

Linear & Non-Linear Dispersive Waves in Optical Fibers

Tommy Zieba

Date, and any other information

Contents

Abstract	ii
1 Introduction	1
1.1 a subsection	1
2 First Section	1
2.1 Brief Description of Topic	1
3 Maxwell's Equations	2
3.1 Gauss's Law for Electric Flux	2
3.2 Gauss's Law for Magnetic Fields	3
3.3 Divergence Theorem & Stoke's Theorem	4
3.4 Faraday's Law	5
3.5 Ampere-Maxwell Law	6
3.6 Maxwell's Equations in Differential Form	8
3.7 Maxwell's Equations in Matter	8
4 Cylindrical Coordinates	9
5 Guided Modes in Cylindrical Waveguides	12
6 Whitham Summary	17
6.1 Chapter 1	17
6.2 Chapter 7	19
6.3 Sample Subsection	19
7 Conclusions	19
A Appendix title	19
A.1 First subsection of appendix	19
References	2

Abstract

abstract

1 Introduction

1.1 a subsection

Introduction

2 First Section

2.1 Brief Description of Topic

The physical concept of interest for this project will be *dispersion*. An optical fiber has specific physical properties that can be described with *Maxwell's equations* and the theory of *waveguides* may be used to interpret these equations under certain assumptions including linear dispersion. Within the theory of waveguides are the concepts of *total internal reflection* and the *Goss-Hanchen shift* which characterize the modal dispersion in a linear, non-dispersive medium. The other source of dispersive effects is the dependence of the refractive index on the frequency (which extends to amplitude dependence). This corresponds to the assumption of a 'fully' dispersive medium having non-linear dispersive effects. Dispersion within an optical fiber are interpreted with partial differential equations (PDEs). Other general points of discussion for this report are:

1. The theory of Maxwell's equations used to arrive at the equations of interest, which are analyzed in other sections.
2. Some theory of hyperbolic waves will be discussed and the concept of dispersive waves will be introduced and discussed.
3. The model used to interpret linear dispersive effects in step-index and graded-index fibers will be derive. The derivation follows a classical method based on the theory of cylindrical waveguides.
4. The TE, TM, Hybrid, and some of LP dispersion modes will be discussed for this derivation. The solutions involved in these guided modes and their relationship with dispersive effects need to be discussed.
5. For non-linear dispersive effects, we require the *non-linear Schrodinger equation* which is derived using Maxwell's equations.
6. Modulation theory is applied to the non-linear schrodinger equation in order to generalize the dispersion relation and other relations relavent to the study of non-linear effects in optical fibers.

7. Such non-linear dispersive effects include the 'balance-like' phenomenon between the group velocity and wave number distribution which produces soliton pulses as possible solutions.

3 Maxwell's Equations

The following four integral equations are known as Maxwell's equations:

$$\oint_S \vec{E} \circ \hat{n} da = \frac{q_{in}}{\varepsilon_0} \quad (3.1)$$

$$\oint_S \vec{B} \circ \hat{n} da = 0 \quad (3.2)$$

$$\oint_C \vec{E} \circ d\vec{l} = -\frac{d}{dt} \int_S \vec{B} \circ \hat{n} da \quad (3.3)$$

$$\oint_C \vec{B} \circ d\vec{l} = \mu_0(I_{enc} + \varepsilon_0 \frac{d}{dt} \int_S \vec{E} \circ \hat{n} da) \quad (3.4)$$

Maxwell's differential equations are required for discussing optical waveguide theory referred to in [4]. This section mainly follows [1] (and other references where required) with the motivation of deriving Maxwell's equations. These equations are derived in integral form, but may be written in differential form by using *Divergence theorem* and *Stoke's theorem*.

3.1 Gauss's Law for Electric Flux

Equation (3.1) is known as Gauss's law for electric fields. Equation (3.1) may be summarized by the following: "electric charge produces an electric field and the flux of that field passing through any closed surface is proportional to the total charge contained in that surface"(p.1)[1]. The electric field is defined by, "the electrical force per unit charge exerted on a charged object"... "thus the electric field may be defined by the relation,

$$\vec{E} = \frac{\vec{F}_e}{q_0} \quad (3.5)$$

where \vec{E} = electrical force on a small charge q_0 "(p.3)[1]. Equation (3.1) also defines a vector field in which, "field lines point in the direction of the field at each point in space"(p.3)[1], and \hat{n} is the unit normal vector that is perpendicular

to a surface S which implies that $\vec{E} \circ \hat{n}$ is the component of the electrical field perpendicular to that surface.

The flux of the electric field is exactly what is defined by the left hand side of (3.1), but for surfaces which are not closed. It is the magnitude of the electrical field passing perpendicularly through such a surface. The electrical flux is also denoted by Φ_E and is equivalent to the number of field lines penetrating the surface. Such densities are referred to as charge densities. For a closed surface, “the number of field lines penetrating that surface must be proportional to the total charge enclosed by that surface” (p.15)[1]. Thus, the electrical flux described by the right hand side of (3.1) can be interpreted using the density of field lines passing through the integral surface, called charge density. Assuming charge density is constant with units C/m^3 , the enclosed charge is defined as:

$$q_{enc} = \rho V = \int_V \rho dV \quad (3.6)$$

The constant of proportionality which describes the electrical flux is the remaining term in (3.1) denoted by ϵ_0 and is referred to as the permittivity of free space. “The permittivity of a material determines its response to an applied electrical field” (p.18)[1]. The subscript, 0, refers to the permittivity relative to vacuum. For optical fibers, the material in question is typically glass. Glass is a dielectric material, meaning that one must also consider “charges that are bound to the atoms of the material [and]”, ..., “the effect of bound charges can be understood by considering what happens when a dielectric is placed in an external electric field” (p.18)[1]. Positive charges are displaced in the direction following the external field passing through a dielectric and negative charges in the opposite direction. This induces an electric field within the dielectric. The induced field is necessarily smaller in magnitude than the external electric field caused by the displaced negative charges. In other words, displaced negative charges produce a smaller net charge enclosed within dielectric material. “The permittivity of a dielectric, ϵ , is often expressed as the relative permittivity” (p.19)[1],

$$\epsilon_r = \frac{\epsilon}{\epsilon_0} \quad (3.7)$$

where $\epsilon_0 = 8.8541878176 \times 10^{-12} \text{ C/Vm}$. Note that permittivity is obtained experimentally.

3.2 Gauss’s Law for Magnetic Fields

Equation (3.2) is known as Gauss’s law for magnetic fields and will be derived in this section. Paraphrasing from p.43 of [1]: Electrical and magnetic fields

differ fundamentally because negative and positive electrical charges can be separated, while magnetic poles only exist as inseparable positive and negative pairs.

The magnetic field \vec{B} (a.k.a. magnetic induction) is defined by the force experienced by a moving charged particle. Following chapter 8 of [2]:

“If both an electric field and a magnetic field are present, the total force on a moving charge is $\vec{F}_e + \vec{F}_m$

$$\vec{F} = q\vec{E} + q(\vec{v} \times \vec{B}) = q(\vec{E} + \vec{v} \times \vec{B}) \quad (3.8)$$

This force is known as the Lorentz force” (p.191).

In (3.8), subscripts denote the electric and magnetic force respectively, q is the charge, and $\vec{v} \times \vec{B}$ is the cross-product of a unit vector in the direction of motion with the magnetic induction. The magnetic induction, \vec{B} , is derived from the definition of the magnetic force experienced by a pair of charged particles with individual velocities. Note that charged particles only experience magnetic force in the component of its velocity that is in the direction of the magnetic field. The justification for (3.2) in [1] is that, “the magnetic force is perpendicular to the velocity at every instant, the component of the force in the direction of the displacement is zero, and the work done by the magnetic field is therefore always zero” (p.45)[1]. The left hand side of (3.2) represents the change in magnetic flux through a closed surface S . But, the right hand side is zero only for closed surfaces, which is discussed further when deriving (3.3).

In order to discuss (3.3) and (3.4), it is important to introduce the divergence theorem and Stoke’s theorem first because the definitions required to formulate these two theorems appear in these equations. These two theorems will also provide a way to write (3.1)-(3.4) in differential form.

3.3 Divergence Theorem & Stoke’s Theorem

Definition 3.1. The **gradient** is a vector valued differential operator which generalizes the derivative for multivariate functions. It is defined by,

$$\vec{\nabla} \equiv \vec{i} \frac{\partial}{\partial x} + \vec{j} \frac{\partial}{\partial y} + \vec{k} \frac{\partial}{\partial z} \quad (3.9)$$

Definition 3.2. The **divergence** of a field \vec{A} is defined by the projection of the field onto the gradient for each infinitesimal point in space. This is equivalent to the dot product of the gradient with \vec{A} given by,

$$\text{div}(\vec{A}) = \vec{\nabla} \circ \vec{A} = \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \oint_S \vec{A} \circ \hat{n} da \quad (3.10)$$

Theorem 3.1. (Divergence Theorem) “The flux of a vector field through a closed surface S is equal to the integral of the divergence of the field over a volume V for which S is a boundary”(p.114)[1]. This may be written for some field \vec{A} as,

$$\oint_S \vec{A} \circ \hat{n} da = \int_V (\vec{\nabla} \circ \vec{A}) dV \quad (3.11)$$

Next, the curl and circulation of a vector field is defined so that Stoke’s theorem may be interpreted.

Definition 3.3. The **curl** of a vector field \vec{A} is a “measure of the vector field’s tendency to circulate about a point”(p.75)[1]. It is defined,

$$\text{curl}(\vec{A}) = \vec{\nabla} \times \vec{A} = \lim_{\Delta S \rightarrow 0} \frac{1}{\Delta S} \oint_C \vec{A} \circ d\vec{l} \quad (3.12)$$

Definition 3.4. The **circulation** of a vector field \vec{A} is the magnitude of cumulative force of the vector field along the direction and length of a closed path C . Formally this is may be written as,

$$\text{Circulation}(\vec{A}) = \oint_C \vec{A} \circ d\vec{l} \quad (3.13)$$

Theorem 3.2. (Stoke’s Theorem) “The circulation of a vector field over a closed path C is equal to the integral of the normal component of the curl of that field over a surface S for which C is boundary”(p.116)[1]. This may be expressed by the equation,

$$\oint_C \vec{A} \circ d\vec{l} = \int_S (\vec{\nabla} \times \vec{A}) \circ \hat{n} da \quad (3.14)$$

Remark. The derivations for the gradient, divergence, and curl and the proofs for the Divergence and Stoke’s theorems can be found in chapter 1 of [2].

3.4 Faraday’s Law

Equation (3.3) is called Faraday’s law. The equation may be summarized by a “changing magnetic flux through a surface [which] induces an emf [electromotive force] in any boundary path of that surface and [as a consequence of changing magnetic flux,] a changing magnetic field induces a circulating electric

field”(p.59)[1].

Note that the left hand side of (3.3) defines the emf and \vec{E} is an induced electric field in this context which only differs from a static electrical field by its structure. In other words, “charge-based electric fields have field lines that originate on positive charge and terminate on negative charge (and thus have zero divergence at those points), while induced electric fields produced by changing magnetic fields have field lines that loop back on themselves”...“(and thus have zero divergence)”(p.62)[1]. The emf represents the cumulative magnitude of force of the electrical field applied in the direction of the closed path C , over its length l . The emf may also be interpreted by the circulation of the electrical field as it is defined. A typical example of this is the voltage along the path of the wire.

On the other hand, the right hand side (ignoring the minus sign) of (3.3) represents the rate of change in flux with respect to time. Note that the magnetic flux through a non-closed surface is not necessarily zero as it is for (3.2). By definition of the static magnetic field described by (3.2), the field lines form closed loops and so a closed surface has equal amounts of magnetic force entering and exiting the surface producing zero net force enclosed by the surface. The magnitude and direction of the magnetic field \vec{B} or the spatial parameters of the surface S may change with respect to time. Hence, the rate of change in flux for a surface may be non-zero since parameters influencing the change in magnetic flux could be changing with the passage of time for an arbitrary surface. Also note that the surface S must contain the path C for this to make sense.

Equation (3.2) implies that the magnetic field is necessarily static at any given moment within *any* closed surface and so an emf necessarily opposes the changing magnetic flux, keeping the magnetic field static. Moreover, this accounts for the minus sign in (3.3) as the opposition of the emf against the changing magnetic flux.

3.5 Ampere-Maxwell Law

Equation (3.4) is called the Ampere-Maxwell law. It can be summarized by “an electric current or a changing electric flux through a surface produce[ing] a circulating magnetic field around any path that bounds that surface”(p.84)[1].

The constant $\mu_0 = 4\pi \times 10^{-7}$ Vs/Am is the permeability of free space. It is the magnetic analogue of ϵ_0 . “Magnetic permeability determines a material’s response to an applied magnetic field” and “the permeability of a magnetic material is often expressed by the relative permeability, $\mu_r = \mu/\mu_0$ ”(p.87)[1].

The enclosed electric current is denoted by I_{enc} and is defined by the net current passing through any surface enclosed by the path C . The enclosed current is independent of surface geometry, except that it must coincide with the path. Note that Ampere's law refers to the equation, $\oint_C \vec{B} \circ d\vec{l} = \mu_0 I_{enc}$ and applies only to steady currents where the change in electric flux with respect to time is zero.

The most common of three methods used for justifying the existence of the change in electric flux, $\frac{d}{dt} \oint_S \vec{E} \circ d\vec{l}$, is the inconsistency of Ampere's law while charging a capacitor. The other two methods involve conservation of charge or special relativity [1]. One may refer to (p.91-94)[1] to derive the change in electric flux using the first method. In that derivation, to account for the permittivity of free space, note that a changing electric flux induces a current on capacitors by the following from (p.93)[1],

$$\int_S \vec{E} \circ \hat{n} da = \int_S \frac{\sigma}{\epsilon_0} da = \frac{Q}{A\epsilon_0} \int_S da = \frac{Q}{\epsilon_0} \quad (3.15)$$

where σ is the area charge density over the surface area A , and Q is the current. Hence, in order to account for the induced current by the change in electric flux, it must be multiplied by the permittivity of free space. Moreover, the total current is given by $I_{enc} + Q$ in this formulation.

In order to convert (3.4) into differential form, it is necessary to define the electric current density. Doing so will also enable one to write I_{enc} as an integral. The current density, denoted by \vec{J} , is "defined as the vector current flowing through a unit cross-sectional area perpendicular to the direction of the current" (p.105)[1]. Formally this can be written as,

$$\vec{J} = n q \vec{v}_d \quad (3.16)$$

where n is the number of charge-carriers, q is the charge per carrier, and \vec{v}_d is the average drift velocity of charge-carriers. "If \vec{J} is nonuniform and not perpendicular to the surface" (p.106)[1], then,

$$I = \int_S \vec{J} \circ \hat{n} da \quad (3.17)$$

where \hat{n} is the unit normal vector for da . One may justify this equation for current by considering 'small' areas of the surface represented by da of which the vector current may be computed for. Summing the components of vector currents perpendicular to da 's ad infinitum results in the integral. Finally, assuming that the circulation of a magnetic field is proportional to the current, we may conclude

that (3.4) indeed describes the circulation of the magnetic field by considering both steady state and time dependent sources of current.

3.6 Maxwell's Equations in Differential Form

The following are Maxwell's equations in differential form:

$$\vec{\nabla} \circ \vec{E} = \frac{\rho}{\varepsilon_0} \quad (3.18)$$

$$\vec{\nabla} \circ \vec{B} = 0 \quad (3.19)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (3.20)$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \left(\vec{J} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \quad (3.21)$$

Applying divergence theorem,(3.11), to the left hand side of (3.1) and taking a derivative with respect to a volume of the entire equality gives the charge volume density, proportional to the constant ε_0 . This also gives (3.18). Since there is a zero on the right hand side of (3.2) then applying (3.11) and differentiating gives (3.19) immediately.

Now Stoke's theorem may be applied to (3.3) by rewriting the contour integral as a surface integral. Note that the differentiation with respect to time can be made partial and brought inside the integral. Doing so allows the differentiation of the entire relation with respect to the surface, which yields (3.20). The same procedure is applied to (3.4) to obtain its differential form (3.21), except one must first express the enclosed current by (3.17).

3.7 Maxwell's Equations in Matter

Whenever dealing with fields inside matter, one must consider two things:

1. The enclosed charge in (3.1) and the current density (3.16) includes both *bound* and *free* charge.
2. The enclosed current (3.17) and the displacement current, $\frac{\partial E}{\partial t}$, includes *bound*, *free*, and *polarization* currents.

Therefore, it is necessary to derive versions of (3.18) and (3.21) which depend only on free charge since it is not apparently easy to determine bound charges and currents in a general dielectric medium [1].

The following discussion is based on (p.10-14)[6] in order to derive the wave equation for the dielectric medium used in optical fibers. When considering Maxwell's equations in a dielectric medium, one must account for the polarization and magnetization of the material. The resulting electric (polarization) and magnetic (magnetization) fields are proportional, by the respective constants of permittivity and permeability, to applied electric and magnetic fields. The superposition of proportioned applied electric and magnetic fields with these medium-specific field quantities is known as the electric displacement field and the magnetic intensity field.

i.e.

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P} \quad (3.22)$$

$$\vec{H} = \mu_0 \vec{B} + \vec{M} = \mu_0 \vec{B} \quad (3.23)$$

The relations characterizing these fields with respect to the applied electric and magnetic fields allow one to write (3.18) and (3.21) with respect to \vec{D} and \vec{H} as,

$$\vec{\nabla} \circ \vec{D} = \rho \quad (3.24)$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} = \frac{\partial \vec{D}}{\partial t} \quad (3.25)$$

Silica fibers are non-conductive and non-magnetic. Therefore, the enclosed current density \vec{J} and magnetism \vec{M} are zero. Isotropic material is one which does not change properties between points. An example of an anisotropic material in this context is a graded index fiber. If the fiber is isotropic (i.e. lossless and homogeneous) then the polarization, \vec{P} , is zero. For non-linear effects, one must consider up to third order susceptibility to derive the non-linear schrodinger equation.

4 Cylindrical Coordinates

To discuss the solutions to the dispersion relation in the context of waveguide theory from [4], we must assume the propagation of a uniform plane wave, having linear dispersion. Otherwise, if the source produces a signal which propagates a non-uniform plane wave then we would expect non-linear dispersive effects due

to the amplitude dependence. In any case we should interpret solutions in cylindrical coordinates based on the structure of silica fiber waveguides. Also note that the direction of propagation, \hat{z} , is fixed for plane wave solutions of the electromagnetic field along the length of the fiber. This section gives a derivation of the curl of a vector field in cylindrical coordinates to be used in the following sections for interpreting and applying Maxwell's equations.

Begin by defining $x = r \cos\theta$, $y = r \sin\theta$, $z = z$. Then it follows that, $x^2 + y^2 = r^2(\cos^2\theta + \sin^2\theta) = r^2$ and $y/x = \tan\theta$, which implies $\sqrt{x^2 + y^2} = r$ and $\tan^{-1}(y/x) = \theta$. A cartesian vector of unital magnitude \vec{u} is given in cylindrical coordinates by,

$$\vec{u} = \begin{pmatrix} r \cos\theta \\ r \sin\theta \\ z \end{pmatrix} \quad (4.1)$$

and so the cylindrical unit vectors may be derived for each direction \hat{r} , $\hat{\theta}$, \hat{z} by applying the operation,

$$\hat{v} = \frac{\frac{d\vec{u}}{dv}}{\left|\frac{d\vec{u}}{dv}\right|} \quad (4.2)$$

to $v = r, \theta, z$ respectively. Then,

$$\begin{aligned} \bullet \hat{r} &= \frac{\frac{d\vec{u}}{dr}}{\left|\frac{d\vec{u}}{dr}\right|} = \frac{1}{\sqrt{\cos^2\theta + \sin^2\theta}} \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \\ \bullet \hat{\theta} &= \frac{\frac{d\vec{u}}{d\theta}}{\left|\frac{d\vec{u}}{d\theta}\right|} = \frac{1}{r\sqrt{\cos^2\theta + \sin^2\theta}} \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix} \\ \bullet \hat{z} &= \frac{\frac{d\vec{u}}{dz}}{\left|\frac{d\vec{u}}{dz}\right|} = \frac{1}{\sqrt{1^2}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

are the unit vectors associated with cylindrical coordinates. In order to define the curl in cylindrical coordinates, we first need to define the operation of the gradient in cylindrical coordinates. By definition, the gradient must satisfy

$$du = \vec{\nabla}u \circ d\vec{r} \quad (4.3)$$

where $u(r, \theta, z)$ is a static scalar field and the operation, “ d ”, is the total derivative and \vec{r} is a cartesian vector. By direct computation it follows that:

$$du = \frac{\partial u}{\partial r} dr + \frac{\partial u}{\partial \theta} d\theta + \frac{\partial u}{\partial z} dz \quad (4.4)$$

$$\vec{\nabla} u \circ d\vec{r} = (\vec{\nabla} u)_r (d\vec{r})_r + (\vec{\nabla} u)_\theta (d\vec{r})_\theta + (\vec{\nabla} u)_z (d\vec{r})_z \quad (4.5)$$

$$d\vec{r} = d[r\hat{r} + z\hat{z}] = \hat{r} dr + r d\hat{r} + \hat{z} dz + z d\hat{z} \quad (4.6)$$

$$\begin{aligned} \frac{\partial \hat{r}}{\partial r} &= \vec{0} & \frac{\partial \hat{r}}{\partial \theta} &= \hat{\theta} & \frac{\partial \hat{r}}{\partial z} &= \vec{0} \\ \frac{\partial \hat{\theta}}{\partial r} &= \vec{0} & \frac{\partial \hat{\theta}}{\partial \theta} &= -\hat{r} & \frac{\partial \hat{\theta}}{\partial z} &= \vec{0} \\ \frac{\partial \hat{z}}{\partial r} &= \vec{0} & \frac{\partial \hat{z}}{\partial \theta} &= \vec{0} & \frac{\partial \hat{z}}{\partial z} &= \vec{0} \end{aligned} \quad (4.7)$$

The relation, $\vec{r} = r\hat{r} + z\hat{z}$, in (4.6) is obtained by observing that any cartesian vector can be represented in this way for some r and z that are real constants. Applying (4.7) and (4.4) to (4.6) gives, $d\vec{r} = \hat{r} dr + \hat{\theta} r d\theta + \hat{z} dz$. Applying (4.6), in its new terms, to (4.5) gives the relation,

$$\vec{\nabla} u \circ d\vec{r} = (\vec{\nabla} u)_r (dr) + (\vec{\nabla} u)_\theta (r d\theta) + (\vec{\nabla} u)_z (dz) \quad (4.8)$$

which is applied with (4.4) to the original assumption, (4.3), to obtain

$$\vec{\nabla} \equiv \hat{r} \frac{\partial}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{z} \frac{\partial}{\partial z} \quad (4.9)$$

after equating the differential components of du with those of $\vec{\nabla} u \circ d\vec{r}$. Now we may define the curl in general by taking the cross product of $\vec{\nabla}$ with a vector field $\vec{A} = \vec{A}_r \hat{r} + \vec{A}_\theta \hat{\theta} + \vec{A}_z \hat{z}$. The curl in cylindrical coordinates is derived as follows:

$$\begin{aligned}
\vec{\nabla} \times \vec{A} &= \left(\hat{r} \frac{\partial}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{z} \frac{\partial}{\partial z} \right) \times \vec{A} \\
&= \left(\hat{r} \frac{\partial}{\partial r} \times \vec{A} \right) + \left(\hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} \times \vec{A} \right) + \left(\hat{z} \frac{\partial}{\partial z} \times \vec{A} \right) \\
&= \left(\hat{r} \times \frac{\partial \vec{A}}{\partial r} \right) + \left(\hat{\theta} \frac{1}{r} \times \frac{\partial \vec{A}}{\partial \theta} \right) + \left(\hat{z} \times \frac{\partial \vec{A}}{\partial z} \right) \\
&= \hat{r} \times \frac{\partial(A_r \hat{r} + A_\theta \hat{\theta} + A_z \hat{z})}{\partial r} + \hat{\theta} \frac{1}{r} \times \frac{\partial(A_r \hat{r} + A_\theta \hat{\theta} + A_z \hat{z})}{\partial \theta} + \hat{z} \times \frac{\partial(A_r \hat{r} + A_\theta \hat{\theta} + A_z \hat{z})}{\partial z} \\
&= [\hat{r} \times \hat{\theta}] \frac{\partial A_\theta}{\partial r} + [\hat{r} \times \hat{z}] \frac{\partial A_z}{\partial r} + [\hat{\theta} \times \hat{r}] \frac{1}{r} \left(\frac{\partial A_r}{\partial \theta} - A_\theta \right) + [\hat{\theta} \times \hat{z}] \frac{1}{r} \frac{\partial A_z}{\partial \theta} + [\hat{z} \times \hat{r}] \frac{\partial A_r}{\partial z} + [\hat{z} \times \hat{\theta}] \frac{\partial A_\theta}{\partial z} \quad (4.10)
\end{aligned}$$

$$= \hat{r} \left(\frac{1}{r} \frac{\partial A_z}{\partial \theta} - \frac{\partial A_\theta}{\partial z} \right) + \hat{\theta} \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) + \hat{z} \left(\frac{1}{r} \frac{\partial A_r}{\partial \theta} + \frac{A_\theta}{r} - \frac{\partial A_\theta}{\partial r} \right) \quad (4.11)$$

In the derivation above, (4.10) is obtained by expanding the partial derivative in the previous step and applying (4.7). Equation (4.11) is the curl given in cylindrical coordinates and follows from (4.10) by writing the cross products of unit vectors in terms of single unit vectors and collecting their coefficients.

5 Guided Modes in Cylindrical Waveguides

In a lossless and homogeneous dielectric medium, Maxwell's equations (3.25) and (3.20) reduce to,

$$\vec{\nabla} \times \vec{e}(t) = -\mu \frac{\partial \vec{h}}{\partial t} \quad (5.1)$$

and

$$\vec{\nabla} \times \vec{h}(t) = \varepsilon \frac{\partial \vec{e}}{\partial t}. \quad (5.2)$$

By assuming that the electric and magnetic fields have plane wave solutions $\vec{e}(t) = \vec{E} e^{i(\omega t - \beta z)}$ and $\vec{h}(t) = \vec{H} e^{i(\omega t - \beta z)}$ and assuming that $\varepsilon = \varepsilon_0 n^2$ and $\mu = \mu_0$, where n = “refractive index”, we can use equation (4.11) that was derived previously to get the system of equations required for the basis of TE, TM, and Hybrid modes. Such a basis depends on the cylindrical components of the electric and magnetic-intensity fields \vec{E} and \vec{H} respectively.

[INSERT] Need to show how the plane wave solutions $\vec{e}(t)$, $\vec{h}(t)$ are derived and identify which assumptions are made for deriving them. Need to relate these

types of solutions to the dispersion relation from the perspective of hyperbolic and dispersive waves.

[INSERT] Also need to show how to get the two equations (5.2) and (5.1) and identify which assumptions are made for deriving them.

The following derivation uses (4.11) with the plane wave solutions $\vec{e}(t)$ and $\vec{h}(t)$ applied to (5.2) and (5.1):

$$\begin{aligned}\vec{\nabla} \times \vec{E}e^{i(\omega t - \beta z)} &= \left[\frac{1}{r} \frac{\partial \vec{e}_z(t)}{\partial \theta} - \frac{\partial \vec{e}_\theta(t)}{\partial z} \right] \hat{r} + \left[\frac{\partial \vec{e}_r(t)}{\partial z} - \frac{\partial \vec{e}_z(t)}{\partial r} \right] \hat{\theta} + \left[\frac{1}{r} \frac{\partial \vec{e}_r(t)}{\partial \theta} + \frac{\vec{e}_\theta(t)}{r} - \frac{\partial \vec{e}_\theta(t)}{\partial r} \right] \hat{z} \\ &= \left[\frac{1}{r} \frac{\partial E_z}{\partial \theta} e^{i(\omega t - \beta z)} + i\beta \frac{\partial E_\theta}{\partial z} e^{i(\omega t - \beta z)} \right] \hat{r} \quad (5.3)\end{aligned}$$

$$+ \left[-i\beta \frac{\partial E_r}{\partial z} e^{i(\omega t - \beta z)} - \frac{\partial E_z}{\partial r} e^{i(\omega t - \beta z)} \right] \hat{\theta} \quad (5.4)$$

$$+ \left[\frac{1}{r} \frac{\partial E_r}{\partial \theta} e^{i(\omega t - \beta z)} + \frac{E_\theta}{r} e^{i(\omega t - \beta z)} - \frac{\partial E_\theta}{\partial r} e^{i(\omega t - \beta z)} \right] \hat{z} \quad (5.5)$$

$$\begin{aligned}-\mu \frac{\partial \vec{h}(t)}{\partial t} &= -\mu \frac{\partial (\vec{H} e^{i(\omega t - \beta z)})}{\partial t} = -\mu \left[i\omega H_r e^{i(\omega t - \beta z)} \right] \hat{r} \\ &\quad - \mu \left[i\omega H_\theta e^{i(\omega t - \beta z)} \right] \hat{\theta} \\ &\quad - \mu \left[i\omega H_z e^{i(\omega t - \beta z)} \right] \hat{z}\end{aligned}$$

By factoring out $e^{i(\omega t - \beta z)}$ from the two relations above and equating the components (5.3)-(5.5) to the respective components (5.6)-(5.8), the resulting differential equations are the following:

$$\frac{1}{r} \frac{\partial E_z}{\partial \theta} + i\beta \frac{\partial E_\theta}{\partial z} = \frac{1}{r} \frac{\partial E_z}{\partial \theta} + i\beta E_\theta = -\mu_0 i\omega H_r \quad (5.9)$$

$$-i\beta \frac{\partial E_r}{\partial z} - \frac{\partial E_z}{\partial r} = -i\beta E_r - \frac{\partial E_z}{\partial r} = -\mu_0 i\omega H_\theta \quad (5.10)$$

$$\frac{\partial E_\theta}{\partial r} + \frac{E_\theta}{r} - \frac{1}{r} \frac{\partial E_r}{\partial \theta} = \frac{1}{r} \left(\frac{\partial (r E_\theta)}{\partial r} - \frac{\partial E_r}{\partial \theta} \right) = -\mu_0 i\omega H_r \quad (5.11)$$

where (5.11) is rewritten using the product rule on $r E_\theta$ with respect to partial differentiation by r and (5.9) and (5.10) are rewritten by observing that partial derivatives with respect to z in the transverse directions θ and r of the fields

E_θ and E_r remain fixed when electric or magnetic-intensity wave propagation is assumed to be in the z direction. A similar procedure for (5.2) produces another set of three differential equations:

$$\frac{1}{r} \frac{\partial H_z}{\partial \theta} + i\beta H_\theta = i\varepsilon_0 \omega n^2 E_r \quad (5.12)$$

$$-i\beta H_r - \frac{\partial H_z}{\partial r} = i\varepsilon_0 \omega n^2 E_\theta \quad (5.13)$$

$$\frac{1}{r} \left(\frac{\partial(rH_\theta)}{\partial r} - \frac{\partial H_r}{\partial \theta} \right) = i\varepsilon_0 \omega n^2 E_z \quad (5.14)$$

The boundary conditions for these sets of coupled differential equations is that the electromagnetic fields are continuous where a change in refractive index occurs. In other words, the tangential components of the electric and magnetic-intensity field are equal between boundaries. In mathematical terms this means:

$$E_t^{(1)} = E_t^{(2)} \quad (5.15)$$

$$H_t^{(1)} = H_t^{(2)} \quad (5.16)$$

$\forall t \in \{r, \theta, z\}$. Now, we may follow chapter 3 of [4]. First, it is required to derive Maxwell's wave equations,

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r} \frac{\partial E_z}{\partial r} + \frac{1}{r^2} \frac{\partial^2 E_z}{\partial \theta^2} + (k^2 n(r, \theta)^2 - \beta^2) E_z = 0 \quad (5.17)$$

$$\frac{\partial^2 H_z}{\partial r^2} + \frac{1}{r} \frac{\partial H_z}{\partial r} + \frac{1}{r^2} \frac{\partial^2 H_z}{\partial \theta^2} + (k^2 n(r, \theta)^2 - \beta^2) H_z = 0. \quad (5.18)$$

Note that the partial derivatives in the wave equations above are only with respect to either r or θ . So, we may assume that the plane wave solutions of the electric and magnetic-intensity fields may be written with $\vec{E} = \overrightarrow{E(r, \theta)}$ and $\vec{H} = \overrightarrow{H(r, \theta)}$ because the electric and magnetic-intensity fields are independent of the propagation direction whenever it is a solution to the wave equations above. This is a consequence of the coupling of the equations (5.9)-(5.11) with (5.12)-(5.14). By isolating E_θ , E_r , H_θ and H_r in (5.9), (5.10), (5.12), and (5.13) respectively,

$$E_\theta = \frac{-i}{\beta} \left(-i\mu_0 \omega H_r - \frac{1}{r} \frac{\partial E_z}{\partial \theta} \right) \quad (5.19)$$

$$E_r = \frac{i}{\beta} \left(-i\mu_0\omega H_\theta + \frac{\partial E_z}{\partial r} \right) \quad (5.20)$$

$$H_\theta = \frac{-i}{\beta} \left(i\varepsilon_0\omega n^2 E_r - \frac{1}{r} \frac{\partial H_z}{\partial \theta} \right) \quad (5.21)$$

$$H_r = \frac{i}{\beta} \left(i\varepsilon_0\omega n^2 E_\theta + \frac{\partial H_z}{\partial r} \right) \quad (5.22)$$

which may be combined together by substituting E_θ , E_r , H_θ and H_r to give the following set of equations:

$$\begin{aligned} E_\theta &= \frac{-i}{\beta} \left(-i\mu_0\omega \left(\frac{i}{\beta} \left(i\varepsilon_0\omega n^2 E_\theta + \frac{\partial H_z}{\partial r} \right) \right) - \frac{1}{r} \frac{\partial E_z}{\partial \theta} \right) \\ &= \frac{-i}{\beta} \left(\frac{i\mu_0\varepsilon_0\omega^2 n^2}{\beta} E_\theta + \frac{\mu_0\omega}{\beta} \frac{\partial H_z}{\partial r} - \frac{1}{r} \frac{\partial E_z}{\partial \theta} \right) \\ \Rightarrow \left(1 - \frac{\mu_0\varepsilon_0\omega^2 n^2}{\beta^2} \right) E_\theta &= \frac{-i}{\beta} \left(\frac{\mu_0\omega}{\beta} \frac{\partial H_z}{\partial r} - \frac{1}{r} \frac{\partial E_z}{\partial \theta} \right) \\ \Rightarrow E_\theta &= \frac{-i}{(k^2 n^2 - \beta^2)} \left(\frac{\beta}{r} \frac{\partial E_z}{\partial \theta} - \omega\mu_0 \frac{\partial H_z}{\partial r} \right) \end{aligned} \quad (5.23)$$

$$\begin{aligned} E_r &= \frac{i}{\beta} \left(-i\mu_0\omega \left(\frac{-i}{\beta} \left(i\varepsilon_0\omega n^2 E_r - \frac{1}{r} \frac{\partial H_z}{\partial \theta} \right) \right) + \frac{\partial E_z}{\partial r} \right) \\ &= \frac{i}{\beta} \left(\frac{-i\mu_0\varepsilon_0\omega^2 n^2}{\beta} E_r + \frac{\mu_0\omega}{r} \frac{\partial H_z}{\partial \theta} + \frac{\partial E_z}{\partial r} \right) \\ \Rightarrow \left(1 - \frac{\mu_0\varepsilon_0\omega^2 n^2}{\beta^2} \right) E_r &= \frac{i}{\beta} \left(\frac{\mu_0\omega}{r\beta} \frac{\partial H_z}{\partial \theta} + \frac{\partial E_z}{\partial r} \right) \\ \Rightarrow E_r &= \frac{-i}{(k^2 n^2 - \beta^2)} \left(\beta \frac{\partial E_z}{\partial r} + \frac{\omega\mu_0}{r} \frac{\partial H_z}{\partial \theta} \right) \end{aligned} \quad (5.24)$$

Similarly,

$$H_\theta = \frac{-i}{(k^2 n^2 - \beta^2)} \left(\frac{\beta}{r} \frac{\partial H_z}{\partial \theta} + \omega\varepsilon_0 n^2 \frac{\partial E_z}{\partial r} \right) \quad (5.25)$$

and

$$H_r = \frac{-i}{(k^2 n^2 - \beta^2)} \left(\beta \frac{\partial H_z}{\partial r} - \frac{\omega\varepsilon_0 n^2}{r} \frac{\partial E_z}{\partial \theta} \right) \quad (5.26)$$

where $k = \omega\sqrt{\mu_0\varepsilon_0}$. Now we have relations for E_θ , E_r , H_θ and H_r with respect to only H_z and E_z . Note that H_z and E_z are the electromagnetic field components in the direction of propagation, z , corresponding to the plane wave solution for

these respective fields. The transverse electric (TE) and transverse magnetic (TM) modes correspond to the assumptions $E_z = 0$ and $H_z = 0$ respectively. The Hybrid mode becomes more complicated in its analysis of solutions for H_z and E_z to deduce the dispersion relation and has the assumption that $E_z \neq 0$ and $H_z \neq 0$. The arguments made for the TE mode assumption may be easily applied to those for the TM mode as stated in [4]. So, these arguments will now be applied to the TM mode assumption and one may refer to those for the TE mode from [4]. Suppose $H_z = 0$. Then, (5.17) and (5.23)-(5.26) become:

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r} \frac{\partial E_z}{\partial r} + \frac{1}{r^2} \frac{\partial^2 E_z}{\partial \theta^2} + (k^2 n(r, \theta)^2 - \beta^2) E_z = 0 \quad (5.27)$$

$$E_\theta = \frac{-i}{(k^2 n^2 - \beta^2)} \frac{\beta}{r} \frac{\partial E_z}{\partial \theta} \quad (5.28)$$

$$E_r = \frac{-i}{(k^2 n^2 - \beta^2)} \beta \frac{\partial E_z}{\partial r} \quad (5.29)$$

$$H_\theta = \frac{-i}{(k^2 n^2 - \beta^2)} \omega \varepsilon_0 n^2 \frac{\partial E_z}{\partial r} \quad (5.30)$$

$$H_r = \frac{-i}{(k^2 n^2 - \beta^2)} \frac{\omega \varepsilon_0 n^2}{r} \frac{\partial E_z}{\partial \theta} \quad (5.31)$$

Before proceeding, we are required to assume that E_z and H_z are separable, i.e. $E_z(r, \theta) = R_1(r)T_1(\theta)$ and $H_z(r, \theta) = R_2(r)T_2(\theta)$ for some continuous functions R_1 , R_2 and T_1 , T_2 . In addition, another boundary condition is required on the the functions T_1 and T_2 , such that these are oscillatory functions, in order for the boundary value problem for the wave equations (5.17) and (5.18) to be well-posed. i.e. $T_1(\theta) = C_1 e^{i(\alpha_1 \theta + \gamma_1)}$ and $T_2(\theta) = C_2 e^{i(\alpha_2 \theta + \gamma_2)}$ for some real constants $C_1, C_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2 \in \mathbb{R}$. Under these assumptions and by letting

$$A = \frac{-i}{(k^2 n^2 - \beta^2)}, \quad B_1 = \frac{\mu_0 \omega}{r}, \quad B_2 = \frac{-\omega \varepsilon_0 n^2}{r},$$

it follows from (5.24) and (5.26) that,

$$\begin{aligned} E_r &= A \left(\beta T_2(\theta) R_2'(r) + B_1 R_1(r) T_1'(\theta) \right) \\ &= A \left(\beta C_2 e^{i(\alpha_2 \theta + \gamma_2)} R_2'(r) + i \alpha_1 B_1 R_1(r) C_1 e^{i(\alpha_1 \theta + \gamma_1)} \right) \\ &= A C_1 e^{i(\alpha_1 \theta + \gamma_1)} \left(\beta C_2' e^{i(\alpha_2' \theta + \gamma_2')} R_2'(r) + i \alpha_1 B_1 R_1(r) \right) \end{aligned} \quad (5.32)$$

and

$$\begin{aligned}
H_r &= A \left(\beta T_1(\theta) R_1'(r) + B_2 R_2(r) T_2'(\theta) \right) \\
&= A \left(\beta C_1 e^{(\alpha_1 \theta + \gamma_1)} R_1'(r) + i \alpha_2 B_2 R_2(r) C_2 e^{(\alpha_2 \theta + \gamma_2)} \right) \\
&= A C_1 e^{i(\alpha_1 \theta + \gamma_1)} \left(\beta R_1'(r) + i \alpha_2 B_2 R_2(r) C_2' e^{(\alpha_2' \theta + \gamma_2')} \right) \tag{5.33}
\end{aligned}$$

for some $C_1', C_2', \alpha_1', \alpha_2', \gamma_1', \gamma_2' \in \mathbb{R}$ such that $C_{1,2}' = C_{1,2} - C_{2,1}$, $\alpha_{1,2}' = \alpha_{1,2} - \alpha_{2,1}$, and $\gamma_{1,2}' = \gamma_{1,2} - \gamma_{2,1}$. Therefore, for the TM mode where $T_2(\theta) R_2(r) = 0$, the equations (5.32) and (5.33) become:

$$E_r = A T_1(\theta) \left(i \alpha_1 B_1 R_1(r) \right) \tag{5.34}$$

$$H_r = A T_1(\theta) \left(\beta R_1'(r) \right) \tag{5.35}$$

which implies that the *real* solution to the boundary value problem posed by (5.29) and (5.31) are,

$$G \tag{5.36}$$

for some functions $g_1(r) = i A \alpha_1 B_1 R_1(r)$ and $g_2(r) = A \beta R_1'(r)$.

6 Whitham Summary

6.1 Chapter 1

A particular non restrictive definition of a wave is “any recognizable signal that is transferred from one part of the medium to another with a recognizable velocity of propagation”, where, ”the signal may be any feature of the disturbance”,..., “provided that it can be clearly recognized and its location at any time can be determined”(p.2)[5]. The book by Whitham is divided between hyperbolic and dispersive waves. The hyperbolic waves arise from hyperbolic partial differential equations. Dispersive waves arise from restrictions on the type of solution one permits.

“The prototype for hyperbolic waves is often taken to be the wave equation,” (p.3)[5]

$$\varphi_{tt} = c_0^2 \nabla^2 \varphi \tag{6.1}$$

When dealing with a linear equation, “a linear dispersive system is any system which admits solutions of the form,

$$\varphi = a \cos(\kappa x - \omega t) \quad (6.2)$$

where the frequency ω is a definite real function of the wave number κ and the function $\omega(\kappa)$ is determined by the particular system. The phase speed is then $\frac{\omega(\kappa)}{\kappa}$ and the waves are usuallu said to be 'dispersive' if this pase speed is not a constant, but depends on κ . The term [dispersion] refers to the fact that a more general solution will consist of the superposition of several modes like 6.2 with different κ .” ... “If the phase speed $\frac{\omega}{\kappa}$ is not the same for all κ , that is $\omega \neq c_0 \kappa$ where c_0 is some constant, the modes with different κ will propagate at different speed; they will disperse.” (p.3)[5]

“Another example [for dispersion] is the classical theory for the dispersive effects of electromagnetic waves in dielectrics; this lead to [the dispersion relation],

$$(\omega^2 - \nu_0^2)(\omega^2 - c_0^2 \kappa^2) = \kappa^2 \nu_p^2 \quad (6.3)$$

where c_0 is the speed of light, ν_0 is the natural frequency of the oscillator, and ν_p is the plasma frequency”(p.9-10)[5]. For linear problems, Fourier series solutions can be obtained in the form,

$$\varphi_\kappa = \int_0^\infty F(\kappa) \cos(\kappa x - W(\kappa)t) d\kappa \quad (6.4)$$

where each value of κ corresponds to some value of $F(\kappa)$. Also note that $F(\kappa)$ can be obtained using the inverse Fourier transform and initial conditions. “The solution in (6.4) is a superposition of wavetrains of different wavenumbers, each traveling with its own phase speed,

$$c(\kappa) = \frac{W(\kappa)}{\kappa}. \quad (6.5)$$

As time evolves, these different component modes 'disperse', with the result that a single concentrated hump, for example, disperses into a whole oscillatory train. **This process is studied by various asymptotic expansions of (6.4).** The key concept that comes out of the analysis is that of the *group velocity* defined as,

$$C(\kappa) = \frac{dW}{d\kappa} \quad (6.6)$$

The oscillatory train arising from (6.4) does not have constant wavelength; the whole range of wave numbers κ is still present. In a sense to be explained, the different values of wave number propagate through this oscillatory train and

the speed of propagation is the group velocity (6.6)”(p.10)[5].

INSERT[Insert from (p.10-11) the derivation of the simple wave equation

$$\frac{\partial k}{\partial t} + C(k) \frac{\partial k}{\partial x} = 0 \quad (6.7)$$

where $k(x, t) \neq \kappa$ and $C(k)$ is the group velocity, **based on the assumption that the nonuniform oscillatory wave is described approximately in the form $\varphi = a \cos \theta$** . Relate this to the physics literature and show that Bessel’s functions provide solutions for guided modes where the modes are determined by assumptions on the axial components of the electric and magnetic intensity fields, E_z and H_z respectively as shown beginning on (p.48) of [4].]

6.2 Chapter 7

6.3 Sample Subsection

To include a graph, un-comment out the following text (remove the percentage signs), assuming that the graph is in a EPS file called graph1.eps that you have created beforehand using any plotting program. Adjust the numbers to change the size of the graph.

7 Conclusions

Summarize the content of your document

A Appendix title

A.1 First subsection of appendix

In the next section I have included a few references to show you how to include a list of references at the end of your report. To cite any of these books or articles any where in the text, you can either write [3], [4], which will give you the reference numbers [1], [2], or you can write “Author (Year)”, e.g. Abramowitz and Stegun (1964), Bacmeister and Pierrehumbert (1988).

References

- [1] Fleisch, Daniel A., 2008, *A Student’s Guide to Maxwell’s Equations*, Cambridge University Press.

- [2] Reitz, John R., Milford, Frederick J., Christy, Robert W., 2009 *Foundations of Electromagnetic Theory, 4th Ed.*, Pearson.
- [3] Knobel, Roger 2000, *An Introduction to the Mathematical Theory of Waves*, The American Mathematical Society.
- [4] Okamoto, Katsunari, 2000, *Fundamentals of Optical Waveguides*, Academic Press.
- [5] Whitham, G.B., 1999, *Linear and Non-Linear Waves*, Wiley.
- [6] Belanger, Pierre A., 1993 *Optical Fiber Theory*, World Scientific.