which correspond to complex resonances of the waveguide system. The improper modal fields decay in the axial direction but, because they represent radiation from a source at $z = -\infty$, oscillate and grow exponentially in the transverse cross-section. They can therefore only be used within restricted regions of space — but within their range of validity, they provide a rapidly convergent representation of the radiation field, the portion of this field not covered by the leaky modes being very small.

We have also investigated methods of orthonormalizing these improper modes and hence of deriving their amplitude coefficients without recourse to the continuous mode spectrum. This can be achieved either by applying the power orthogonality condition to truncated modes and then obtaining corrections due to coupling between these modified modes or by abandoning power orthogonality in favour of a more general orthogonality condition which holds exactly for all modes but loses the physical significance of power orthogonality.
REFERENCES


4. V.V. Shevchenko, Continuous Transitions in Open Waveguides (Golem Press, Boulder, Colorado, 1971).


40. Dr. D. Marcuse suggested this approach and performed the slab calculation.


5.1 INTRODUCTION

Having examined the origin and behaviour of leaky modes in general terms, we must now study the detailed characteristics of these modes as a prelude to their use in specific excitation problems. The basic quantity to be determined is the transverse propagation constant in the core, $U$.

As has been mentioned a number of times above, the allowed values of $U$ are found as the zeros of the transverse resonance relation (or eigenvalue equation)

$$ M(U) = \frac{H_2^2(\lambda)}{\lambda H_2^2(\lambda)} - \frac{\mu_1}{\mu_2} \frac{J_\lambda'(U)}{U J_\lambda(U)} \left( \frac{H_2^2(\lambda)}{\lambda H_2^2(\lambda)} - \frac{\varepsilon_1}{\varepsilon_2} \frac{J_\lambda'(U)}{U J_\lambda(U)} \right) $$

$$ - \left( \frac{\psi}{U \lambda} \right)^4 \left( \frac{\rho}{k_2} \right)^2 = 0, \hspace{1cm} (1.1) $$

where $U$ and $\lambda$ are related by

$$ \psi^2 = U^2 - \lambda^2. \hspace{1cm} (1.2) $$

and $M$ is now explicitly considered a function of $U$ rather than $\lambda$.

For bound modes, $\lambda = -iW$ where $W$ is real and positive so that the cladding field has a $K_\lambda(WR)$ radial dependence. The value of $W$ at which a bound mode is cut off (i.e. at which $W > 0$) is given by $^1$
\[ J_0(V) = 0 \quad \text{for TE}_{0m}, \text{TM}_{0m} \text{ modes} \quad (1.3) \]

\[ J_\ell(V) = 0, \quad V > 0 \quad \text{for EH}_{\ell m} \text{ modes} \quad (1.4) \]

\[ J_1(V) = 0 \quad \text{for HE}_{1m} \text{ modes} \quad (1.5) \]

and

\[ J_{\ell-2}(V) = -\frac{\delta}{2-\delta} J_\ell(V), \quad V > 0 \quad \text{for HE}_{\ell m} \text{ modes} (\ell \geq 2), (1.6) \]

where the subscript \( m \) orders the infinity of solutions of each of Eqs. (1.3-6) for a particular value of \( \ell \). These mode labels are omitted throughout this chapter except where required to avoid ambiguity.

Equation (1.1) can be simplified considerably in the case where \( \mu_1 = \mu_2 \) and \( \varepsilon_1 = \varepsilon_2 \). We put

\[ M_1 = H_\ell - J_\ell - \ell \left( \frac{V}{U\lambda} \right)^2 \quad (1.7) \]

\[ M_2 = M_1 - \frac{\delta}{1-\delta} \left( J_\ell - \frac{\ell}{U^2} \right) \quad (1.8) \]

and

\[ M_3 = \frac{\ell^2}{1-\delta} \left( \frac{V}{U\lambda} \right)^4 \left( 1 - \frac{U^2\delta}{V^2} \right) \quad (1.9) \]

where \( H_\ell \) and \( J_\ell \) are defined by Eq. (4.3.4). Then

\[ M(U) = M_1M_2 - M_3 \]

\[ = \left\{ M_1 - \ell^2 \left( \frac{V}{U\lambda} \right)^4 \right\} - \frac{\delta}{1-\delta} \left\{ \left( M_1 + \ell \left( \frac{V}{U\lambda} \right)^2 \right) \left( J_\ell - \frac{\ell}{U^2} \right) - \ell J_\ell \left( \frac{V}{U\lambda} \right)^2 \right\} \quad (1.10) \]

Thus in the limit \( \delta \to 0 \), we have the two equations

\[ M_1 = \ell \left( \frac{V}{U\lambda} \right)^2 \]

or

\[ H_\ell - J_\ell = 0, \quad \text{HE}_{\ell m} \text{ modes} \quad (1.11) \]

\[ = 2\ell \left( \frac{V}{U\lambda} \right)^2, \quad \text{EH}_{\ell m} \text{ modes} \quad (1.12) \]
Cutoff condition Eq. (1.6) for the $HE_{\lambda m}$ mode also simplifies to

$$J_{\lambda-2}(V) = 0.$$  (1.13)

Examination of Eqs. (1.4, 11-13) then shows that in this limit, the $HE_{\lambda+2,m}$ and $EH_{\lambda m}$ modes are degenerate.¹

In the following sections we shall begin by looking at the numerical and asymptotic solution of the simplified eigenvalue equation, Eq. (1.11), below cutoff and then proceed to a study of the leaky-mode solutions of the exact equation, Eq. (1.1). This study supplements the large volume of work which has already been carried out on the solution of the corresponding eigenvalue equations for a number of dielectric structures, but which has been concerned predominantly with the bound modes. Some examples are the works dealing with

(a) bound modes of slab²-⁴ and rectangular⁵,⁶ waveguides,

(b) leaky modes of slab waveguides,³,⁷-¹⁰

(c) bound modes of homogeneous circular fibres¹¹-¹⁴ (in infinite cladding),

(d) bound modes of doubly-clad circular fibres with low-index outer cladding,¹⁵-¹⁷

(e) bound modes of radially inhomogeneous circular fibres,¹⁸-²²

(f) bound and leaky modes of doubly-clad circular fibres with low-index inner cladding,²³-²⁵

(g) leaky modes of homogeneous circular fibres.²⁶-³⁰

The leaky modes mentioned under (f) above are not of the same type as those slowly radiating modes with which we are concerned here. If their refractive indices are $n_1$ (core), $n_2$ (inner cladding) and $n_3$
(outer cladding), then in the case \( n_3 > n_1 > n_2 \), these structures behave like a hollow waveguide so there are no bound modes and all modes are "refracting". In the case \( n_1 > n_3 > n_2 \), modes with longitudinal propagation constant, \( \beta \), between \( k_1 \) and \( k_3 \) are bound while those with \( \beta \) between \( k_3 \) and \( k_2 \) are refracting (\( k_1, k_2, k_3 \) being the wavenumbers in the core and cladding media).

Of the papers concerned with the leaky modes on an infinitely-clad dielectric rod, ((g) above), only Refs. 28-30 are relevant here. Arnbak\(^26\) solves Eq. (1.1) numerically for the circularly symmetric \( \text{TE}_{0m} \) and \( \text{TM}_{0m} \) modes for a large-\( \phi \) fibre (\( \varepsilon_1 = 15 \varepsilon_2 \)) which are very rapidly leaky and do not display any of the important features of the weakly-leaky tunnelling modes while James\(^27\) is concerned with the symmetry of the positions of zeros of Eq. (1.1) on the \( \lambda \)-surface without reference to the modes themselves. The papers of Snyder and Mitchell\(^28,29\) derive analytic approximations to the solution of Eq. (1.11) for very weakly leaky modes and Ref. 30 is a preliminary version of parts of this chapter.

For any of the above structures, once the eigenvalue equation has been solved, one knows immediately the eigenfunctions, or modal field distribution along the waveguide (Eqs. (4.3.11-16)) and other modal characteristics such as the longitudinal propagation constant, \( \beta \) (Eq. (2.2.8)), the group velocity\(^31\) and, in the case of leaky modes, the attenuation coefficient, \( \alpha \) (Eq. (4.4.6)).

Another parameter which has proven very useful in the study of bound modes is \( \eta \), the fraction of the modal power which propagates within the core.\(^{13,14,16}\) For small \( \delta \), this is given by\(^{13}\)
\[ n_p = \left( \frac{\xi_p}{v} \right)^2 \left( \frac{1}{\xi_p^2} - \frac{\lambda_p^2}{v^2} \right), \]  

(1.14)

where the subscript \( p \) refers to a particular mode, and \( \xi_p \) is defined by Eq. (2.2.35) or, in terms of \( \lambda_p \),

\[ \xi_p = \frac{H_{\lambda-2}^{(2)}(\lambda_p)}{H_{\lambda-2}^{(2)}(\lambda_p)} \cdot \frac{H_{\lambda}^{(2)}(\lambda_p)}{H_{\lambda-1}^{(2)}(\lambda_p)} \]  

(1.15)

While Eq. (1.14) can be continued below cutoff, it is clear that the physical interpretation cannot be strictly the same for leaky modes because in that case the total modal power is infinite. A generalized \( n \) is therefore introduced for modes below cutoff which represents the fraction of a mode's total guided power within the core. That is, we define

\[ n_p = \frac{\text{Power of mode } p \text{ in } r < p}{\text{Power of mode } p \text{ in } r < r_{tp}} \]

\[ = \frac{\int_{S_F} \text{ } \hat{e}_p \times \hat{h}_p^* \cdot \hat{z} \text{ } da}{\int_{S_{tp}} \text{ } \hat{e}_p \times \hat{h}_p^* \cdot \hat{z} \text{ } da} \]  

(1.16)

or equivalently,

\[ \eta_p = \frac{\int_{S_F} \text{ } \hat{e}_p \times \hat{h}_p^* \cdot \hat{z} \text{ } da}{\int_{S_{\infty}} \text{ } \hat{e}_p \times \hat{h}_p^* \cdot \hat{z} \text{ } da} \]  

(1.17)

where \( r_{tp} \) is defined by Eq. (4.3.23), \( S_F \) is the core cross-section, \( S_{tp} \) is the cross-sectional disc of radius \( r_{tp} \), \( S_{\infty} \) is the infinite cross-section and the fields in Eq. (1.17) are the truncated modes defined in Section 4.5.
For bound modes Eqs. (1.14) and (1.17) are identical since then \( r_p = 0 \). For refracting modes, \( \eta_p \) is identically one since the turning point lies on the core-cladding boundary; but for tunnelling modes, we shall see in Section 5.2 (by numerical example) that Eq. (1.16) gives results in very close agreement with the analytic continuation of Eq. (1.14) below cutoff.

Asymptotic expressions for \( \eta \) as well as the other characteristics of weakly-leaky and high-order modes are also derived in Section 5.2 while Section 5.3 contains detailed numerical results from the exact eigenvalue equation.

In Section 5.4, we use the characteristics found in the preceding sections to investigate the excitation and propagation of leaky modes using the two source fields introduced in Chapter 3. The results of this analysis are then compared with corresponding results found using the leaky ray theory\(^{32}\) and the range of validity of the latter approximation is discussed.

Sections 5.5 and 5.6, respectively, are concerned with the number of leaky modes with a given attenuation rate and the influence of material absorption on this attenuation.

5.2 APPROXIMATE EIGENVALUE EQUATION\(^{30}\)

5.2.1 Numerical Solution

The initial difficulty in solving Eq. (1.11) below cutoff numerically is that of finding starting values for the imaginary part of the eigenvalue, \( U \). Except in the case of the \( HE_{lm} \) modes which cease to exist in an interval immediately below cutoff,\(^{29}\) this imaginary part is
very small near cutoff and so a purely numerical search is difficult.

One can overcome this problem by using perturbation theory.

We define

\[ G_{\ell}(U) = \frac{UJ_{\ell}(U)}{J_{\ell-1}(U)} - \frac{\lambda H_{\ell}(2)(\lambda)}{H_{\ell-1}(\lambda)} \]  

\[ = G_{\ell_{\text{re}}}(U) + iG_{\ell_{\text{im}}}(U) \]  

where \( G_{\ell_{\text{re}}}, G_{\ell_{\text{im}}} \) are the real and imaginary parts of \( G_{\ell} \), respectively.

Now suppose \( G_{\ell_{\text{re}}} \) has a zero at the real point \( U_0 \). Making a Taylor expansion about this point gives

\[ G_{\ell}(U) \approx G_{\ell}(U_0) + \Delta U G_{\ell}'(U_0) = 0, \]  

where \( U = U_0 + \Delta U, \Delta U = \Delta U_{\text{re}} + i\Delta U_{\text{im}} \) and the prime indicates differentiation with respect to \( U \).

Equating real and imaginary parts in Eq. (2.3) and recalling that \( G_{\ell_{\text{re}}}(U_0) = 0 \), we find

\[ \Delta U_{\text{re}} = -\frac{G_{\ell_{\text{im}}}(U_0) G_{\ell_{\text{im}}}'(U_0)}{|G_{\ell}'(U_0)|^2} \]  

and

\[ \Delta U_{\text{im}} = -\frac{G_{\ell_{\text{re}}}(U_0) G_{\ell_{\text{re}}}'(U_0)}{|G_{\ell}'(U_0)|^2}. \]  

Since, near cutoff, \( U_0 \approx V \), we can easily solve the equation \( G_{\ell_{\text{re}}}(U_0) = 0 \) and hence find starting values for the solution of Eq. (1.11) via Eqs. (2.4 and 2.5). This initial approximation to the eigenvalue can then be improved either by minimizing \( |G_{\ell}(U)| \) on a two-dimensional grid\(^{15}\) or by Newton-Raphson iteration.\(^{33}\)
In the latter approach, if we wish to find a zero of a function \( \mathcal{D}(U) \) and have an approximate zero, \( U^{(n)} \), then a better approximation, \( U^{(n+1)} \), is found by making a Taylor expansion of \( \mathcal{D} \) about \( U^{(n)} \), retaining only first order terms and equating the result to zero.

Thus we put

\[
\mathcal{D}^{(n+1)} = \mathcal{D}^{(n)} + \Delta U^{(n)} \frac{\partial \mathcal{D}^{(n)}}{\partial U} = 0 ,
\]

where \( \mathcal{D}^{(n)} = \mathcal{D}(U^{(n)}) \), \( \Delta U^{(n)} = U^{(n+1)} - U^{(n)} \) and primes indicate differentiation with respect to \( U \). Then

\[
\Delta U^{(n)} = - \frac{\partial \mathcal{D}^{(n)}}{\partial U} \cdot \frac{U^{(n)}}{|\mathcal{D}^{(n)}|} \]

or, in terms of real and imaginary parts,

\[
U_r^{(n+1)} = U_r^{(n)} - \frac{\mathcal{D}_r^{(n)} U_r^{(n)} - \mathcal{D}_i^{(n)} U_i^{(n)}}{|\mathcal{D}^{(n)}|^2} \quad (2.8)
\]

\[
U_i^{(n+1)} = U_i^{(n)} - \frac{\mathcal{D}_i^{(n)} U_r^{(n)} - \mathcal{D}_r^{(n)} U_i^{(n)}}{|\mathcal{D}^{(n)}|^2} \quad (2.9)
\]

In the present case \( \mathcal{D} \equiv G_\lambda \)

\[
G_\lambda' = U \left[ H_\lambda^{(2)}(\lambda) \frac{H_{\lambda-2}^{(2)}(\lambda)}{H_{\lambda-1}^{(2)}(\lambda)} - \frac{J_{\lambda}^{(U)} J_{\lambda-2}^{(U)}}{(J_{\lambda-1}^{(U)})^2} \right] = \frac{U \lambda^2}{\lambda^2} \frac{J_{\lambda}^{(U)} J_{\lambda-2}^{(U)}}{J_{\lambda-1}^{(U)}^2} \quad (2.10)
\]

and iteration via Eqs. (2.8 and 2.9) gives the results shown in Fig. 5.1 for the first six \( HE_{\lambda\lambda} \) leaky modes. As indicated, the top set of curves represents the real part of \( U \) and the lower set, the imaginary part.

Each of the curves is divided into three regions representing bound, tunnelling and refracting modes. The boundary between bound and
tunnelling mode regions occurs at the point $U_{\ell 1} = V$ and between tunnelling and refracting at $\text{Re}(\lambda_{\ell 1}) = \ell - 1$ (or $r_{tp} = \rho$).

Fig. 5.1: The real and imaginary parts of the eigenvalue $U_{\ell 1}$ as functions of $V$ for the HE_{\ell 1} mode where the value of $\ell$ is that given in the circle. The dashed portion of each curve represents the tunnelling mode region. Values of $V$ below this region belong to the refracting region while values above correspond to trapped modes. In this case, $\delta = 0$.

In Fig. 5.2 the eigenvalues for the HE_{3 1}, HE_{4 1} and HE_{5 1} modes are compared with the first-order approximations given by Eqs. (2.4 and 2.5). The perturbation solution, indicated by dotted curves, gives $U_r$ accurately throughout the tunnelling mode region but the imaginary part, $U_i$, is only accurate very close to cutoff.
Fig. 5.2: Comparison of the eigenvalue $U_{\ell l}$ found by numerical solution of Eq. (1.11) with the perturbation solution used to find starting values, Eqs. (2.4 and 5). The dotted curve gives the perturbation solution and the circled figure indicates the value of $\ell$.

While the eigenvalues plotted in Fig. 5.1 are calculated from the $\delta = 0$ equation, some finite value for $\delta$ must be assumed in order to calculate an attenuation coefficient for the mode since, using Eqs. (2.2.8, 18, 27 and 4.4.6),

$$\alpha = \sqrt{2}\left[\left(\frac{\nu^2}{\delta} - U_r^2\right)^2 + \left(2U_r U_{\ell l}\right)\right]^{1/2} - \left(\frac{\nu^2}{\delta} - U_r^2\right)^{1/2}. \quad (2.11)$$

When $\delta = 0.01$, the maximum error obtained in the real part of $U$ by using Eq. (1.11) rather than Eq. (1.1) is less than about 0.15% throughout the
Fig. 5.3: The attenuation coefficient $a_{\ell l}$ as a function of $V$ for $HE_{\ell l}$ modes with $\ell$ values shown. The value of the parameter $\delta$ has been taken to be 0.01. As in Fig. 5.1, the dashed portion of each curve represents the tunnelling mode region.

tunnelling mode region and the maximum error in the imaginary part is less than about 1.5%. We therefore substitute $\delta = 0.01$ and the values of $U_r$ and $U_i$ given in Fig. 5.1 into Eq. (2.11) to obtain $\alpha$ as shown in Fig. 5.3. Once again each of the curves is divided into tunnelling and refracting mode regions (the attenuation coefficient for bound modes being zero, of course, in a non-absorbing waveguide).

If we define

$$\Delta v_{\ell l} = v_{c_{\ell l}} - v_{a=0.01}^{\ell l},$$

(2.12)
where $V_{c,\ell}^{\perp}$ is the cutoff value for mode $HE_{\ell,1}$ and $V_{c,\ell}^{\perp,\alpha=0.01}$ is the value of $V$ at which the $HE_{\ell,1}$ mode has an attenuation of $\alpha_{\ell,1} = 0.01$, then Fig. 5.4 shows the rate of increase of $\Delta V_{\ell,1}^{\perp,\ell}/V_{c,\ell}^{\perp}$ as $\ell$ increases. This confirms that, for fixed $V$, the more highly skew a mode is, the more slowly it attenuates.

The value of $\eta_{\ell,1}$ calculated from Eq. (1.16) for the first seven $HE_{\ell,1}$ modes is shown in Fig. 5.5 where, again the dashed portion of each curve represents the tunnelling mode region. To be more explicit, if we use the small $\theta_p$ form for the modes (Eqs. 2.2.27 - 35), then the
Fig. 5.5: \( \eta_{l1} \) calculated from Eq. (1.16) for the \( HE_{l1} \) modes with values as shown. In the refracting mode region, \( \eta_{l1} \) is identically 1.

The quantity plotted is

\[
\eta_{lm} = \frac{I_2}{I_1 + I_2},
\]

(2.13)

where

\[
I_1 = \frac{1}{|J_{l-1}(\lambda_{lm})|^2} \int_0^1 J_{l-1}(\lambda_{lm} R) J_{l-1}(\lambda_{lm}^* R) R \, dR
\]

(2.14)

and

\[
I_2 = \frac{1}{|H_{l-1}(\lambda_{lm})|^2} \int_1^R t_p H_{l-1}^{(2)}(\lambda_{lm} R) H_{l-1}^{(1)}(\lambda_{lm}^* R) R \, dR
\]

(2.15)

In Fig. 5.6, the results obtained from Eq. (1.16) are contrasted with those obtained by analytic continuation of Eq. (1.14)
Fig. 5.6: Comparison of $\eta$ calculated using Eq. (1.16) (-----) and Eq. (1.14) (-------).

below cutoff, which represents the "normalized" leaky-mode power within the core. The close agreement between two sets of curves confirms the usefulness of the turning-point concept and of the truncated leaky mode.
5.2.2 Asymptotic Results

In a multimode communication fibre, one is interested in large values of $l$ and $V$ so that numerical solution of the eigenvalue equation is, to say the least, inconvenient. As we shall see in Section 5.4, the ray-optical approximation provides a very useful method of circumventing this problem. But, under these circumstances, the modal analysis itself can be simplified by the use of asymptotic forms for the eigenvalue equation and characteristic mode parameters.

To begin with the eigenvalue equation itself, Snyder has shown that Eq. (1.11) is equivalent to

$$\frac{dU}{dV} = \frac{U}{V} \left(1 - \frac{1}{\xi}\right), \quad (2.16)$$

where $\xi$ is given by Eq. (1.15). Under the assumptions that $l >> 1$ and $l - \lambda > l^{1/3}$ or $\lambda >> l$ and $\lambda - l >> l^{1/3}$ (with $\lambda$ approximately real) it is shown in Appendix 5A that

$$\frac{1}{\xi} = 1 - \left[(l-1)^2 - \lambda_r^2\right]^{-1/2} + O\left((l-1)^2 - \lambda_r^2\right)^{-1}. \quad (2.17)$$

Substituting in Eq. (2.16) we have, therefore,

$$\frac{dU}{U} \approx \frac{dV}{V \left[(l-1)^2 - U^2 + V^2\right]^{1/2}}. \quad (2.18)$$

As we are interested chiefly in the weakly-leaky modes (i.e. those close to cutoff), we can obtain a first-order approximation to the real part of the eigenvalue $U$ by substituting its cutoff value in the denominator of the right-hand side of Eq. (2.18) and integrating:

$$U \approx V_c \left\{ \exp \left( \frac{1}{a} \cos^{-1}\left(\frac{a}{V}\right) - \cos^{-1}\left(\frac{a}{V_c}\right) \right) \right\}, \quad (2.19)$$

where
Fig. 5.7: Comparison of $\text{Re}(U_{\ell\ell})$ calculated using approximations (2.19) (·····) and (2.21) (----) with the exact numerical results (---).

$$a^2 = V_c^2 - (\ell-1)^2$$

and $V_c$ is the cutoff value of the mode under consideration.

Figure 5.7 shows that the agreement in the tunnelling-mode region between this simple form for the real part of $U$ and the numerical solution of Eq. (1.11) improves as $\ell$ increases and, even by the time $\ell = 7$, is very good in the weakly-leaky region.
Progressively simpler expressions for the real part of \( U \) (whose range of validity is progressively more restricted to the region near cutoff) can be found by making the \( U = V \) approximation earlier in the above derivation. Thus, putting \( U = V \) \((\lambda = 0)\) in Eq. (2.17) gives

\[
\frac{dU}{Ur} = \frac{dV}{(l-1)V} \tag{2.20}
\]

or

\[
U_r = V_c \left( \frac{V}{V_c} \right)^{1/(l-1)}, \quad l > 1 \tag{2.21}
\]

which approximation is also plotted in Fig. 5.7.

The expression given by Snyder and Mitchell,\(^{29}\)

\[
\Delta U_r = \Delta V/(l-1), \quad l \geq 3, \tag{2.22}
\]

where \( \Delta U_r = U_r - V_c \) and \( \Delta V = V - V_c \), is obtained by also putting \( U = V \) on the right hand side of Eq. (2.16). For \( l = 2 \), Snyder and Mitchell\(^{29}\) give an additional term

\[
\Delta U_r = \Delta V(1 + 1/\log(-\Delta V)), \quad l = 2 \tag{2.23}
\]

while for \( l = 1 \) there is no solution immediately below cutoff\(^{29}\) (as is the case for the modes of a slab waveguide). Sufficiently far below cutoff, the eigenvalue equation for the \( HE_{1m} \) modes simplifies to that of the slab waveguide, apart from a phase factor. The solutions of this equation have been investigated by Marcuse.\(^3\)

Provided attention is restricted to the weakly-leaky region, the imaginary part of \( U \) can now be found by a perturbation method.\(^{28}\) We assume \( U_i < U_r \) and make a Taylor expansion of \( G_\ell(U) \) about \( U_r \), retaining the first two terms to give

\[
G_\ell(U) = G_\ell(U_r) + iU_iG_\ell'(U_r) = 0. \tag{2.24}
\]
Equating real and imaginary parts gives

\[
G_{kr}(U_r) - U_i G'_{li}(U_r) = 0 \quad (2.25)
\]

and

\[
G_{li}(U_r) + U_i G'_{kr}(U_r) = 0 \quad ,
\]

from which we find

\[
U_i = - \frac{G_{li}(U_r)}{G'_{kr}(U_r)} = \frac{G_{kr}(U_r)}{G'_{li}(U_r)} . \quad (2.27)
\]

Using the first half of Eq. (2.27) one can then derive the expression\(^{28}\)

\[
U_i = \frac{2 U_r}{\pi v^2} \left| H_\ell^2(\lambda_r) \frac{H_{\ell-2}^2(\lambda_r)}{\lambda_r} \right|^{-1} , \quad (2.28)
\]

where \(\lambda_r\) is calculated by assuming \(U\) to be real.

The second half of Eq. (2.27) provides a check on the self-consistency of this perturbation analysis, being an alternative equation for the real part of \(U\). Figures 5.8(a-c) compare the values of \(U_r\) found from this equation with those found from Eq. (1.11). Within the range of validity (i.e. where \(U_i \ll U_r\)), agreement is very good; but there is marked (and rapidly growing) disagreement beyond this range.

We turn now to the longitudinal propagation constant, \(\beta\), and the attenuation coefficient, \(\alpha\). From Eq. (2.2.8), we can write

\[
\rho^2 \beta_r \beta_i = - U_r U_i \quad (2.29)
\]

and

\[
2(\rho \beta_r)^2 = \{ [ (\rho k_1)^2 + U_i^2 - U_r^2 ]^2 + 4 U_2^2 U_i^2 \}^{1/2} - [ (\rho k_1)^2 + U_i^2 - U_r^2 ] \quad . \quad (2.30)
\]

From Eq. (4.4.6), the attenuation coefficient \(\alpha\) can then be calculated immediately using above expressions for \(U_r\) and \(U_i\):

\[
\alpha = \frac{2 U_r U_i}{\rho \beta_r} . \quad (2.31)
\]
Fig. 5.8: Comparison of numerical solutions of Eq. (1.11) (---) and Eq. (2.27) (-----).
5.2.2

But in the case where $\beta_r = k_1$ and $\delta << 1$, this expression for $\alpha$ can be simplified using asymptotic results appropriate to modes in the refracting, tunnelling or transition regions. Noting that when $\lambda_r >> 1$ or $\lambda_r >> 1$

\[
\left| H_{\lambda-2}^{(2)}(\lambda_r) H_{\lambda}^{(2)}(\lambda_r) \right| \approx \left| H_{\lambda-1}^{(2)}(\lambda_r) \right|^2,
\]  

(2.32)

we rewrite Eq. (2.31) in the form

\[
\alpha \approx \frac{4}{\pi} \frac{u_\lambda}{p_\lambda} \frac{\sqrt{\delta}}{v_\lambda} \left| H_{\lambda-1}^{(2)}(\lambda_r) \right|^{-2},
\]  

(2.33)

where

(i) for refracting modes with $\lambda_r >> 1$ and $\lambda_r - \lambda >> 1/3$,

\[
\left| H_{\lambda-1}^{(2)}(\lambda_r) \right|^2 \approx \frac{2}{\pi} \left( \lambda_r^2 - (\lambda-1)^2 \right)^{-1/2};
\]  

(2.34)

(ii) for modes in the transition region with $\lambda >> 1$ and $\lambda - 1 = \lambda_r$,

\[
\left| H_{\lambda-1}^{(2)}(\lambda-1) \right|^2 \approx \frac{4a^2}{(\lambda-1)^{2/3}},
\]  

(2.35)

where

\[
a = \frac{\Gamma(1/3)}{2^{1/3} 3^{1/6} \pi (\lambda-1)^{1/3}};
\]

(iii) for tunnelling modes with $\lambda >> 1$ and $\lambda - \lambda_r >> 1/3$,

\[
\left| H_{\lambda-1}^{(2)}(\lambda_r) \right|^2 \approx \frac{2}{\pi} \left( (\lambda-1)^2 - \lambda_r^2 \right)^{-1/2} \exp \left\{ 2(\lambda-1) \cosh^{-1} \left( \frac{\lambda-1}{\lambda_r} \right) \right. \\
- 2 \left[ (\lambda-1)^2 - \lambda_r^2 \right]^{1/2} \left( 1 - \frac{\lambda_r^2}{(\lambda-1)^2} \right)^{3/2} + \ldots \right\},
\]  

(2.36)

\[
\approx \frac{2}{\pi} \left( (\lambda-1)^2 - \lambda_r^2 \right)^{1/2} \exp \left\{ \frac{2(\lambda-1)}{3} \left( 1 - \frac{\lambda_r^2}{(\lambda-1)^2} \right)^{3/2} + \ldots \right\},
\]  

\[
\lambda \approx \lambda_r.
\]  

(2.37)

A uniformly valid asymptotic expression in terms of Airy functions can also be found.
The remaining mode characteristic for which we find asymptotic expressions is the parameter $\eta$. For weakly-leaky modes, $\eta$ is simply the analytic continuation of the bound mode result, Eq. (1.14) below cutoff. Then using Eq. (2.17) when $\ell - \lambda_r \gg \ell^{1/3}$, $\ell \gg 1$ we have

$$
\eta \approx \frac{u^2}{v^2} \left[ 1 - \frac{1}{[(\ell-1) - \lambda_r^2]^{1/2}} - \frac{\lambda_r^2}{v^2} \right]$

$$
\approx 1 - [(\ell-1)^2 - \lambda_r^2]^{-1/2}
$$

when $U \approx V$. \hspace{1cm} (2.38)

Snyder and Mitchell\(^{29}\) also derive forms appropriate to smaller values of $\ell$ near cutoff. In particular, for $\ell \leq 4$ they give

$$
\eta_{\ell m} \approx \left[ \frac{\ell-2}{\ell-1} \right] \left[ 1 + \frac{U_{r c}}{V_{c m} (\ell-1)(\ell-3)} + \frac{1}{V_{c m}} \right], \hspace{1cm} (2.39)
$$

where $V_{c m}$ is the cutoff value for the $HE_{\ell m}$ mode and $U_{r}$ is given by Eq. (2.22).

For modes at the opposite end of the tunnelling region, i.e. near the refracting mode boundary, we return to the definition of $\eta_{\ell m}$ given by Eq. (2.13). As long as $U_{r} \gg U_{i}$,

$$
I_1 \approx \frac{\ell}{2} \left[ 1 - \frac{\lambda_{\ell m}^2}{U_{r}^2} \right] \frac{J_{\ell-2}(U_{r} \lambda_{\ell m})}{J_{\ell-1}(U_{r} \lambda_{\ell m})} \left/ J_{\ell-1}^2(U_{r} \lambda_{\ell m}) \right.$$

$$
= \frac{\ell}{2} \left[ 1 - \frac{\lambda_{\ell m}^2}{U_{r}^2} \right] \frac{H_{\ell}^{(2)}(\lambda_{\ell m})}{H_{\ell-2}^{(2)}(\lambda_{\ell m})} \left/ H_{\ell-1}^{(2)}(\lambda_{\ell m}) \right]^2.
$$

For large $\ell$, we have therefore

$$
I_1 \approx \frac{\ell}{2} \frac{\lambda_{\ell m}^2}{U_{r}^2} \hspace{1cm} (2.40)
$$

Since we are looking at the region around $R_{tp} \approx 1$, we can also approximate
5.3 \[ I_2 = R_{tp}^{\lambda m} - 1 \]
\[ = \frac{R_{tp}}{2} \left[ \left( R_{tp}^{\lambda m} \right)^2 - 1 \right] \] (2.40)
\[ = \frac{1}{2} \left[ \left( \frac{\lambda - 1}{\lambda \lambda m} \right)^2 - 1 \right]. \] (2.41)

Substituting Eqs. (2.40 and 2.42) in Eq. (2.13), we have then

\[ \eta_{\lambda m} = \frac{V^2}{V^2 + U_{\lambda m}^2} \left[ \left( \frac{\lambda - 1}{\lambda \lambda m} \right)^2 - 1 \right] \]
\[ \sim 1 - \frac{U_{\lambda m}^2}{V^2} \left[ \left( \frac{\lambda - 1}{\lambda \lambda m} \right)^2 - 1 \right]. \] (2.42)

5.3 EXACT EIGENVALUE EQUATION - DETAILED MODE CHARACTERISTICS

Once the simplified eigenvalue equation has been solved numerically, there ought in principle be no particular difficulty in applying precisely the same method to the solution of the more complicated exact equation. The fact that the expressions involved in the latter case are considerably more complicated, however, means that the solution does become more difficult - particularly near cutoff (which is the area of greatest interest).

The reason for this difficulty is that numerical errors which inevitably occur in the very small imaginary parts of the solution are magnified by the calculation of the Bessel and Hankel functions and the multitude of other operations so that the iterative procedure is occasionally unstable and does not converge properly. Consequently it is necessary to make a good "guess" at the starting value.
For small values of $\delta$, one can achieve this by first solving Eq. (1.11) at the required value of $V$ and then using this solution as a starting value for the exact equation. For large values of $\delta$ (say $\delta \geq 0.5$), even this is occasionally not good enough. The procedure then is to solve the eigenvalue equation some distance further below cutoff where the imaginary part is not so sensitive. $V$ is then increased to the required value by small increments, the eigenvalue at each step being calculated using the solution at the previous step as the starting value.

Fortunately, these complications only arise in a small number of (apparently random) modes at values of $V$ very close to cutoff. Apart from these pathological cases, however, the Newton-Raphson iteration converges reliably even when quite inaccurate starting values are used.

We set up the formalism necessary for application of Eqs. (2.8 and 2.9) by defining

$$D = M$$  \hspace{1cm} (3.1)

and

$$D' = M' = M_1'M_2 + M_1M_2' - M_3'$$,  \hspace{1cm} (3.2)

where

$$M_1' = \frac{U}{\lambda} f(\lambda) - f(U)$$,  \hspace{1cm} (3.3)

$$M_2' = M_1' - \frac{\delta}{1 - \delta} f(U)$$,  \hspace{1cm} (3.4)

$$M_3' = -\frac{2\delta^2}{1 - \delta} \left( \frac{V}{U\lambda} \right)^{\lambda} \left\{ \frac{U\delta}{V^2} + 2 \left( 1 - \frac{U^2\delta}{V^2} \right) \left( \frac{1}{U} + \frac{V}{\lambda^2} \right) \right\}$$,  \hspace{1cm} (3.5)

$M_1$, $M_2$ and $M_3$ are defined by Eqs. (1.7-9) and $f$ by

$$f(x) = \frac{1}{x} \left[ \frac{2\lambda}{x^2} - 1 + \frac{2(\lambda - 1)}{x} \frac{Z_{\lambda - 1}}{Z_{\lambda}} - \frac{Z_{\lambda - 1}^2}{Z_{\lambda}^2} \right]$$,  \hspace{1cm} (3.6)
where \( Z_{k-1} \) and \( Z_\lambda \) stand for \( J_{k-1}(U), J_\lambda(U) \) when \( x = U \) and \( H_{k-1}^{(2)}(\lambda), H_\lambda^{(2)}(\lambda) \) when \( x = \lambda \).

Using the methods outlined above, Eq. (1.1) has been solved for the cases \( \mu_1 = \mu_2 \) and \( \delta = 0.01, 0.05, 0.1 \) and 0.5. Figures 5.9 and 5.10 show the real and imaginary parts, respectively, of the \( HE_{\lambda m} \) mode eigenvalues for modes cut off at values of \( V \) less than 0.1 in the case \( \delta = 0.1 \). Figures 5.11 and 5.12 show the same eigenvalues when \( \delta \) has been increased to 0.5.

A clearer picture can be formed of the effect of changing \( \delta \) by inspection of Figs. 5.13–5.18. These show the variation in the real and imaginary parts of \( U_{\lambda m} \) with variation in \( \delta \) for the \( HE_{61} \), \( HE_{42} \) and \( HE_{91} \) modes. Figures 5.19–5.21 give the attenuation coefficients for each of these modes, as calculated from Eq. (2.11).

Clearly the attenuation coefficients decrease as the "skewness" of the mode increases. As a further illustration of this, Fig. 5.22 shows the power of the \( HE_{13,2} \) mode inside the core as a function of \( V \), for a number of values of \( \frac{Z}{\rho} \). The mode only has a sharp "cutoff" when viewed at a distance of about \( 10^5 \rho \) from the source.

Finally, Figs. 5.23 and 5.24 illustrate the splitting of the degeneracy between \( HE_{k+2,m} \) and \( EH_{\lambda m} \) modes as \( \delta \) is increased in the specific case of the \( HE_{41} \) and \( EH_{21} \) modes.
Fig. 5.9: $\text{Re}(U_{\ell m})$ calculated from Eq. (1.1) with $\delta = 0.1$. The modes are $\text{HE}_{\ell 1}$ (---), $\text{HE}_{\ell 2}$ (----) and $\text{HE}_{\ell 3}$ (.....).
Fig. 5.10: Imaginary parts of the solutions corresponding to Fig. 5.9.
Fig. 5.11: As for Fig. 5.9 with $\delta = 0.5$. 
Fig. 5.12: As for Fig. 5.10 with $\delta = 0.5$. 
Fig. 5.13: \( \text{Re}(U_{61}) \) calculated from Eq. (1.1) for various values of \( \delta \).
Fig. 5.14: $\text{Im}(U_{61})$ calculated from Eq. (1.1) for various values of $\delta$. 
Fig. 5.15: $\text{Re}(U_{42})$ for various values of $\delta$. 
Fig. 5.16: $\text{Im}(U_{42})$ for various values of $\delta$. 
Fig. 5.17: $\text{Re}(U_{31})$ for various values of $\delta$. 
Fig. 5.18: $\text{Im}(\mathcal{U}_{91})$ for various values of $\delta$. 
Fig. 5.19: $\frac{\alpha_{61}}{\sqrt{\delta}}$ calculated from Eq. (2.11) for various values of $\delta$. 
Fig. 5.20: $\alpha_{42}/\sqrt{\delta}$ for various values of $\delta$. 
Fig. 5.21: $\frac{\alpha_{g1}}{\sqrt{\delta}}$ for various values of $\delta$. 
Fig. 5.22: Variation of HE\(_{13,2}\) modal power within the core with \(V\) and distance from the source.
Fig. 5.23: Comparison of real parts of $EH_{21}$ and $HE_{41}$ mode eigenvalues for various values of $\delta$. When $\delta = 0$ (---) the modes are degenerate.
Fig. 5.24: Imaginary parts of eigenvalues given in Fig. 5.23.
5.4 EXCITATION AND PROPAGATION OF THE UNBOUND FIELD—LEAKY MODE AND LEAKY RAY ANALYSES\textsuperscript{35}

The launching and propagation of leaky modes under excitation by truncated-plane-wave or totally incoherent sources (as described in Chapter 3) can now be investigated using mode characteristics calculated by the methods of the preceding two sections. For large V fibres, light propagation under these conditions has previously been studied using the approximation of leaky-ray theory.\textsuperscript{32} Comparison of results obtained by these two methods will, therefore, be useful in providing criteria for the range of validity of modified geometric optics.

The calculations will be performed at values of $V=10$ and 20 and $\delta=10^{-2}$ and $10^{-4}$ (corresponding to critical angles of 0.1 and 0.01 radians). At these relatively low values of $V$ it is inappropriate to compare bound modes and bound rays separately from leaky modes and leaky rays because, as Fig. 5.25 (taken from Ref. 36) shows, the modal result is subject to resonance effects which the ray theory neglects by assuming a continuum of modes. It is therefore only meaningful to compare the total power within the core as calculated by the mode and ray theories.

5.4.1 Leaky Mode Analysis

As we have seen in Chapter 4, the electric field within the care of a dielectric waveguide can be written to a very good approximation as

$$E = \sum_p a_p \tilde{e}_p(x,y) e^{-i\beta z}$$

(4.1)

where the sum extends over both bound and leaky modes, the leaky mode
Fig. 5.25: The per cent error of meridional-ray optics, for use in determining the summed time-averaged power of the bound modes transmitted within a circular optical fibre illuminated by an incoherent source is given by the solid line. The dotted line gives an asymptotic approximation to the error and the fluctuations of the actual error about this smooth curve are due to mode resonance effects. [From Ref. 36.]

terms being simply the analytic continuation of the corresponding bound mode terms below cutoff. Because of the small values of $\delta$ involved, the approximate degeneracy of the $HE_{k+2,m}$ and $EH_{km}$ modes is assumed throughout this section although the exact eigenvalues are used in calculating the attenuation coefficients of the HE modes. For added simplicity, the $\beta=k$ modes (Eqs. 2.2.28 - 35) are used. The validity of this approximation will be justified below.
The one modification which might possibly have to be made to Eq. (4.1) arises in the case where a pole, either bound or leaky, lies very close to the saddle point in the \( \psi \)-plane (Sections 4.2 and 4.4). Then the amplitude coefficient for that mode will be modified at all points within a certain distance of the source, \( a_p \) being replaced by \( a_p [1 + Q(z)] \) where \( Q \) is the error function term in Eq. (4.4.9). In the present calculations, the only time this modification approaches any significance is the case of the \( HE_{17,1} \) bound mode at \( V = 20 \) where \( |\beta - k_2| \approx 0.007 \). But as is indicated by Fig. 5. 26 which shows \( \log_{10} |Q(z/p)|^2 \) as a function of \( z/p \), even in this case the error-function term only adds significantly to the power within the waveguide at relatively short distances from the source. It will therefore be ignored in the calculations of this section.

The time-averaged power propagating within the waveguide can now be calculated from

\[
\langle P_{in} \rangle = \pi p^2 \text{Re} \left\{ \sum_{p,q} a_{p,q}^* e^{-i(\beta - \beta^*)z} C_{pq} \right\}
\]

(4.2)

in the case of the truncated plane wave and

\[
\langle P_{in} \rangle = \pi p^2 \text{Re} \left\{ \sum_{p,q} (a_{p,q}^*) e^{-i(\beta - \beta^*)z} C_{pq} \right\}
\]

(4.3)

for the quasi-monochromatic incoherent source, where \( C_{pq} \) is defined by Eq. (3.2.38) and the amplitude coefficients in both cases are given in Chapter 3.

When the \( \beta = k \) modes are used, \( \langle a_{p,q} a_{p,q}^* \rangle = 2C_{pq} \) so that Eq. (4.3) can be rewritten as

\[
\langle P_{in} \rangle = 2\pi p^2 \text{Re} \left\{ \sum_{p,q} C_{pq} e^{-i(\beta - \beta^*)z} \right\}
\]

(4.4)
Equations (4.2) and (4.4) define the quantities to be compared with the corresponding results from leaky-ray theory.

Fig. 5.26: The modification which must be applied to the $HE_{17,1}$ mode contribution to account for the proximity of the saddle point (Eq. (4.4.9)).

5.4.2 Leaky Ray Analysis

The leaky ray approach to the study of light propagation in dielectric waveguides is basically a modification of classical ray tracing techniques to take account of the fact that some of the rays which classical geometric optics predicts to be trapped in the core are, in fact, radiating.
A ray travelling along the waveguide is characterized by two angles: \( \theta_z \) is the inclination of the ray to the waveguide axis and \( \phi \) is the angle between the ray projection onto the cross-sectional plane containing the point of incidence at the fibre boundary and the tangent at the point of incidence in that plane. Trapped rays have \( \theta_z < \theta_c \), where \( \theta_c \) is the complement of the critical angle, i.e.

\[
\theta_c = \sin^{-1}\left(1 - \left(\frac{n_2}{n_1}\right)^2\right) = \sin^{-1}(\sqrt{\nu}) .
\] (4.5)

Refracted rays make an angle \( \theta_N = \cos^{-1}(\sin \theta_z \sin \phi) \) with the local normal which is less than \( \frac{\pi}{2} - \theta_c \) and hence leave the fibre by refraction loss. The rays with \( \theta_z > \theta_c \) but \( \theta_t = \frac{\pi}{2} - \theta_N < \theta_c \) are not refracted rays, but nor are they trapped — these are the tunnelling rays which correspond to tunnelling leaky modes.

The attenuation of leaky rays can be described by an attenuation coefficient \( \alpha \) so that the power \( P(z) \) of a narrow, parallel ray bundle after travelling a distance \( z \) along the fibre is

\[
P(z) = P(0) e^{-\alpha z/\rho} \] (4.6)

just as in the case of leaky modes. In ray terms, the attenuation coefficients given by Eqs. (2.33-37) can be rewritten for

(i) refracting rays with \( (\theta_t^2 - \theta_c^2) \gg [(\sin^2 \theta_z - \theta_c^2)/\rho k]^{2/3} \)

\[
\alpha \approx 2 \left(\frac{\sin^2 \theta_z}{\theta_c^2 \cos \theta_c} \right) \left(\frac{\sin \theta_t}{\theta_c} \right)^2 - 1 \right)^{1/2} ;
\] (4.7)

(ii) rays at the transition point, \( \theta_t = \theta_c \)

\[
\alpha \approx 1.59 \left(\frac{\theta_c}{\cos \theta_z} \right) \left(\frac{1}{\sqrt[3]{\nu}} \right) \left(\frac{\sin \theta_z}{\theta_c} \right)^2 \left[\left(\frac{\sin \theta_z}{\theta_c} \right) - 1 \right]^{1/2} ;
\] (4.8)
(iii) tunnelling rays

\[ \alpha = 2 \left( \frac{\sin^2 \theta}{\theta \cos \theta \theta} \right) \left( 1 - \frac{\sin^2 \theta}{\theta \theta} \right) \frac{1}{2} \exp \left\{ 2\rho k \left( \frac{\sin^2 \theta}{\theta \theta} \right) \right\} \]

\[ \times \left[ \frac{\theta \theta}{\sin^2 \theta \theta - \sin^2 \theta \theta} \right] \frac{1}{2} - \cosh^{-1} \left( \frac{\sin^2 \theta}{\theta \theta - \sin^2 \theta \theta} \right) \right\}. \quad (4.9) \]

When \( \theta_\phi = 0 \) and \( \theta_z \) is small, Eq. (4.9) becomes

\[ \alpha = 2 \frac{\theta \theta}{\theta \theta} \left( \frac{\theta \theta}{\theta \theta^2} \right) \frac{1}{2} \left( \frac{\theta \theta}{\theta \theta} \right) \frac{1}{2} \exp \left[ -2\rho k \left( \frac{\theta \theta}{\theta \theta - \sin^2 \theta \theta} \right) \right] \] \( (4.10) \)

and when \( \theta_t \approx \theta_c \),

\[ \alpha = 2 \left( \frac{\sin^2 \theta}{\theta \cos \theta \theta} \right) \left( 1 - \left( \frac{\theta \theta}{\theta \theta} \right) \right) \frac{1}{2} \exp \left[ -2\rho k \frac{\theta \theta}{\theta \theta - \sin^2 \theta \theta} \right] \right\}. \quad (4.11) \]

In practice, Eq. (4.11) is a good approximation throughout most of the tunnelling mode region.

Once these attenuation coefficients are known, it only remains to specify the angular distribution and intensity of rays at the source and then to trace the rays down the fibre as in classical geometric optics, but incorporating the above attenuations. This has been done by Pask and Snyder\(^\text{32}\) for Lambertian and truncated plane wave sources and also for totally incoherent sources.\(^\text{37}\) The leaky ray results to be used for comparison with the leaky mode calculations are taken directly from their work.

5.4.3 Numerical Results

Before looking at this comparison, some of the "technicalities" associated with the mode calculation will be touched on briefly. To
Fig. 5.27: Normalized power inside a waveguide excited by an incoherent source as a function of axial distance from the source for three closely separated values of V. The difference between the curves arises from discrete-mode effects. (The 19.5 and 20.5 results are not identical at large distances but are indistinguishable on the scale of this diagram.)

begin with the difficulty mentioned above concerning our inability to make a direct comparison between the unbound parts of the field as calculated by mode and ray analyses, Fig. 5.27 is a further illustration of this point. The curves show the z-dependence of the power within a small fibre excited by an incoherent source (normalized by the input power calculated from Eq. (3.3.13)). The difference between the curves for the three closely spaced values of V arises because, coincidentally, there are more modes with cutoffs close to V = 19.5 and V = 20.5 than to
V = 20. So at large z/p, the V = 19.5 and V = 20.5 fibres have almost twice as much of their unbound power within the core as the V = 20 fibre. On the other hand the ray results are almost imperceptibly changed by this small change in V. At larger V-values where the density of modes is much greater, this will obviously cease to be a problem.

The other detail to be considered is the number of leaky modes which have to be included in the sums in Eqs. (4.2) and (4.4). Figure 5.28 shows the normalized power inside the core of a fibre with V = 20, \( \delta = 0.01 \) illuminated by a totally incoherent source. The curve labelled "20" gives the bound mode power while the others include all leaky modes with cutoffs between V = 20 and \( V_c^{\text{max}} \), where \( V_c^{\text{max}} \) is the number on the left of each curve. Clearly, many more modes are required to ensure convergence of the sum close to the source than are required at greater distances. Figure 5.29 shows the percentage error,

\[
\frac{P_{30} - P_{V_c^{\text{max}}}}{P_{30}} \times 100
\]

as a function of z/p, where \( P_{V_c^{\text{max}}} \) is the power in the core calculated by including all modes with cutoff less than or equal to \( V_c^{\text{max}} \) and \( P_{30} \) is the power in the case \( V_c^{\text{max}} = 30 \).

Proceeding now to a comparison between the results obtained by the mode and ray analyses, we plot in Figs. 5.30 and 5.31 the normalized power within the waveguide core as a function of z/p when the fibres are illuminated by a totally incoherent, quasi-monochromatic source. In both figures, the solid curves represent the ray solutions and the broken curves, the mode solutions. The values of V used are 10 in Fig. 5.30 and 20 in Fig. 5.31 while the values of \( \theta_c^{1/2} \) (\( \delta^c \)) are as indicated on
Fig. 5.28: The effect on power calculated by Eq. (4.4) of changing the number of leaky modes included (see text for further explanation). \[ \bar{P} = \frac{P_{in}}{P_s} \times 10^3. \]

Fig. 5.29: Percentage difference between the curves shown in Fig. 5.28.
Fig. 5.30: Power inside a waveguide excited by an incoherent source, calculated using ray theory (——) and mode theory (-----) for the cases $V=10$ and $\theta_c = 0.1$ and 0.01. $P = P_{in}/P_s \times 10^n$ where $n=3$ for $\theta_c = 0.1$ and $n=5$ for $\theta_c = 0.01$.

The diagrams. The percentage differences between the mode and ray results are shown in Fig. 5.32 where, now, the solid curves correspond to the case $\theta_c = 0.1$ and the broken curves to $\theta_c = 0.01$.

There is a certain amount of confusion in the $\theta_c = 0.01$ results at short distances because neither refracted rays nor saddle-point corrections to the mode results have been included. Pask has studied the contribution of the refracted rays in the case of a Lambertian source and found that for small $\theta_c$, the initial power is divided among the three ray types as follows:
Fig. 5.31: As for Fig. 5.30 with V increased to 20.

Fig. 5.32: Percentage differences between the ray and mode results shown in Figs. 5.30 and 5.31.
\[
\frac{P_{\text{trap}}(0)}{P_{\text{total}}} = \frac{2}{C}
\]
\[
\frac{P_{\text{tunn}}(0)}{P_{\text{total}}} \approx \frac{2}{C} \left( 1 - \frac{3\theta}{\pi} \right)
\]
\[
\frac{P_{\text{ref}}(0)}{P_{\text{total}}} \approx 1 - \frac{2}{C} \left( 1 - \frac{3\theta}{\pi} \right)
\]

He has also found an upper bound on the refracted ray contribution as a function of \(z\). Our results here are consistent with there being a similar type of behaviour for the totally incoherent source.

Beyond this near-source region, the errors are ordered by

\[
\begin{align*}
\left( V = 20 \right) &< \left( V = 20 \right) \left( \theta_c = 0.01 \right) < \left( V = 10 \right) \left( \theta_c = 0.1 \right) < \left( V = 10 \right) \left( \theta_c = 0.1 \right)
\end{align*}
\]

indicating that

1. as the radius-to-wavelength ratio \((\rho/\lambda)\) for fixed \(\theta_c\) increases, the ray approximation increases in accuracy, as expected, and

2. the magnitudes of both \(V\) and \(\rho/\lambda\) must be taken into account in assessing the usefulness of ray theory.

The fact that the error is less than 10\% for \(V = 20\) even when \(\theta_c\) is a realistic figure like 0.1 indicates that, at least in the case of diffuse sources, the leaky ray theory can be confidently used for \(V \gg 20\). This is important because, while the additional contribution to the power from the unbound field is relatively small for \(V = 20\), Fig. 5.33 (which was obtained by Pask and Snyder\(^{32}\) using ray theory and a Lambertian source) shows that for large values of \(V\) the leaky mode contribution can by no means be ignored, and since the \(V = 20\) calculations already involved 116 distinct modes (excluding the EH-HE mode degeneracy and the two possible polarizations for each mode) the
Fig. 5.33: Fraction of incident power remaining within a fibre illuminated by an incoherent, Lambertian source (calculated by Pask and Snyder\textsuperscript{32} using a leaky ray analysis). The initial power is divided 52\% into trapped rays and 48\% into leaky rays.

Ray approximation is very nearly essential, in practical terms, for the larger values of $V$.

Before looking at the corresponding results for coherent excitation, we return to the justification of the use of $\beta \approx k$ modes. To do this we again use a result calculated by Pask\textsuperscript{37} for the incoherent source. Figure 5.34 shows the percentage of initial tunnelling ray power with $$(\sin \theta / \sin \theta_c) \ll y$$ and confirms that most of the tunnelling rays have $\theta$ quite small — corresponding to modes with $\beta \approx k$. 
Fig. 5.34: The percentage of power launched into tunnelling rays with \( \frac{\sin \theta}{\sin \theta_c} \leq y \), given by \(^{37}\)

\[
P(y) = \frac{J(\theta_c, \theta)}{J(\theta_c, \pi/2)} \times 100,
\]

where

\[
J(\theta_c, \theta) = \int_{\theta_c}^{\theta} \sin t \left( 1 + \cos^2 t \right) \left[ \pi - \psi(t) - \sin \psi(t) \right] dt,
\]

\[
\psi(t) = 2 \cos^{-1} \left( \frac{\sin \theta_c}{\sin t} \right).
\]
5.4.3

Turning now to the case of a coherently excited fibre, Figures 5.35-38 show the power inside the waveguide core when the illuminating field is a finite portion of a plane wave incident over the fibre aperture at angle $\theta$ to the axis. The figures labelling each pair of curves (corresponding to ray (---) and mode (-----) results) give the value of $\log_{10}(z/p)$. The first of these figures, Fig. 5.35, is for the case $V=20$, $\theta_c = 0.01$ and, again, there is very good agreement between the two methods except at large distances where the divergence is attributable to numerical difficulties in the ray calculations which lead to an underestimate of the rate of leakage of the most weakly tunnelling rays. These difficulties decrease as $V$ increases. In Fig. 5.36, the value of $\theta_c$ is increased to 0.1 and agreement is not so good, even at the shorter distances. When $V$ is decreased to 10, agreement is quite poor even when $\theta_c = 0.01$ (Fig. 5.37) and is at best "qualitative" when $\theta_c = 0.1$ (Fig. 5.38).

The reason for the more restricted range of validity of the ray analysis in the case of the coherent source would seem to be that this analysis ignores the effects of diffraction at the aperture and these become more pronounced as $V$ is decreased or $\theta_c$ is increased.

The conclusions to be drawn from the results of this section are, therefore, that leaky ray theory gives an adequate description of the waveguide field on a fibre with $V \geq 20$ and $\theta_c$ of the order of 0.1 or less with a diffuse (incoherent) source but that when the excitation is coherent, either $\theta_c$ must be decreased by an order of magnitude or $V$ increased.
Fig. 5.35: Power inside a waveguide illuminated by a finite portion of plane wave incident over the aperture at angle θ. The figures labelling each pair of curves (calculated using rays (——) and modes (—–)) give the value of $\log_{10}(z/p)$. The curve labelled "0" gives the power in the trapped and tunnelling rays at $z=0$. The parameters in this case are $V=20$ and $\theta_c=0.01$. 
Fig. 5.36: As for Fig. 5.35 with $V = 20$ and $\Theta_c = 0.1$. Curves for the larger distances are not shown because agreement is considerably worse than in the cases which are shown and their inclusion would confuse the picture.
Fig. 5.37: As for Fig. 5.36 with $V=10$ and $\theta_c = 0.01$.

Fig. 5.38: As for Fig. 5.38 with $V=10$ and $\theta_c = 0.1$. 
5.5 NUMBER OF LEAKY MODES

At the relatively low values of \( V \) treated in the previous section, it was possible to calculate the individual mode characteristics numerically in order to find which modes would still be significant after a certain distance. For larger values of \( V \), and for inhomogeneous refractive index profiles, where such exact calculations may be impractical, it is of some help to be able to estimate the number of leaky modes with less than a given attenuation constant.

Extending the Gloge-Marcatili\(^{39}\) theory of multimode graded-core fibres to include leaky modes, Gloge\(^{40}\) and Stewart\(^{41}\) have provided a method of obtaining such estimates by calculating the number of modes whose turning points lie at or beyond a particular radius. The turning point radius and attenuation coefficient can then be approximately related\(^{40}\) to convert this to the number of modes with attenuation coefficient less than or equal to a given amount. The calculation is outlined below:

Consider a dielectric waveguide consisting of a core of radius, \( \rho \), with wave number \( k_1(r) \), possibly non-uniform, and a cladding with uniform wave number, \( k_2 \). The local wave vector corresponding to a mode of this waveguide can be decomposed, at any point, into its components in the \( r, \phi \) and \( z \) directions. The radial component has the form

\[
\begin{align*}
\frac{u_1^2(r)}{r^2} &= k_1^2(r) - \beta^2 - \frac{\beta^2}{r^2} \\
\frac{u_2^2(r)}{r^2} &= k_2^2 - \beta^2 - \frac{\beta^2}{r^2}
\end{align*}
\]

(5.1)

in the core region and

\begin{align*}
\frac{u_2^2(r)}{r^2} &= k_2^2 - \beta^2 - \frac{\beta^2}{r^2}
\end{align*}

(5.2)

in the cladding.

For a particular bound or tunnelling mode, with \( \lambda \) and \( \beta \) fixed, there are two radii, \( r_1 \) and \( r_2 \), at which \( u_1(r) \) vanishes. These define
an annulus in which $u_1$ is real and the modal field is oscillatory. In order that this oscillatory field be associated with a mode, it must form a standing wave in the radial direction — i.e. the total phase change between $r_1$ and $r_2$ must add up, at least approximately, to an integer number of half-periods. So, if $m$ denotes the number of half-periods

$$m\pi = \int_{r_1}^{r_2} u_1(r) \, dr . \quad (5.3)$$

To count the number of modes in any particular category, it is now necessary to specify the limits on $\ell$, $m$ and $\beta$. For bound modes, $\beta \gg k_2$ so that, from Eq. (5.1), the largest value of $\ell$ which occurs when $m=0$ and $\beta=k_2$ is

$$\ell_{\text{max}} = r[k_1^2(r) - k_2^2]^{\frac{1}{2}} . \quad (5.4)$$

The largest value of $m$ occurs when $\ell=0$ and $\beta=k_2$.

So we find the total number of bound modes, $N_B$, by summing Eq. (5.3) over all values of $\ell$ between 0 and $\ell_{\text{max}}$ with $\beta=k_2$. Provided $\nu \gg 1$, $\ell_{\text{max}} \gg 1$ so that we can consider $\ell$ a continuous variable and integrate to find

$$N_B = \frac{4}{\pi} \int_0^{\ell_{\text{max}}} \int_{r_1(\ell)}^{r_2(\ell)} \left( k_1^2(r) - k_2^2 - \frac{\beta^2}{r^2} \right) \frac{dr}{\ell} . \quad (5.5)$$

The factor 4 allows for HE and EH modes and for two polarizations of each. Interchanging orders of integration,

$$N_B = \frac{4}{\pi} \int_0^{\rho} \int_0^{\ell_{\text{max}}} \left( k_1^2(r) - k_2^2 - \frac{\beta^2}{r^2} \right) \frac{dl}{\ell} dr \quad (5.6)$$

$$= \int_0^{\rho} [k_1^2(r) - k_2^2] r \, dr . \quad (5.7)$$

This can also be written in the form
5.5

\[ N_B = \int_0^\rho \lambda_{\text{max}} [k_1^2(r) - k_2^2]^{1/2} \, dr , \]  \hspace{1cm} (5.8)

where \( \lambda_{\text{max}} \) is given by Eq. (5.4).

For the step-index fibre, Eq. (5.8) gives

\[ N_{BS} = \frac{v^2}{2} \]  \hspace{1cm} (5.9)

while for a graded-index fibre with,

\[ k_1^2(r) = k_1^2(0) \left\{ 1 - \delta \left( \frac{r}{\rho} \right)^q \right\} , \quad (q > 0) \]  \hspace{1cm} (5.10)

\[ N_{Bq} = \frac{q}{q + 2} N_{BS} . \]  \hspace{1cm} (5.11)

But now suppose we want all modes with a turning point or outer caustic at a radius greater than or equal to \( r_0 > \rho \). Then \( \beta \) is no longer restricted by being greater than \( k_2 \). Instead, imposing the conditions that

\[ u_2^2(r_0) = 0 \]

and

\[ u_2^2(r) < 0 , \quad \rho < r < r_0 \]

we find the lower limit on \( \beta \) is given by

\[ \beta \geq \left( k_2^2 - \frac{\rho^2}{r_0^2} \right)^{1/2} . \]  \hspace{1cm} (5.12)

This gives the maximum value of \( \ell \) as

\[ \ell_{\text{max}} = r r_0 \left( k_1^2(r) - k_2^2 \right)^{1/2} \]  \hspace{1cm} (5.13)

in place of Eq. (5.4). Substituting these values for the maximum value of \( \ell \) and the minimum value of \( \beta \) in Eq. (5.3) and summing over all \( \ell \) gives
\[ N(r_0) = \frac{4}{\pi} \int_0^\rho \int_0^{\max} \left( k_1^2(r) - k_2^2 - \frac{1}{r_0^2} \right) \frac{1}{r^2} r^2 \, dl \, dr \]

\[ = \int_0^\rho \frac{k_1^2(r) - k_2^2}{r_0^2 - r^2} r_0 r \, dr. \quad (5.14) \]

Inspection of Eq. (5.14) shows that \( N(r_0) \) is given by the same expression as \( N_B \) (Eq. (5.8)) provided \( \ell_{\max} \) is adjusted.

In the step-index case, Eq. (5.14) can be integrated to give

\[ N(r_0)_S = 2\left(\frac{r_0}{\rho}\right)^2 \left\{ 1 - \left[ 1 - \left(\frac{\rho}{r_0}\right) \right]^{\frac{1}{2}} \right\} N_{BS}. \quad (5.15) \]

When \( r_0 = \rho \), all tunnelling rays are included and

\[ N(\rho)_S = 2N_{BS}. \quad (5.16) \]

That is, there are equal numbers of bound and tunnelling modes. On the other hand, when \( r_0 \gg \rho \) so that we only include weakly leaky modes,

\[ N(r_0)_S = \left\{ 1 + \frac{1}{4} \left( \frac{\rho}{r_0} \right)^2 \right\} N_{BS}. \quad (5.17) \]

For the case of the graded profile given by Eq. (5.10), Eq. (5.14) gives

\[ N(r_0)_q = I \times N_{BS}, \quad (5.18) \]

where

\[ I = \frac{2r_0}{\rho^2} \int_0^\rho \frac{1 - (r/\rho)^q}{r_0^2 - r^2} r^2 \, dr. \quad (5.19) \]

In general, integration of Eq. (5.19) gives an incomplete beta-function, but when \( q \) is an integer, integration in terms of elementary functions is possible. In the parabolic index case \((q = 2)\),

\[ I = 2\left(\frac{r_0}{\rho}\right)^2 - \frac{4}{3} \left(\frac{r_0}{\rho}\right)^4 \left\{ 1 - \left[ 1 - \left(\frac{\rho}{r_0}\right)^2 \right]^{3/2} \right\}. \quad (5.20) \]
When \( r_0 = \rho, \ I = \frac{2}{3} \) so that the total number of bound and tunnelling modes is

\[
N(\rho)_{q=2} = \frac{2}{3} N_{BS} = \frac{4}{3} N_{B,q=2} \tag{5.21}
\]

while when \( r_0 \gg \rho \), the number of bound and weakly leaky modes is

\[
N(r_0)_{q=2} = \frac{1}{2} \left\{ 1 + \frac{1}{6} \left( \frac{\rho}{r_0} \right)^2 \right\} N_{BS} = \left\{ 1 + \frac{1}{6} \left( \frac{\rho}{r_0} \right)^2 \right\} N_{B,q=2} \tag{5.22}
\]

Contrasting Eq. (5.21) with Eq. (5.16) and Eq. (5.22) with Eq. (5.17) shows that the ratio of the numbers of leaky and bound modes on a parabolic index fibre is smaller than that on a step-index fibre. Leaky modes may nevertheless be more important on the graded index fibres because it is much more difficult to avoid launching leaky modes on such a fibre than it is on a step-index fibre.\(^41,42\)

While the value of \( \frac{r_0}{\rho} \) gives an indication of the rate of attenuation of a leaky mode, in the case of the step-index fibre, we can use the asymptotic expressions for the attenuation coefficient given in Sections 5.2 or 5.4 to obtain a more explicit relationship for the number of modes with attenuations less than or equal to a given amount.\(^40\)

Beginning with Eq. (4.10) and putting \( \frac{\theta Z}{\theta c} = \frac{U}{V} \), we have

\[
\alpha = 2 \theta_c \left[ \frac{U}{V} \right]^2 e^{2V \left( \frac{U - V}{U + V} \right)^{U-1}} \tag{5.23}
\]

When \( V \) is large and we consider only weakly leaky modes (\( \lambda \approx U \approx V \)), this reduces to
\[ \alpha = 2 \theta \left( \frac{u^2 - v^2}{4l^2} \right)^V. \]  

Modes with a turning point at \( r_0 \) or further have \( \lambda^2 = u^2 - v^2 \leq \frac{2l^2}{r_0^2} \), so that the upper limit on \( \alpha \) is given by

\[ \alpha = 2 \theta \left( \frac{e^2}{4} \left( \frac{p}{r_0} \right)^2 \right)^V. \]  

From Eq. (5.17) we have, therefore, the number of modes with attenuation less than or equal to \( \alpha \) given by

\[ N(\alpha)_s = \left[ 1 + e^{-\frac{\alpha}{2\theta}} \left( \frac{\alpha}{2\theta} \right)^{1/V} \right] N_{BS}. \]  

5.6 MODE ATTENUATION DUE TO MATERIAL ABSORPTION

Up to this point, the thesis has been concerned entirely with ideal, non-absorbing waveguides. Before launching into a discussion of visual photoreceptors, systems whose function is to absorb light, it seems appropriate to discuss, at least briefly, the additional attenuation of leaky modes due to material absorption in the core and cladding.

First we consider the attenuation of bound modes due to absorption loss. (For early work on this subject, see Refs. 43, 44.) We define a power absorption coefficient \( \gamma_p \) for mode \( p \) by

\[ \gamma_p = - \frac{1}{P_p} \frac{dP_p}{dz} \cdot \frac{P_p}{P_{tot}}. \]  

where \( P_{tot} \) is the total modal power. Then intuitively, \( \gamma_p \) should equal the absorption coefficient \( \gamma_1 \) of the core material in bulk multiplied by the fraction \( \eta_p \) of the mode's power passing through the core plus the
absorption coefficient \( \gamma_2 \) of the cladding material in bulk multiplied by the fraction \((1 - \eta_p)\) of the mode's power passing through the cladding, i.e.

\[
\gamma_p = \gamma_1 \eta_p + \gamma_2 (1 - \eta_p) \tag{6.2}
\]

where

\[
\gamma_1 = \frac{2k_1 n_{1i}}{n_{1r}} = \frac{4\pi n_{1i}}{\lambda} \tag{6.3}
\]

\[
\gamma_2 = \frac{2k_2 n_{2i}}{n_{2r}} = \frac{4\pi n_{2i}}{\lambda} \tag{6.4}
\]

\( n_1 = n_{1r} + in_{1i} \) and \( n_2 = n_{2r} + in_{2i} \) are the core and cladding refractive indices, \( k_1 \) and \( k_2 \) are the corresponding wave numbers and \( \lambda \) here is the wavelength in vacuo.

In fact, the intuitive result, Eq. (6.2) is valid only when the absorption losses are small \((n_1 << n_2)\) for both core and cladding, when \( n_{1r} \) is only slightly greater than \( n_{2r} \) and when the axial component \( \beta_p \) of the modal and vector approximately equals the wave number \( k_1 \) of the core medium.\(^{45}\) When \( n_{1r} = n_{2r} \), the bound modes have \( \beta_p = k \).

Now consider the extension of these arguments to find \( \gamma_p \) for a leaky mode. The only obstacle is the possibility that slight absorption due to the cladding region beyond the caustic, \( r \geq r_{tp} \), contributes significant additional attenuation to the guided portion of the leaky mode. But in fact it does not since the fields for \( r > r_{tp} \) represent energy that has already been lost from the guided region. Thus using the definition of \( \eta_p \) applicable to both bound and leaky modes, Eq. (1.17), we see that Eq. (6.2) also holds for leaky modes under the conditions listed above.
Modes for which $r_{tp}$ is approaching $\rho$ have very high intrinsic attenuation compared to which the absorption loss is quite small. However for weakly leaky tunnelling modes with $r_{tp} \gg \rho$, absorption loss is more significant and the attenuation coefficient of such a mode must be found by summing the power absorption coefficient (normalized by $\rho$) and the intrinsic leakage coefficient.

For those tunnelling modes, substituting Eq. (2.38) into Eq. (6.2) gives

$$\gamma_p = \gamma_1 + \frac{(\gamma_2 - \gamma_1)}{[(\lambda - 1)^2 - \lambda^{2.5}]^{1/2}}. \quad (6.5)$$

This result is identical with that found from elementary plane wave theory for the case of "paraxial" rays using only Fresnel's equations and thus provides an independent check on the validity of the generalized $\eta_p$ defined by Eq. (1.17).

5.7 RÉSUMÉ

In this chapter, the basic characteristics of leaky modes have been calculated, both numerically for low-order modes and asymptotically for higher-order modes. These characteristics have then been used to study the excitation and propagation of the radiation field within the core of an optical fibre. This same problem has previously been investigated using the leaky ray approximation so comparison of the results obtained by the two different methods can be used to estimate the range of validity of the leaky ray theory. On making this comparison, it is found that leaky rays provide a good description of the power propagating along a typical multimode fibre excited by an incoherent source but that slightly more stringent conditions are
required in the case of coherent excitation because of the effects of diffraction.

A method for calculating the number of leaky modes with attenuation less than a given amount on a fibre of either homogeneous or inhomogeneous refractive index profile has also been reviewed and the effect of material absorption on the modal attenuation briefly discussed.
APPENDIX 5A

ASYMPTOTIC FORMS IN THE NEAR DEBYE LIMIT

In this section, asymptotic forms are found for the function

\[ \xi_\nu(x) = \frac{H^{(2)}_\nu(x) H^{(2)}_{\nu-2}(x)}{H^{(2)}_{\nu-1}(x)^2} \]  

\[ = \left( \frac{\nu-1}{x} \right)^2 - \left( \frac{H^{(2)}_\nu(x)}{H^{(2)}_{\nu-1}(x)} \right)^2 \]

in the limits (i) \( \nu \gg 1, \nu-x \gg \nu^{1/3} \)

and (ii) \( x \gg 1, x-\nu \gg \nu^{1/3} \).

(i) \( \nu \gg 1 \) and \( \nu-x \gg \nu^{1/3} \)

From Eqs. (9.3.8-13) of Ref. 47 we can write

\[ \frac{Y^{(2)}_\nu(x)}{Y^{(1)}_\nu(x)} = -\sin \alpha \cdot \frac{A}{B} \]  

where \( Y_\nu, Y^{(2)}_\nu \) are respectively the Bessel function of the second kind and its derivative

\[ A = 1 - \frac{v_1}{\nu} + \frac{v_2}{\nu^2} - \frac{v_3}{\nu^3} + O(\nu^{-4}) \]  

\[ B = 1 - \frac{u_1}{\nu} + \frac{u_2}{\nu^2} - \frac{u_3}{\nu^3} + O(\nu^{-4}) \]

\[ \sinh \alpha = \left( \nu^2 - x^2 \right)^{1/2}/x \]

and the \( u_i \) and \( v_i \) are polynomials in \( t = \coth \alpha \) given by Eqs. (9.3.9) and
(9.3.13). By binomial expansion,

\[
\frac{A}{B} = 1 + \frac{u_1 - v_1}{v} + \frac{u_1(u_1 - v_1) - (u_2 - v_2)}{v^2}
\]

\[+ \frac{u_1(v_2 - u_1v_1 - u_2 + u_1^2) - u_2(u_1 - v_1) + u_3 - v_3}{v^3} + O(v^{-4}). \quad (A.7)
\]

After a little algebra, Eq. (A.7) reduces to

\[
\frac{A}{B} = 1 - \frac{x^2}{2(v^2 - x^2)^{3/2}} + O[(v^2 - x^2)^{-2}], \quad (A.8)
\]

so that, combining Eqs. (A.3, 6 and 8) we have

\[
\frac{H_{\nu}^{(2)\prime}(x)}{H_{\nu}^{(2)}(x)} \approx \frac{Y_{\nu}^{\prime}(x)}{Y_{\nu}(x)} = - \frac{(v^2 - x^2)^{1/2}}{x} \left\{ 1 - \frac{x^2}{2(v^2 - x^2)^{3/2}} + O[(v^2 - x^2)^{-2}] \right\}
\]

\[= - \frac{(v^2 - x^2)^{1/2}}{x} + \frac{x}{2(v^2 - x^2)} + O[(v^2 - x^2)^{-3/2}] \quad (A.9)
\]

Finally, from Eqs. (A.2) and (A.9) (with \(v + v-1\)),

\[
\xi_{\nu} (x) = 1 + [(v-1)^2 - x^2]^{1/2} + O[(v^2 - x^2)^{-1}] \quad (A.10)
\]

or

\[
\xi_{\nu}^{-1} (x) = 1 - [(v-1)^2 - x^2]^{1/2} + O[(v^2 - x^2)^{-1}] \quad (A.11)
\]

(ii) \(x >> 1\) and \(x-v >> v^{1/3}\)

When the argument of the Hankel functions is no longer small, we need to retain both the real and imaginary parts of the Hankel functions. We now use Eqs. (9.3.15-22) of Ref. 47 to write

\[
\frac{H_{\nu}^{(2)\prime}(x)}{H_{\nu}^{(2)}(x)} = \frac{J_{\nu}^{\prime}(x) - iY_{\nu}^{\prime}(x)}{J_{\nu}(x) - iY_{\nu}(x)}
\]

\[= - i \sin \beta \left[ \frac{N - iO}{L + iM} \right], \quad (a.12)
\]

where \(\sin \beta = (x^2 - v^2)^{1/2}/x\),
\[ N - i\Omega = 1 - \frac{iv_1}{\nu} + \frac{v_2}{\nu^2} - \frac{iv_3}{\nu^3} + O(\nu^{-4}) \]  
\[ \text{(A.13)} \]

and
\[ L + i\Omega = 1 + \frac{iu_1}{\nu} + \frac{u_2}{\nu^2} + \frac{iu_3}{\nu^3} + O(\nu^{-4}) \]  
\[ \text{(A.14)} \]

\( u_i \) and \( v_i \) now being polynomials in \( \cot \beta \).

Correct to order \( x^{-1} \), binomial expansion in Eq. (A.12) now gives
\[ \frac{N - i\Omega}{L + i\Omega} = 1 - \frac{ix^2}{2(x^2 - \nu^2)^{3/2}} + O(x^{-2}) \]  
\[ \text{(A.15)} \]

whence
\[ \frac{H_{\nu}^{(2)'}}{H_{\nu}^{(2)}}(x) = -\frac{i(x^2 - \nu^2)^{1/2}}{x} - \frac{x}{2(x^2 - \nu^2)} + O(x^{-2}) \]  
\[ \text{(A.16)} \]

So from Eqs. (A.2) and (A.16) (with \( \nu + \nu - 1 \)),
\[ \xi_{\nu}(x) = 1 - i[x^2 - (\nu - 1)^2]^{-1/2} + O(x^{-2}) \]  
\[ \text{(A.17)} \]

or
\[ \xi_{\nu}^{-1}(x) = 1 + i[x^2 - (\nu - 1)^2]^{-1/2} + O(x^{-2}) \]  
\[ \text{(A.18)} \]
REFERENCES


34. J.D. Love (Personal communication).


37. C. Pask (Personal communication).


CHAPTER 6
THE ROLE OF UNBOUND MODES IN
VISUAL PHOTORECEPTORS

6.1 INTRODUCTION

To conclude, we shall consider the application of some of the ideas introduced in the preceding chapters to the study of light propagation in visual photoreceptors.

The fact that photoreceptors are capable of behaving as light guides has been known for at least 130 years, since the work of Brücke. But it was apparently not until 1948, following attempts by a number of workers to explain the Stiles-Crawford effects using geometric optics, that Toraldo di Francia pointed out that, because photoreceptors have radii of the order of a wavelength of light, the wave nature of light plays a significant role in these structures.

Since that time, the waveguide behaviour of photoreceptors has been experimentally verified by the observation of mode patterns at the ends of both the outer segments of receptors in vertebrate retinae and the rhabdoms of insect photoreceptors.

Such observations, and the variety of intriguing phenomena encountered in the study of visual systems, have aroused a great deal of

† Ernst Wilhelm von Brücke was one of the leaders of the "biophysical movement" which had its beginnings in Berlin in 1847 and also included Ludwig, du Bois-Reymond and Helmholtz.
theoretical interest in photoreceptor optics and, particularly, in the application of waveguide concepts to these minute optical fibres (see, for example, the many articles in Ref. 12). But while the principles of light propagation are the same in both photoreceptors and glass fibres, the practice in the case of the former is rather more complicated. Although the photoreceptors of most animals with a high degree of acuity and sensitivity are roughly long, narrow cylinders, many are far from circular in cross section. They are birefringent and dichroic and, in the case of insect rhabdoms, may consist of several sub-units (rhabdomeres) containing different pigments, all fused into a composite unit. Moreover, their properties may change with the level of illumination.

Nevertheless, a considerable degree of success has been achieved in investigations of such things as polarization-, angular- and spectral-sensitivities, the effects of birefringence and dichroism, and optical cross talk (see, e.g. Ref. 13). Many of these investigations lend support to the idea that waveguide effects play a functionally significant role in vision, particularly in setting the limiting optical performance of photoreceptors.

In general, it is the bound-mode field which provides the major contribution to light absorption in a photoreceptor and so, in all the theoretical studies referred to above, attention has been directed exclusively at this portion of the field, assuming the unbound field to be negligibly small. But in the relatively short fibres involved in

† Rather than attempt to give an exhaustive list of references to sources containing information on each particular aspect of the visual systems, we refer the reader to Refs. 13-17 which contain comprehensive collections of papers on the subject.
visual systems, the unbound energy may remain inside the fibre for a significant proportion of its length and therefore contribute appreciably to energy absorbed by the photopigment.

This could explain the discrepancies which occur in calculations like those of Snyder and Pask\textsuperscript{18} concerning the Stiles-Crawford effect, for example, where bound mode theory predicts a much more "discontinuous" behaviour of the photoreceptors with changing frequency than is realized in practice – the discontinuities being associated with mode cutoffs. If one allows for the fact that energy does not immediately cease to flow in the photoreceptor once cutoff is reached but gradually leaks out, then a smoother, more realistic result is obtained.

The purpose of Section 6.2 is to estimate the contribution of the unbound modes to light absorption in an ideal photoreceptor of circular cross section. Because of the short distances involved, the exact, rather than leaky-mode, analysis is used.

In Section 6.3 we then discuss a situation in which the unbound modes, because of their "smoothing" effect, make a definite negative contribution in terms of convenience – by obscuring the modal cutoff frequencies and thus preventing the use of cutoff observations to determine $V$ experimentally.

Although we have investigated some of the effects of birefringence and dichroism in photoreceptors,\textsuperscript{19-21} a discussion of these would extend beyond the scope of the present thesis so the following sections are concerned with photoreceptors which, apart from having a slight loss, are ideal.
6.2 CONTRIBUTION OF UNBOUND MODES TO LIGHT ABSORPTION IN VISUAL PHOTORECEPTORS

In order to determine the light absorption in a photoreceptor exactly, it would be necessary to find the eigenvalues and eigenfunctions of the structure with the dielectric constant being given an imaginary part to account for the absorption. Such an analysis has been performed by Röhler and Fischer for the bound modes on an axially homogeneous photoreceptor of circular cross section.

But visual photopigments absorb a very small amount of energy in a distance comparable to the wavelength of light. The electromagnetic fields of an absorbing photoreceptor are therefore, to a very good approximation, identical with those of a non-absorbing cylinder, apart from the inclusion of a slow-exponential-decay factor. This slight difference in the fields, on the one hand leads to numerical difficulties of the same type as those encountered in solving the eigenvalue equation for leaky modes near cutoff but, on the other hand, allows a perturbation analysis to be used to determine the light absorption approximately.

We obtain an approximate expression for the power absorbed in the photoreceptor by assuming that the differential change in power, \( \text{d}P_{\text{in}} \), due to absorption in a differential length, \( \text{d}z \), is related to the power inside by

\[
\text{d}P_{\text{in}}(z) = -\xi(z) P_{\text{in}}(z) \text{d}z ,
\]

where \( \xi(z) \) is the photoreceptor absorption coefficient. In the case of the bound field where the only change in \( P_{\text{in}} \) with \( z \) (apart from minor fluctuations due to beating between modes) is due to this absorption, Eq. (2.1) can be integrated to give the total power absorbed in a photoreceptor of length \( L \) from the bound field, \( \Omega_B(L) \), as
Of course this is only an approximation because, in the presence of lossy material, there is a net energy flow in the radial direction which will also contribute to the change in \( P_{in}(z) \). But for small absorptions, this can be neglected. If we also assume that \( \xi \) does not vary along the photoreceptor, then Eq. (2.2) becomes

\[
Q_B(L) = P_{in}^B(0) \left[ 1 - e^{-\xi L} \right].
\]  (2.3)

The corresponding expression for power absorbed from the unbound field, again assuming small, \( z \)-independent absorption coefficient \( \xi \) is found by integrating Eq. (2.1) to give

\[
Q_U(L) = \xi \int_0^L \left( P_{in}^{UB}(z) + P_{in}^{cross}(z) \right) dz,
\]  (2.4)

where \( P_{in}^{UB} \) and \( P_{in}^{cross} \) are defined by the second and third terms of Eq. (3.2.37), respectively.

The quantity in which we are interested is then the percentage contribution of \( Q_U(L) \) to the total power absorbed in length \( L \), that is,

\[
\frac{Q_U(L)}{Q_t(L)} \times 100,
\]  (2.5)

where

\[
Q_t = Q_U + Q_B.
\]  (2.6)

The calculation of \( Q_B \) is a straight-forward matter given the expressions for \( P_{in}^B \) as found in Chapter 3. But \( Q_U \) is not so easily calculated because the integral in Eq. (2.4) is rather difficult to
evaluate. An adequate approximation for our purposes can, however, be
obtained by graphically determining the area under curves for $P_{UB}$ and $P_{cross}$ as given, for example, in Figs. 3.4-3.7 for the case of plane wave sources.

Before doing any calculations, however, we must determine the values of the physiological parameters of interest.

The value of $V$ for most visual photoreceptors lies approximately within the range 0.5 to 10. In the fly eye, for example, rhabdomeres 1 to 6 have $V$ varying between 3.48 and 1.74 for wavelengths in the range 0.3 $\mu$m to 0.6 $\mu$m, whereas rhabdomeres 7 and 8 have $V$ in the range 1.74 to 0.87. Over the same range of wavelengths, $V$ for the rhabdom of the worker bee varies between approximately 3 and 6. Finally, in the human eye, rod outer segments have values of $V$ ranging from 3.9 to 2.2 for light of wavelengths between 0.4 $\mu$m and 0.7 $\mu$m.

The second parameter of interest is the length of the optical fibre involved. In the case of the fly, rhabdomeres 1 to 6 are approximately 220 $\mu$m in length (with a radius of 1 $\mu$m), rhabdomere 7 in approximately 130 $\mu$m (radius 0.5 $\mu$m) and rhabdomere 8 is 70 $\mu$m in length (with radius 0.5 $\mu$m). The length of the worker-bee rhabdom is about 350 $\mu$m (with a radius of about 2 $\mu$m) and the rod outer segment of the human eye varies in the range 40 - 80 $\mu$m (with a radius of about 0.5 $\mu$m).

The magnitude of the final parameter, $\xi$, is still open to question. Kirschfeld has suggested a rough estimate of 0.5%/ $\mu$m for the fly rhabdomere. In the bee, Shaw suggests a maximum absorption coefficient of about 1.8%/ $\mu$m. For vertebrate photoreceptors, a still-higher estimate of 3%/ $\mu$m has been obtained by Liebman and Entine.
Fig. 6.1: Percentage contribution of propagating unbound power to power absorbed in a photoreceptor as a function of $V$. The power absorbed within a distance $\lambda \rho$ of the receptor aperture is calculated where $\lambda = 20, 100, \text{and } 300$. If we assume the receptor radius, $\rho$, to be 1 $\mu$m then the absorption coefficient $\xi$ is $1/\mu$m.

Figures 6.1 and 6.2 show the results obtained by calculating expression (2.5) when the illuminating field is a finite portion of a plane wave of radius equal to the photoreceptor radius and normally incident at the aperture (i.e. corresponding to Figs 3.4 and 3.5).

Figure 6.1 shows the percentage contribution of unbound modes to the total power absorbed by a photoreceptor for $V$ in the range 1 to 3 and $L/\rho = \lambda$ in the range 20 to 300. The absorption coefficient in this case
Fig. 6.2: Percentage contribution of propagating unbound power to power absorbed in a length 100 $\rho$ of a photoreceptor. If the receptor radius, $\rho$, is taken to be 1 $\mu$m, the absorption coefficient varies between 0.1%/µm and 5%/µm. It is assumed to be 1%/µm (assuming also that $\rho = 1$ $\mu$m). The effect of these results of varying the absorption coefficient in the range $0.1\% \leq \xi \leq 5\%/\mu$m is shown in Fig. 6.2 for $\ell = 100$. The results are seen to be insensitive to this very large variation in absorption coefficient.

From Figs. 6.1 and 6.2, it would seem that unbound modes can only be expected to be important when $V \leq 1.5$ which, in terms of the examples mentioned above, excludes everything but rhabdomeres 7 and 8 of
the fly rhabdom (for which the contribution of unbound modes is over 30% of the power absorbed). But for a number of reasons, these curves represent considerable underestimates of the radiation field contribution.

First, we have ignored completely the contribution from evanescent modes of the continuous spectrum which, although they decay very rapidly, might be expected to make some contribution on a fibre of the size under consideration here.

Then there is the fact that a real photoreceptor is inhomogeneous and anisotropic and that light passes through a number of sections with different properties. A larger proportion of the propagated electromagnetic field will therefore be carried in the unbound modes than in an ideal fibre.

But the most important factor is the source we have used. It has already been noted in Chapter 3 that a significant increase in the amount of power launched into the unbound modes can be achieved by illuminating an area greater than the fibre aperture (Fig. 3.10) or by using off-axis illumination (Figs. 3.8 and 3.9). Yet another interesting possibility arises in the system illustrated in Fig. 6.2 where the photoreceptor radius changes (as in, for example, a mammalian retinal cone).

Consider the situation where the (larger) inner segment has $V \approx 3$ and the outer segment $V \leq 2.4$. Then if a $\text{TM}_{01}$ mode, for example, is propagating along the structure alone, once it reaches the "neck", the whole of the field is converted into radiation because the $\text{TM}_{01}$ mode is now below cutoff. We examine the radiation field propagating along the outer segment by using the portion of the $\text{TM}_{01}$ mode field within the outer segment radius as the source field in Eq. (3.2.3). With the
Fig. 6.3: Schematic representation of a mammalian cone receptor.

Fig. 6.4: Propagating unbound mode power within a photoreceptor excited by a truncated \( TM_{01} \) mode.
incident field normalized to unity, the resulting z-directed power is shown in Fig. 6.4. The difference between $P_{\text{tot}}^{UB}$ and $P_{\text{in}}^{UB}$ at $z=0$ is again because of evanescent modes.

By calculating the area under the curve in Fig. 6.4, we find that 12 times more power is absorbed from the propagating absorbed modes in this case than at the same value of $V$ for an on-axis plane wave source and even $2\frac{1}{2}$ times as much as when $V$ is reduced to unity with that source. Moreover, there is no bound mode contribution. So in systems like this, which appear to arise in discussions of the Stiles-Crawford effect, for example, the unbound modes cannot be ignored and their range of importance is not restricted to $V<1.5$ photoreceptors.

Finally we turn to another situation where unbound modes need to be considered.

6.3 AMBIGUITY IN DETERMINATION OF PHOTORECEPTOR PARAMETERS BY OBSERVATION OF MODE CUTOFFS

The waveguide properties of a photoreceptor depend on the value of $V$ which can be written in the form

$$ V = 2\pi \rho n_1 \frac{\theta_c}{\lambda}, \quad (3.1) $$

where

$$ \theta_c = \cos^{-1}\left(\frac{n_2}{n_1}\right), \quad (3.2) $$

$\rho, n_1, n_2$ have their usual meanings and $\lambda$ is the wavelength of light in vacuo.

It is extremely difficult to determine $V$ directly from Eq. (3.1) because of the difficulty in obtaining sufficient accurate values of $n_1$ and $n_2$. Since, in general, $n_1 \approx n_2$, small errors in the refractive
indices in Eq. (3.2) lead to large errors in $\theta_c$ and hence to large errors in $V$.

Kirschfeld and Snyder\textsuperscript{28} have measured $V$ directly in live fly eyes but the method is not applicable to vertebrate eyes. In fact there appears to be no reliable, direct determination of $V$ in vertebrate eyes, although indirect assessments of $n_1$ and $n_2$ can be made by estimating the concentration of protein and lipid in the photoreceptor membrane.\textsuperscript{21}

Enoch\textsuperscript{9} has proposed a method for determining $V$ by observing the frequency at which various modes become cut off, on the assumption that modes have well-defined cutoff characteristics associated with a particular value of $V$. Biernson and Kinsley\textsuperscript{35} have also discussed this approach in their paper on bound mode characteristics.

Figure 6.5, which shows the power of a number of low-order modes inside the core as a function of $V$ with varying $z/p$ underlies the fact that on short photoreceptors, such a determination can only be useful if the HE$_{21}$ (or TM$_{01}$, TE$_{01}$) mode is used. The other modes just do not have sharp cutoffs. They do fall off steeply below cutoff but the fall-off is far from vertical. There is therefore no unambiguous criterion for deciding when the mode is "cut off" except at large distances and errors of up to 25% may be made in the value of $V$.

6.4 RéSUMÉ

In this chapter we have briefly examined the possible role of unbound modes in contributing to both light absorption and experimental uncertainty in idealized visual photoreceptors. We have found that whether or not the unbound field makes a significant contribution to light absorption is dependent on the individual photoreceptor but also
Fig. 6.6: Modal power within a circular optical fibre as a function of $V$ and of the length of the fibre.

strongly dependent on the type of illumination. We have also found that care must be taken in interpreting mode cutoff observations because of the presence of leaky modes.
REFERENCES

1. E.W. von Brücke, Archiv für Anat. und Physiol. 11, 444 (1843); 14, 225; 387; 479 (1847) [quoted in Ref. 5].


28. K. Kirschfeld and A.W. Snyder, "Waveguide mode effects, birefringence and dichroism in fly photoreceptors", in Ref. 13, pp.56-77.


CHAPTER 7

SUMMARY AND CONCLUSIONS

The object of this thesis has been to study the radiation field propagating along a homogeneous, circular dielectric cylinder immersed in an infinite, homogeneous cladding, using both spectral (continuum) and non-spectral (leaky mode) representations.

To this end, we began in Chapter 2 by repeating Snyder's derivation of an orthonormal set of continuum modes using a decomposition of the radiation field into "incident" and "scattered" components. These modes have the advantages of being cast in the form of solutions of a well-known scattering problem, of separating naturally into two orthogonal sets determined by the nature of the incident fields and of obeying orthonormality conditions which apply to the much simpler incident fields. The last property follows from a general result we have proven linking the perturbed and unperturbed fields on any dissipationless, open, guiding structure.

In Chapter 3, the continuum modes were used to study the excitation and propagation of the radiation field on small-V fibres. The sources used were truncated plane waves and a quasi-monochromatic, totally (spatially) incoherent source. The plane wave results indicate that when the incident beam is parallel to the waveguide axis and has radius less than or equal to the core radius, very little power is launched into the unbound modes unless V is very small ($\xi < 2$). If the
beam is inclined at an angle to the axis or its radius is increased beyond the core radius, the amount of power in the unbound modes can be significantly increased. At small values of $V (~ 10)$, this unbound power leaves the core quite quickly however so it will probably only be important on fibres of the scale of visual photoreceptors.

The plane wave results also show that when the difference in dielectric constant between the core and cladding is small, the far field radiation pattern from the waveguide is very similar to the Fraunhofer diffraction pattern of a plane wave incident at a circular aperture — indicating that for the radiation field at infinity, the waveguide is only a small perturbation of free space.

When the plane waves are replaced by an incoherent source, the unbound mode power is confined to the waveguide core over somewhat larger distances but, once again, at these small values of $V$ it is insignificant on waveguides of the scale of optical fibres.

Unfortunately calculations using the spectral unbound modes have had to be restricted to values of $V$ smaller than ten by practical limitations. To evaluate the continuum mode power crossing a finite area, a double integration is required — and this becomes progressively more difficult as $V$ and the length of the fibre are increased. For this reason, we turned our attention in Chapter 4 to an alternative, asymptotic representation of the waveguide field.

The first step in deriving this asymptotic representation was to transform from the transverse spectral representation (in which we summed over the transverse component of the modal wave vectors) to a longitudinal one which is the Fourier transform in $z$, the longitudinal coordinate. In this way, the explicit separation between bound and
unbound modes was removed and an integral obtained which could be approximated using steepest-descent techniques. On transforming, in turn, to this steepest-descent representation, it was found that bound modes arise from poles corresponding to real solutions of the usual transverse resonance relation (or eigenvalue equation) but that contributions from complex poles must also be included within restricted regions of space which depend on the observation point. These complex poles correspond to improper, leaky modes which are the analytic continuations of bound modes below their cutoff frequencies.

The leaky modes decay exponentially in the axial direction and have bound-mode type radial behaviour out to some finite turning point but then begin to oscillate and grow to infinity. They are therefore not normalizable in the usual sense and power orthogonality cannot be used to determine the amplitudes of leaky modes excited by a particular source. This problem can however be overcome without the necessity to calculate residues in the continuous spectrum representation. One method is to define a new set of truncated modes which are artificially cut off at the turning point. Between weakly leaky modes the usual power orthogonality condition is then approximately satisfied and, for others, coupled mode theory can be used to obtain corrections to the amplitude coefficients obtained by assuming orthogonality. The alternative method is to define a more general form of inner product between the modal fields in which we integrate over a complex "radius" to introduce an exponential decay in the integrand. This approach loses the physical picture of power orthogonality but is exact and treats bound and leaky modes in the same way.
Of course the pole contributions do not give the total field—there is also a "space wave" contribution from the integral along the steepest-descent path. Except in the case where a pole lies close to the saddle point and complicates matters slightly, the space wave only has an algebraic decay in the z-direction, whereas the leaky modes have an exponential decay. However, the dominant leaky modes on an optical fibre have such small attenuation coefficients that their contribution does not fall below that of the space wave until a distance is reached where the whole of the radiation field within the fibre is insignificant. Thus after a relatively short distance (on the scale of communication fibres), the whole of the electromagnetic field within the fibre may be accurately approximated by a finite sum over discrete modes.

In Chapter 5, we quantified these ideas by calculating (both numerically and asymptotically) the detailed characteristics of leaky modes. These modes were then used to study the propagation of the electromagnetic field in an optical fibre illuminated by either truncated-plane-wave or totally incoherent sources as introduced in Chapter 3. The calculations were performed on fibres with \( V \leq 20 \).

Since ray analyses using similar sources have already been performed, a comparison of the results obtained by the two methods enabled us to estimate the ranges of \( V \) and the critical angle over which ray methods could be expected to be accurate.

We found that when an incoherent source was used, ray and mode calculations differed by at most 10% on a \( V = 20 \) fibre after a distance of \( z/p = 10^3 \) and by 15 - 20% on a \( V = 10 \) fibre, with the larger error occurring at the larger value of the critical angle. However, when the plane wave was used, it was found that the range of validity of the ray
analysis was reduced because diffraction effects then become important. For $V = 20$ and $\theta_c = 0.01$, the two methods agreed very well but in practical situations, $\theta_c$ would be somewhat larger and so a correspondingly larger value of $V$ would be necessary to ensure the accuracy of ray calculations.

We then outlined Gloge's method for estimating the number of leaky modes with attenuation coefficients in a given range based on the position of their turning points and concluded the chapter by discussing the additional mode attenuation due to material absorption.

Finally, the sixth chapter sought to estimate the role of unbound modes in visual photoreceptors. It was found that because of the small values of $V$ and small lengths involved, a significant proportion of the light absorbed by a photoreceptor may be absorbed from unbound modes. As a result, many of the discontinuous, resonance effects predicted to occur in visual systems by bound mode theory would, in fact, be smoothed and experimental results more closely reproduced. It was also found that this smoothing effect due to the presence of unbound modes has the disadvantage of "blurring" the cutoff frequency for most bound modes so that difficulties may arise in the experimental determination of the value of $V$ by observation of mode cutoffs.

To sum up then, this work has shown that

(i) in dielectric waveguides with $V \leq 10$, the radiation field leaves the core very quickly. But in short fibres where this field may nevertheless be important, its contributions can be calculated using the continuous spectrum of radiation modes.

(ii) in fibres with $V > 20$, the radiation field may persist in the core for quite large distances. When a diffuse source is used, the
radiation field (and in fact the whole of the waveguide field) may be calculated using modified geometric optics but when a collimated source is used, somewhat larger values of V (or very small critical angles) are required to ensure the validity of the ray theory.

(iii) in the intermediate region, leaky modes provide a very good approximation to the radiation field in the core except very close to the source where a space wave contribution must also be included (and very far from the source where both the leaky mode and space wave terms are insignificant but the latter dominates what little unbound field is present).

On the other hand, the work has not examined in detail the contribution of the evanescent continuum modes to the radiation field very close to the source nor the exact form of the space wave contribution to the leaky mode expansion. More importantly, it has not considered at all the effect of a finite cladding on the leaky mode analysis and, except in some very general results, has been confined to homogeneous media so that nothing has yet been said about the relative attenuations of leaky modes on graded- and step-index fibres. The last two questions are particularly worthy of further attention.