Coupled Mode Theory in Low-Contrast Nonlinear Photonic Crystals

Jamin L. Sheriff

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF APPLIED SCIENCE

DIVISION OF ENGINEERING SCIENCE
FACULTY OF APPLIED SCIENCE AND ENGINEERING
UNIVERSITY OF TORONTO

September 2003

Supervisor : Edward H. Sargent

Abstract

Nonlinear photonic crystals may enable all-optical signal processing. Developed herein is an analytical method by which the propagation of light in low-contrast nonlinear photonic crystals can be understood.

In solving Maxwell's equations in nonlinear photonic crystals, their constituent periodicity and nonlinearity are treated as perturbations to an underlying homogeneous medium. The method of multiple scales is then used to obtain coupled mode equations governing the evolution of resonantly coupled normal modes in the perturbed medium, thus providing an approximation to the electric field therein.

The method developed here applies to photonic crystals of any dimensionality and reproduces as a special case the one-dimensional theory of which it is a generalization. Being explicitly intended for low-contrast photonic crystals, this method provides a more direct means for their analysis than studying methods intended for high-contrast media in the limit of low-contrast. It is also more general than previous instances of multi-dimensional nonlinear coupled mode theory.

By facilitating the analysis and understanding of the behaviour of light in nonlinear photonic crystals, the formalism developed in this work may help expedite their application in all-optical signal processing.

Acknowledgements

I would foremost like to thank Professor Edward Sargent for his guidance, encouragement and for trusting me with a challenging and engaging problem.

I thank Professor Dmitry Pelinovsky of McMaster University, as well as Lukasz Brzozowski, Mathieu Allard and Emanuel Istrate for many insights.

Finally, I thank my family for their support. To them I can wholly attribute whatever success I may find.

Contents

List of Symbols and Abbreviations

$\mathbf{E}$	Electric field vector
$\mathbf{D}$	Electric flux density vector
n	Index of refraction
Δ	Total index perturbation
Δ^l, Δ^{nl}	Linear and nonlinear components of the index perturbation
$\Delta^l_{\mathbf{G}}, \Delta^{nl}_{\mathbf{G}}$	Fourier components of Δ^l and Δ^{nl}
ξ	Perturbation parameter
$\mathbf{k}$	Wavevector
ω	Angular frequency
$\varphi_{\mathbf{k},\omega}$	Plane wave with wavevector $\mathbf{k}$ and frequency ω
$\mathbf{A}_{\mathbf{k}}$	Mode envelope
$\mathbf{X}_F, T_F, \mathbf{X}_S, T_S$	Fast and slow time and space variables
∇_S, ∇_F	Gradients with respect to the slow and fast space variables
$\langle , \rangle$ or $\langle \vert \rangle$	Inner product
$\vert S\vert$	Cardinality of set S
CME	Coupled mode equations
MSM	Method of multiple scales
RLV	Reciprocal lattice vector

Chapter 1

Introduction

Telecommunications networks are the foundation of the Internet. Information in these networks is transmitted as light in optical fibres but processed electronically when it must be actively operated on. Routing, for example, is performed by high speed electronics appearing intermittently between lengths of optical fibre.

As demand for bandwidth grows and as electronic technology inexorably approaches fundamental limits in its size and so its speed, the need to eliminate it from telecom networks becomes increasingly urgent. An ideal network must be universally optical if it is to have any optical component at all.

This in turn requires a new paradigm of devices capable of performing optically and with equal efficacy the operations currently carried out electronically. Briefly, these devices must be capable of all-optical signal processing.

As will be discussed below, the properties of nonlinear photonic crystals make them suitable candidates on which these devices could be based and for this reason are the focus of this work. We first discuss the constitutive features of nonlinear photonic crystals, periodicity and nonlinearity, before proceeding to acknowledge how fruitful their combination can be and what bearing it may have on the creation of all-optical networks.

1.1 Periodic Media

In the context of photonics, a periodic medium is one whose dielectric permittivity varies periodically with position. This is expressed explicitly as $\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} + \mathbf{R_i})$ for $i = 1, 2, 3$. The vectors $\mathbf{R}_i$ are the basis for a lattice in which the periodic medium consists. Periodic dielectric materials are for this reason known as photonic crystals.

The effect of a dielectric lattice on the behaviour of light closely parallels the influence of periodic potentials on electrons in regular crystals. This similarity is manifest in the form assumed by the governing equations in each case, Maxwell's and Schrodinger's respectively. In a periodic dielectric structure, the stationary Maxwell's equations can be cast into the form of a eigenvalue problem that closely resembles Schrodinger's equation in a periodic potential and in which the similarities between periodic potentials and permittivities is made patent.

The wave equation in each case is amenable to the Bloch-Floquet theory and, as is well known from solid state physics [1], thereby admit stationary solutions with the form $\mathbf{u}(\mathbf{r})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ where the function $\mathbf{u}(\mathbf{r})$, scalar in Schrodinger's equation and vectorial in Maxwell's, has the periodicity of the dielectric or potential lattice so that $\mathbf{u}(\mathbf{r}) = \mathbf{u}(\mathbf{r} + \mathbf{R_i})$. Solutions possessing this form of a plane wave modulated by a periodic function are called Bloch modes. We immediately see that exacting the analogy between Schrodinger's and Maxwell's equations in periodic media has availed to the study of light the well established results of solid state physics and quantum mechanics.

Central to the study of Bloch modes and so to that of light in photonic crystals are dispersion relations. Dispersion relations in photonic crystals are constructed by determining the conditions on ω and $\mathbf{k}$ under which the Bloch mode $\mathbf{u}(\mathbf{r})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ does in fact solve Maxwell's equations. The dispersion relations are typically represented by ω as a function of $\mathbf{k}$ and are written as $\omega(\mathbf{k})$. It is the dispersion relations in which the salient property of periodic media is found, namely, bandgaps.

For particular lattice geometries there exist frequencies at which no light can propagate in the photonic crystal, a phenomenon that is manifest as gaps in the dispersion relations at these frequencies. For no value of $\mathbf{k}$ is the dispersion relation $\omega(\mathbf{k})$ equal to these frequencies. These frequencies collectively constitute the photonic bandgaps of the crystal.

The physical mechanism responsible for bandgaps lies in the periodicity of the dielectric permittivity in the medium. The resulting dielectric lattice can diffract electromagnetic waves in a manner similar to that in which x-rays are diffracted by semiconductor crystals. For particular wavelengths and directions, waves can interfere constructively with the diffracted waves that they create, in turn creating standing waves that cannot propagate.

Periodicity, however, is not sufficient for the formation of gaps in the dispersion relations. The existence of photonic bandgaps in fact depends strongly on the combination of the scattering induced by the dielectric lattice, called Bragg scattering, and that caused by the individual elements constituting the lattice [2]. Strictly, photonic crystals are dielectric lattices that do exhibit photonic bandgaps.

The photonic bandgap gives rise to novel and unique physical phenomenon and consequently have been singularly responsible for generating a wealth of research. One of the contexts in which it was first conceived was the inhibition of spontaneous emission [3], one result of which would be more efficient semiconductor lasers. Photonic crystals were simultaneously proposed as a means of localizing electromagnetic waves by introducing defects into them [4]. Modes whose frequencies lie in the bandgap of the crystal but that are allowed to exist in the defects, will be localized therein by their inability to propagate through the rest of the crystal. The evanescence of modes whose frequencies lie the bandgap similarly leads to applications such as antennae substrates [5] and waveguides [6, 7] that, through the bandgap, can very precisely control the propagation of light and so be substantially more efficient than their ordinary dielectric counterparts.

The applications mentioned here, however, have only a passive dependence on the photonic bandgap inasmuch as they rely only on its existence. It suffices in these contexts for the bandgap to be static and spatially uniform. Optical signal processing, however, requires some means of actively controlling the propagation of light in a medium, which in turn requires the band structure of the medium to be dynamic [8]. Linear photonic crystals cannot, for this reason, be suitable for all of the operations associated with optical signal processing.

We discuss next one means of obtaining the mutability of band structures that is essential to all-optical signal processing, namely, nonlinearities.

1.2 Nonlinear Media

Linear materials are appropriately characterized as having responses to applied electric fields, or more precisely, polarizations that depend linearly on those fields. However, when subject to fields with sufficiently high intensities, this dependence can, even in nominally linear materials, become nonlinear.

The details of the response of a material to an applied electric field can be determined from its susceptibility, which relates the polarization in the material to that field. This relationship can be written as a power series in the components of the field $\mathbf{E}$ [9],

$$P_i = \epsilon_0(\chi^{(1)}_{ij}E_j(\mathbf{r}) + \chi^{(2)}_{ijk}E_jE_k + \chi^{(3)}_{ijkl}E_jE_kE_l + \cdots)$$

in which it is clear that the nonlinear terms corresponding to $\chi^{(n)}$ for $n \geq 2$ cannot be ignored for high field intensities. The tensor $\chi^{(n)}$, with components $\chi_{ijk\cdots}$ is the n^{th} order susceptibility tensor which, in isotropic media, becomes a scalar. In what follows, we will be concerned only with isotropic media.

In a material possessing inversion symmetry, such that $\chi^{(n)}(\mathbf{r}) = \chi^{(n)}(-\mathbf{r})$, all even-ordered susceptibilities must vanish [10], making the third order susceptibility $\chi^{(3)}$ the lowest order nonlinearity. We focus here on materials possessing inversion symmetry and for which $\chi^{(3)}$ is the dominant nonlinearity, so that $\mathbf{P} \approx \epsilon_0(\chi^{(1)}\mathbf{E} + \chi^{(3)}|\mathbf{E}|^2\mathbf{E})$. Materials in which this is the case are said to possess Kerr nonlinearities.

While, in general, the effect of a Kerr nonlinearity is to produce nonlinear anisotropy [9], it, in isotropic media, introduces a polarization that depends on the local light intensity but that remains collinear with the applied field. The effects of the intensity dependence of the polarization are more conveniently embodied in the form of an intensity dependent index of refraction in the following way. Susceptibility is related to the index of refraction by $n = \sqrt{1+\chi}$. In a Kerr nonlinear material, $\mathbf{P} = \epsilon_0(\chi^{(1)}+\chi^{(3)}|\mathbf{E}|^2)\mathbf{E} = \epsilon_0\chi\mathbf{E}$, so that

$$\begin{aligned} n &= \sqrt{1+\chi^{(1)}+\chi^{(3)}|\mathbf{E}|^2} \\ &\approx n_l + n_{nl}|\mathbf{E}|^2 \end{aligned} \tag{1.1}$$

where $n_l = \sqrt{1+\chi^{(1)}}$ and $n_{nl} = \frac{1}{2}\frac{\chi^{(3)}}{1+\chi^{(1)}}$. A Kerr nonlinearity thus has the effect of introducing a nonlinear component, $n_{nl}|\mathbf{E}|^2$ into the index of refraction. Although third-order susceptibility nonlinearities are responsible for a variety of phenomenon such as third harmonic generation and intensity dependent absorption, the intensity dependence of the index of refraction is the consequence most relevant to this work.

Light in a Kerr nonlinear medium can, via that nonlinearity, induce local changes to the medium that in turn affect the propagation of light therein. Optical nonlinearities thus allow light to affect its own propagation by causing local changes in the dispersion relation of the medium it is in. That light can affect its own behaviour is manifest in the intensity dependent index of refraction obtained above. Specifically, light propagating through a nonlinear medium has a phase component that is intensity dependent. The light can thereby alter its own phase, an effect known as self-phase modulation (SPM). The phenomenon of cross-phase modulation (XPM) is the extension of this notion to situations involving multiple sources of light wherein each source can affect the phase of another.

SPM has the effect of introducing new frequency components into pulses propagating in nonlinear media, and thereby broadening their spectrum [11]. This in turn has the effect of compressing the pulse itself, inasmuch as the width of a pulse and that of its Fourier transform are inversely related. Dispersion, however, has the effect of introducing a time and wavevector dependent phase into the spatial Fourier transform of a pulse. In a nondispersive medium, this phase has the effect of translating the pulse without disturbing its shape. In general however, this phase has the effect of broadening the pulse as the ensemble of normal modes constituting the pulse travel at different speeds.

In a dispersive medium, the effects of SPM can counteract those of the dispersion and can, under certain conditions, balance precisely with them. The products of this balance are solitons, pulses capable of propagating through a medium without experiencing any changes to their shape.

There exists an important class of solitons consisting of those formed in media that are dispersive for no reason other than their periodicity. Solitons formed in a periodic nonlinear material are termed Bragg solitons and can propagate undisturbed through periodic structures as a result of the balance between SPM and the dispersion caused by periodicity [12].

Inasmuch as light can locally affect the dispersion relation of the medium it is in, it should be expected that light in a periodic nonlinear one be able to locally affect its band structure and in particular its bandgaps. Indeed, this is the process by which gap solitons [13, 14] are able to propagate in a nonlinear periodic structure. Gap solitons are Bragg solitons whose frequency content lies entirely in the bandgap of the periodic medium in which they are formed. Such a pulse would be unable to propagate inside a linear periodic structure because of this defining property. However, in a nonlinear periodic medium, a sufficiently intense pulse can change the bandstructure local to it enough to remove its frequency content from the bandgap, thereby allowing it to propagate as a gap soliton.

From the example of gap solitons it is clear that the combination of periodicity and nonlinearity leads to phenomena unique to their union and exclusive of each property separately. This fact has been exploited in proposals for fundamental optical devices such as limiters [15, 16] and switches [17, 18, 19, 20, 21, 22] from which more complicated operations, such as logic functions can be obtained, [23].

Nonlinearity may thus provide the control over, or tunability, of bandstructures of periodic materials that is essential to all-optical signal processing. This property of nonlinear periodic media has been the impetus for their study and that of the behaviour light within them. It is similarly what compelled this work.

1.3 Motivation

The behaviour of light in any medium is determined by the solution to Maxwell's equations therein. The combination of nonlinearity and periodicity, however, can leave Maxwell's equations analytically intractable. This difficulty is typically circumvented with the use of perturbative methods, for example, [24]. We describe below how and why we develop such a method that is explicitly intended for multi-dimensional low-contrast photonic crystals.

Weak nonlinearities are typically treated as perturbations to an underlying linear medium. Whatever the strength of the nonlinearity, it precludes the independent evolution of the normal modes of the unperturbed linear medium. However, for a sufficiently weak nonlinearity, the lengths over which these normal modes do interact will be greater than the length characterizing the linear medium, its periodicity [25]. Similarly, the time after which the normal modes would have interacted substantially will be greater than the characteristic rates at which these modes oscillate. Consequently, two distinct scales can be associated with the propagation of waves in weakly nonlinear media, one corresponding to the periodicity of the unperturbed medium, and the other to the interaction of its normal modes. For this reason, an asymptotic method suitable for the analysis of nonlinear periodic structures is the method of multiple scales (MSM) [24].

In the MSM, to the lowest order of approximation, the electric field is expressed as some linear combination of the normal modes of the unperturbed linear medium. Normally, in non-stationary perturbation theory [26], the coefficients in this combination are made to vary in space and time. While this is the case in the MSM, what distinguishes it from other perturbative methods is the formal introduction of new time and space scales over which the coefficients characteristically vary. The coefficients in the normal mode expansion can then be made to vary only on the scale on which the normal modes interact.

These coefficients, or envelopes, can be thought of as modulating the normal modes and it is in those envelopes that the effects of a periodic nonlinearity can be found. The envelopes can thereby be sufficient for understanding the nature of light propagation in periodic nonlinear media, for example, [27, 28, 29]. The MSM avails us of a means of determining these envelopes.

This method of studying the large scale behaviour of the unperturbed normal modes was initially employed in optics in the study of deep nonlinear one-dimensional gratings [24, 12], where the envelope modulating a Bloch mode of the periodic unperturbed medium was found to satisfy a nonlinear Schrodinger equation. More recently, the technique has been employed in studying high contrast nonlinear photonic crystals [27, 28].

1.4 Objective

The focus of this work is on linearly periodic media that addtionally possess periodic Kerr nonlinearities. In particular, we use the MSM to study coupled modes in low contrast nonlinear photonic crystals, those in which the change in the dielectric permittivity over one period of the crystal is small. Low contrast photonic crystals are currently of experimental interest [30] but have not previously been the explicit subject of any maximally general treatments.

Our objective is to develop, in the framework of the MSMs, a method that will yield coupled mode equations (CME) for isotropic low-contrast nonlinear periodic media having any geometry in one, two or three dimensions.

The CMEs derived here will, for the first time, permit analysis of the intensity-dependent behaviour of multidimensionally-periodic low-contrast nonlinear media. The applicability of the CMEs to multidimensional media will facilitate the extension of previous work in 1-d devices to higher dimensions.

The CMES derived here describe resonant light incident from any direction on a medium having any periodic index of refraction profile. This is in contrast to some previous works wherein the index profile was fixed [32, 33]. Also, because it is explicitly intended for low-contrast photonic crystals, the method presented here may provide a more direct means for their analysis than the alternative of considering high contrast methods [29] in the limit of low index contrast.

The details of the derivation of the CMES and some of their general consequences are presented in Chapter 2. For illustration, this formalism is applied in Chapter 3 to particular scenarios in three-dimensional media.

Chapter 2

A Perturbative Approach

In this chapter, we enlist the method of multiple scales in obtaining nonlinear coupled mode equations for any low-contrast nonlinear periodic structure. The central result of this chapter will be generalized couple mode equations whose final form depends only on the modes that are coupled and the Fourier series representation of the periodic index of refraction.

What is essentially required for the understanding of Kerr nonlinear periodic structures is an approximate solution to Maxwell's equations. The particular structures that we endeavour to study are those possessing small, periodic spatial variations in both the linear and nonlinear components of their indices of refraction, as given by (1.1).

It is by treating the small variation in the index of the structure as a perturbation to a homogeneous linear medium that we obtain an approximate solution to Maxwell's equations. The approximation consists in only considering the envelopes modulating the normal modes of interest. This treatment leads naturally to the use of the MSM.

2.1 Generalized Coupled Mode Equations

Inside the perturbed medium, that having the position dependent and nonlinear index of refraction, the electric field must satisfy the wave equation,

$$\nabla^2 \mathbf{E} - \nabla(\nabla \cdot \mathbf{E}) = \frac{n^2(\mathbf{x}, |\mathbf{E}|^2)}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{2.1}$$

which is obtained with the assumption that

$$||n^2(\mathbf{x}, |\mathbf{E}|^2) \frac{\partial^2 \mathbf{E}}{\partial t^2}|| \gg ||2 \frac{\partial \mathbf{E}}{\partial t} \frac{\partial n^2}{\partial t} + \mathbf{E} \frac{\partial^2 n^2}{\partial t^2}||$$

Equivalently, the components of the field must satisfy the projection of (2.1) onto their respective axes, which we denote by $j = 1, 2, 3$, so that,

$$\nabla^2 E_j - \sum_{n=1}^{3} \frac{\partial^2 E_n}{\partial x_j \partial x_n} = \frac{n^2(\mathbf{x}, |\mathbf{E}|^2)}{c^2} \frac{\partial^2 E_j}{\partial t^2} \tag{2.2}$$

Inside the perturbed medium, the index of refraction is written as

$$n(\mathbf{x}, |\mathbf{E}|^2) = n_0 + \Delta(\mathbf{x}, |\mathbf{E}|^2) \tag{2.3}$$

where n_0 is the index in the linear, homogeneous unperturbed medium and $\Delta(\mathbf{x}, |\mathbf{E}|^2)$ the index perturbation. The index perturbation Δ contains all of the position dependence and nonlinearity found in the perturbed index $n(\mathbf{x}, |\mathbf{E}|^2)$ and is written as $\Delta(\mathbf{x}, |\mathbf{E}|^2) = \Delta^l(\mathbf{x}) + |\mathbf{E}|^2 \Delta^{nl}(\mathbf{x})$. The terms $\Delta^l(\mathbf{x})$ and $|\mathbf{E}|^2 \Delta^{nl}(\mathbf{x})$ are respectively the linear and nonlinear components of the perturbation. That the perturbed medium is Kerr nonlinear is manifest in the second term.

The relevant quantity in the wave equation (2.1) is the square of the total index which, using (1.1) and the full expression for $\Delta(\mathbf{x},|\mathbf{E}|^2)$, becomes

$$\begin{aligned} n^2(\mathbf{x},|\mathbf{E}|^2) &= n_0^2+2n_0\Delta(\mathbf{x},|\mathbf{E}|^2)+\Delta^2(\mathbf{x},|\mathbf{E}|^2) \\ &= n_0^2+2n_0\big(\Delta^l(\mathbf{x})+\Delta^{nl}(\mathbf{x})|\mathbf{E}|^2\big)+\Delta^2(\mathbf{x},|\mathbf{E}|^2) \end{aligned} \tag{2.4}$$

The periodicity of $\Delta^l(\mathbf{x})$ and $\Delta^{nl}(\mathbf{x})$ in the perturbed medium implies the existence of vectors $\mathbf{T}_i, i = 1,2,3$ such that $\Delta^l(\mathbf{x}) = \Delta^l(\mathbf{x}+\mathbf{T_i})$ and $\Delta^{nl}(\mathbf{x}) = \Delta^{nl}(\mathbf{x}+\mathbf{T_i})$. These vectors, called lattice vectors, form the lattice constituting the periodic medium. An important consequence of the periodicity of this is that the components of the perturbation can be as Fourier series in the following way,

$$\begin{aligned} \Delta^l(\mathbf{x}) &= \sum_{\substack{all\mathbf{G}\\ \mathbf{G}\neq\mathbf{0}}} \Delta^l_{\mathbf{G}} e^{-i(\mathbf{G}\cdot\mathbf{x})} \\ \Delta^{nl}(\mathbf{x}) &= \sum_{all\mathbf{G}} \Delta^{nl}_{\mathbf{G}} e^{-i(\mathbf{G}\cdot\mathbf{x})} \end{aligned} \tag{2.5}$$

The vectors $\mathbf{G}$ are reciprocal lattice vectors (RLVs) defined in the usual way in terms of the $\mathbf{T}_i$ [1]. The RLV $\mathbf{G}=\mathbf{0}$ is omitted from the series for $\Delta^l(\mathbf{x})$ because $\Delta^l_{\mathbf{0}}$ represents the component of the linear perturbation that is constant in space. We assume that there exists no linear index mismatch between the perturbed and unperturbed media, so that $\Delta^l_{\mathbf{0}}$ is accounted for already by n_0.

The Fourier coefficients $\Delta^l_{\mathbf{G}}$ are given by [1]

$$\Delta^l_{\mathbf{G}} = \frac{1}{V_{pc}}\int_{pc} d^3r\,\Delta^l(\mathbf{r})e^{i(\mathbf{G}\cdot\mathbf{r})}$$

where pc denotes the primitive cell of the real lattice and V_{pc} its volume. The components $\Delta^{nl}_{\mathbf{G}}$ are given by an analogous expression. It is useful to note that, if the index perturbations are real and possess inversion symmetry then, inasmuch as the primitive cell of a lattice must always be symmetric about the origin of the lattice, the Fourier components are actually given by

$$\Delta^l_{\mathbf{G}} = \frac{1}{V_{pc}} \int_{\frac{pc}{2}} d^3 r \Delta^l(\mathbf{r}) \cos(\mathbf{G}\cdot\mathbf{r})$$

where $\frac{pc}{2}$ denotes half of the primitive cell.

That the perturbation $\Delta(\mathbf{x}, |\mathbf{E}|^2)$ is small can be quantified by the introduction of a perturbation parameter ξ, such that (2.4) becomes

$$n^2(\mathbf{x}, |\mathbf{E}|^2) = n_0^2 + 2n_0\big(\widetilde{\Delta}^l(\mathbf{x}) + \widetilde{\Delta}^{nl}(\mathbf{x})|\mathbf{E}|^2\big)\xi + \Delta^2\xi^2 \tag{2.6}$$

where $\widetilde{\Delta}^l$ and $\widetilde{\Delta}^{nl}$ are $O(1)$. In what follows, the symbols $\Delta^l(\mathbf{x})$ and $\Delta^{nl}(\mathbf{x})$ will be used to refer to $\widetilde{\Delta}^l$ and $\widetilde{\Delta}^{nl}$. Because we do not consider corrections to the field that are $O(\xi^2)$, the final term in (2.6) is simply omitted in the analysis that follows.

The solutions of the wave equation (2.1) are parameterized through ξ. Furthermore, when $\xi = 0$, those solutions are known to be plane waves, given that the unperturbed medium is charge free. Accordingly, the field is expanded in an asymptotic series about $\xi = 0$,

$$\mathbf{E}(\mathbf{r}, t) = \sum_{m=0}^{\infty} \mathbf{E}_m(\mathbf{x}, t)\xi^m \tag{2.7}$$

from which an approximation to the field can be obtained by truncating the series at a sufficiently high order. Because the series is defined as being asymptotic, increasingly higher order terms in the expansion introduce decreasingly large deviations from the field in the unperturbed medium. This will be important in obtaining the CME.

As is additionally done in the MSM, new time and space scales are defined using ξ, namely $\mathbf{X}_p = \xi^p\mathbf{x}$ and $T_p = \xi^p t$ for $p = 0, 1, 2 \ldots$. Because our analysis is limited to $O(\xi)$ corrections, only the scales corresponding to $p = 0$ and $p = 1$ require consideration. In particular, $\mathbf{X}_F = \mathbf{x}$ is the fast space variable corresponding to the periodicity of the index lattice, and $\mathbf{X}_S = \xi\mathbf{x}$ is the slow space variable corresponding to the interaction of the normal modes induced by the nonlinearity. Similarly, $T_F = t$ and $T_S = \xi t$ are the fast and slow time variables. The terms $\mathbf{E}_m(\mathbf{x}, t)$ will be assumed to be functions of all of these scales and will be written as $\mathbf{E}_m(\mathbf{X}_\beta, T_\beta)$ with the dependence on all of the scales implicit in the use of β.

Treating each of these new scales as independent variables gives the following expressions for the derivatives relevant to the wave equation (2.2),

$$
\begin{aligned}
\frac{\partial^2}{\partial x_i^2} &= \frac{\partial^2}{\partial X_{F,i}^2} + 2\xi \frac{\partial^2}{\partial X_{F,i} \partial X_{S,i}} + \xi^2 \frac{\partial^2}{\partial X_{S,i}^2} \\
\frac{\partial^2}{\partial x_i \partial x_j} &= \frac{\partial^2}{\partial X_{F,i} \partial X_{F,j}} + \xi \Big(\frac{\partial^2}{\partial X_{S,i} \partial X_{F,j}} + \frac{\partial^2}{\partial X_{F,i} \partial X_{S,j}} \Big) + \xi^2 \frac{\partial^2}{\partial X_{S,i} \partial X_{S,j}} \\
\frac{\partial^2}{\partial t^2} &= \frac{\partial^2}{\partial T_F^2} + 2\xi \frac{\partial^2}{\partial T_S \partial T_F} + \xi^2 \frac{\partial^2}{\partial T_S^2}
\end{aligned} \tag{2.8}
$$

It should be noted that the index perturbations $\Delta^l(\mathbf{x})$ and $\Delta^{nl}(\mathbf{x})$ are functions only of the fast variables, so that $\Delta^l(\mathbf{x}) = \Delta^l(\mathbf{X}_F)$ and $\Delta^{nl}(\mathbf{x}) = \Delta^{nl}(\mathbf{X}_F)$. The Fourier series representing each of the perturbations retains the same form as (2.5) but with $\mathbf{x}$ replaced by $\mathbf{X}_F$.

Equipped with the expansion (2.7) and the new derivatives (2.8), a recurrence equation for the $\mathbf{E}_m(\mathbf{X}_\beta, T_\beta)$ can be obtained by substituting the expansions into the wave equation (2.2) and then using the linear independence of the powers of ξ. As will be shown, the CME will be a direct consequence of the existence of a solution to these recurrence relations.

Beginning by collecting terms proportional to ξ^0 in (2.2), it is found that, to first order, the recurrence equation is the wave equation in the unperturbed medium,

$$
\sum_{n=1}^{3} \frac{\partial^2 E_{0,j}}{\partial X_{F,n}} + \sum_{n=1}^{3} \frac{\partial^2 E_{0,n}}{\partial X_{F,j} \partial X_{F,n}} = \frac{n_0^2}{c^2} \frac{\partial^2 E_{0,j}}{\partial T_F^2}
$$

which is equivalent to the vector equation,

$$\nabla_F^2 \mathbf{E}_0 - \nabla_F(\nabla_F \cdot \mathbf{E}_0) = \frac{n_0^2}{c^2}\frac{\partial^2 \mathbf{E}_0}{\partial T_F^2} \tag{2.9}$$

where $\nabla_F = \sum_{n=1}^{3} \mathbf{e}_n \frac{\partial}{\partial X_{F,n}}$. $\nabla_F^2, \nabla_F\cdot$ are defined similarly.

Equation (2.9) differs slightly from the wave equation expected in a charge-free homogeneous medium because of the presence of the $\nabla_F(\nabla_F \cdot \mathbf{E}_0)$ term. However, taking into account another of Maxwell's equations, it will now be shown that that term is identically zero. Using the expansion (2.7), the divergence of the electric field is given by

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \nabla \cdot \Big(\sum_{m=0}^{\infty} \mathbf{E}_m \xi^m \Big) \\
&= \nabla \cdot \mathbf{E}_0 + \nabla \cdot \Big(\sum_{m=1}^{\infty} \mathbf{E}_m \xi^m \Big) \\
&= \nabla_F \cdot \mathbf{E}_0 + \xi \nabla_S \cdot \mathbf{E}_0 + \nabla \cdot \Big(\sum_{m=1}^{\infty} \mathbf{E}_m \xi^m \Big)
\end{aligned}$$

The expansion $\nabla \cdot = \nabla_F \cdot + \xi \nabla_S \cdot$ has been used in the last line, where $\nabla_F \cdot = \sum_{n=1}^{3} \frac{\partial}{\partial X_{F,n}}$ and $\nabla_S \cdot$ is defined similarly.

Now, invoking the constitutive relation $\mathbf{D} = n^2(\mathbf{r}, |\mathbf{E}^2|)\mathbf{E}$, and noting that the absence of charge in the medium requires $\nabla \cdot \mathbf{D} = 0$, we find that

$$\begin{aligned}
\nabla \cdot n^2 \mathbf{E} &= 0 \\
&= 2n(\nabla n)\dot{\mathbf{E}} + n^2 \nabla \cdot \mathbf{E} \\
&= 2\xi n(\nabla_F n) \cdot \mathbf{E} + n^2 \Big(\nabla_F \cdot \mathbf{E}_0 + \xi \nabla_S \cdot \mathbf{E}_0 + \nabla \cdot \Big(\sum_{m=1}^{\infty} \mathbf{E}_m \xi^m \Big) \Big) \\
&= n_0^2 \nabla_F \cdot \mathbf{E}_0 + O(\xi)
\end{aligned}$$

From this it follows that $\nabla_F \cdot \mathbf{E} = 0$ identically. The third line above follows from the index being dependent only on the fast scale $\mathbf{X}_F$, so that $\nabla_S n = \mathbf{0}$.

Applying this result to (2.9) yields the following wave equation for $\mathbf{E}_0$,

$$\nabla_F^2 \mathbf{E}_0 = \frac{n_0^2}{c^2} \frac{\partial^2 \mathbf{E}_0}{\partial T_F^2} \tag{2.10}$$

the general solution to which is a superposition of plane waves $\mathbf{e}_{\mathbf{k},\lambda,} e^{i(\mathbf{k}\cdot\mathbf{X_F}-\omega T_F)}$ where ω and $\mathbf{k}$ satisfy $||\mathbf{k}||^2 = \omega^2 \frac{n_0^2}{c^2}$. Here $\mathbf{e}_{\mathbf{k},\lambda}$ denotes one of two unit vectors labelled by λ that, together with $\mathbf{k}$ form an orthogonal triad. These two linearly independent normal mode polarizations are required to create a complete set of normal modes. So,

$$\mathbf{E}_0(\mathbf{X}_\beta, T_\beta) = \int d^3k \sum_{\lambda=1}^{2} A_{\mathbf{k}}(\mathbf{X}_S, T_S)\mathbf{e}_{\mathbf{k},\lambda} e^{i(\mathbf{k}\cdot\mathbf{X}_F-\omega(\mathbf{k})T_F)}$$

where the integral is over all $\mathbf{k}$ and the plane wave coefficients $A_{\mathbf{k}}(\mathbf{X}_S, T_S)$ are functions only of the slow variables $\mathbf{X}_S$ and T_S. These coefficients are in fact the mode envelopes that we seek to determine through the CME. The zeroth order approximation to the field thereby consists of normal modes of the unperturbed medium modulated by envelopes that, for $\xi = 0$, do not vary in space or time, as expected. What is intended by requiring that $\frac{\partial \mathbf{A}_{\mathbf{k}}}{\partial X_{S,i}} = 0$ is that rapid variations in the electric field be embodied in the plane waves $e^{i(\mathbf{k}\cdot\mathbf{X}_F-\omega(\mathbf{k})T_F)}$.

The integrand above can be written as $\mathbf{A}_{\mathbf{k}}(\mathbf{X}_S, T_S)\varphi_{\mathbf{k},\omega}$ where

$$\mathbf{A}_{\mathbf{k}}(\mathbf{X}_S, T_S) = \sum_{\lambda=1}^{2} A_{\mathbf{k},\lambda}(\mathbf{X}_S, T_S)\mathbf{e}_{\mathbf{k},\lambda}$$

and $\varphi_{\mathbf{k},\omega} = e^{i(\mathbf{k}\cdot\mathbf{X}_F-\omega(\mathbf{k})T_F)}$. We only consider monochromatic fields, so that, if $\mathbf{E}_0$, is monochromatic at a frequency ω_0, then the integral above can be written as a surface integral over the sphere in k-space centered at the origin and having radius $\frac{\omega_0 n_0}{c}$. For notational convenience, this integral will be written as a sum, so that

$$\mathbf{E}_0(\mathbf{X}_\beta, T_\beta) = \sum_{\mathbf{k}}^{(S)} \mathbf{A}_{\mathbf{k}}(\mathbf{X}_S, T_S)\varphi_{\mathbf{k}} \tag{2.11}$$

In the case of monochromaticity, ω can be omitted from the normal mode notation because the modes of interest are $\varphi_{\mathbf{k},\omega_0}$ with only $\mathbf{k}$ varying. Equation (2.11) is the form of the leading order term that we will use throughout the remainder of the derivation.

Having obtained the leading order term, the next term, $\mathbf{E}_1(\mathbf{X}_\beta, T_\beta)$, can be determined by collecting in the wave equation (2.2) all terms proportional to ξ and requiring that the resulting coefficient of ξ vanish. In doing so, it is useful to see that the intensity of the electric field, using the asymptotic expansion of the electric field, is given by

$$\begin{aligned}
|\mathbf{E}(\mathbf{x},t)|^2 &= \sum_{m,m'=0}^{\infty} \mathbf{E}_m \overline{\mathbf{E}}_{m'} \xi^{m'+m} \\
&= |\mathbf{E}_0|^2 + \sum_{m=0,m'=1}^{\infty} \mathbf{E}_m \overline{\mathbf{E}}_{m'} \xi^{m'+m} + \sum_{m=1,m'=0}^{\infty} \mathbf{E}_m \overline{\mathbf{E}}_{m'} \xi^{m'+m}
\end{aligned}$$

It follows that in the wave equation (2.2), the only contribution to terms proportional to ξ from the nonlinear term $\xi\Delta^{nl}(\mathbf{X}_F)|\mathbf{E}|^2$ is $\xi\Delta^{nl}(\mathbf{X}_F)|\mathbf{E}_0|^2$.

So, collecting all of the terms in the wave equation that are proportional to ξ yields the following equation for $\mathbf{E}_1(\mathbf{X}_\beta, T_\beta)$ in terms of $\mathbf{E}_0(\mathbf{X}_\beta, T_\beta)$.

$$\begin{aligned}
& \sum_{m=1}^{3} \frac{\partial^2 E_{1,i}}{\partial X_{F,m}^2} - \sum_{m=1}^{3} \frac{\partial^2 E_{1,m}}{\partial X_{F,m}\partial X_{F,i}} - \frac{n_0^2}{c^2}\frac{\partial^2 E_{1,i}}{\partial T_F^2} \\
= & -2\sum_{m=1}^{3} \frac{\partial^2 E_{0,i}}{\partial X_{F,m}\partial X_{S,m}} + \sum_{m=1}^{3}\left(\frac{\partial^2 E_{0,m}}{\partial X_{F,m}\partial X_{S,i}} + \frac{\partial^2 E_{0,m}}{\partial X_{F,i}\partial X_{S,m}}\right) + \frac{2n_0^2}{c^2}\frac{\partial^2 E_{0,i}}{\partial T_F \partial T_S} + \\
& \frac{2n_0}{c^2}\Delta^l(\mathbf{X}_F)\frac{\partial^2 E_{0,i}}{\partial T_F^2} + \frac{2n_0}{c^2}\Delta^{nl}(\mathbf{X}_F)|\mathbf{E}_0|^2\frac{\partial^2 E_{0,i}}{\partial T_F^2}
\end{aligned} \tag{2.12}$$

Note that the left side of this equation is the i^{th} component of the vector expression

$\nabla_F^2 \mathbf{E}_m - \nabla_F\big(\nabla_F \cdot \mathbf{E}_m\big) - \frac{n_0^2}{c^2}\frac{\partial^2 \mathbf{E}_m}{\partial T_F^2}$ with $m = 1$. As is explained in section 2.4, this is the case for all m in the recurrence relations determining the $\mathbf{E}_m$.

Now, using the solution obtained for $\mathbf{E}_0$ and recalling that the mode envelopes $\mathbf{A}_\mathbf{k}(\mathbf{X}_S, T_S)$ are functions only of the slow variables, the equation above becomes

$$
\begin{aligned}
&\Big(\nabla_F^2 \mathbf{E}_1 - \nabla_F\big(\nabla_F \cdot \mathbf{E}_1\big) - \frac{n_0^2}{c^2}\frac{\partial^2 \mathbf{E}_1}{\partial T_F^2}\Big)_i \\
= \; & \sum_{\mathbf{k}}^{(S)} \Big[\sum_{m=1}^{3} -2ik_m \frac{\partial A_{\mathbf{k},i}}{\partial X_{S,m}} \varphi_\mathbf{k} + \sum_{m=1}^{3} \Big(ik_m \frac{\partial A_{\mathbf{k},m}}{\partial X_{S,i}} + ik_i \frac{\partial A_{\mathbf{k},m}}{\partial X_{S,m}} \Big) \varphi_\mathbf{k} - i\frac{2n_0^2\omega_0}{c^2}\frac{\partial A_{\mathbf{k},i}}{\partial T_S}\varphi_\mathbf{k} - \\
& \frac{2n_0\omega_0^2}{c^2}\Delta^l(\mathbf{X}_F) A_{\mathbf{k},i}\varphi_\mathbf{k} - \frac{2n_0\omega_0^2}{c^2}\Delta^{nl}(\mathbf{X}_F)|\mathbf{E}_0|^2 A_{\mathbf{k},i}\varphi_\mathbf{k}\Big] \qquad (2.13)
\end{aligned}
$$

We see that in (2.13), an expression for the intensity of the leading order term is required in terms of the mode envelopes. The intensity of the leading order term is given by

$$
\begin{aligned}
|\mathbf{E}_0|^2 &= \mathbf{E}_0 \overline{\mathbf{E}}_0 \\
&= \sum_{\mathbf{k}', \mathbf{k}''}^{(S)} \mathbf{A}_{\mathbf{k}'} \cdot \overline{\mathbf{A}_{\mathbf{k}''}} e^{i(\mathbf{k}'-\mathbf{k}'')\cdot \mathbf{X}_F}
\end{aligned}
$$

since $\varphi_{\mathbf{k}'}\overline{\varphi}_{\mathbf{k}''} = e^{i(\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{X}_F}$. Using this expression for the leading order intensity as well as the Fourier series for $\Delta^l(\mathbf{X}_F)$ and $\Delta^{nl}(\mathbf{X}_F)$, (2.5), and noting that $e^{-i\mathbf{G}\cdot\mathbf{X}_F}\varphi_\mathbf{k} = \varphi_{\mathbf{k}-\mathbf{G}}$, (2.13) becomes

$$
\begin{aligned}
&\Big(\nabla_F^2\mathbf{E}_1 - \nabla_F\big(\nabla_F\cdot\mathbf{E}_1\big) - \frac{n_0^2}{c^2}\frac{\partial^2\mathbf{E}_1}{\partial T_F^2}\Big)_i \\
= &\sum_{\mathbf{k}}^{(S)}\Big[-i\Big(\mathbf{k}\cdot\nabla_S A_{\mathbf{k},i} - \frac{1}{2}\sum_{m=1}^{3}\Big(k_m\frac{\partial A_{\mathbf{k},m}}{\partial X_{S,i}} + k_i\frac{\partial A_{\mathbf{k},m}}{\partial X_{S,m}}\Big) + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_{\mathbf{k},i}}{\partial T_S}\Big)\varphi_{\mathbf{k}} - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{G}}\Delta^l_{\mathbf{G}}A_{\mathbf{k},i}\varphi_{\mathbf{k}-\mathbf{G}} - \frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{k}',\mathbf{k}''}^{(S)}\sum_{\mathbf{G}}\Delta^{nl}_{\mathbf{G}}\big(\mathbf{A}_{\mathbf{k}'}\cdot\overline{\mathbf{A}}_{\mathbf{k}''}\big)A_{\mathbf{k},i}\varphi_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''-\mathbf{G}}\Big] \\
= &\sum_{\mathbf{k}}^{(S)}\Big[-i\Big(\mathbf{k}\cdot\nabla_S A_{\mathbf{k},i} - \frac{1}{2}\mathbf{k}\cdot\frac{\partial\mathbf{A}_{\mathbf{k}}}{\partial X_{S,i}} - \frac{1}{2}k_i\nabla_S\cdot\mathbf{A}_{\mathbf{k}} + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_{\mathbf{k},i}}{\partial T_S}\Big)\varphi_{\mathbf{k}} - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{G}}\Delta^l_{\mathbf{G}}A_{\mathbf{k},i}\varphi_{\mathbf{k}-\mathbf{G}} - \frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{k}',\mathbf{k}''}^{(S)}\sum_{\mathbf{G}}\Delta^{nl}_{\mathbf{G}}\big(\mathbf{A}_{\mathbf{k}'}\cdot\overline{\mathbf{A}}_{\mathbf{k}''}\big)A_{\mathbf{k},i}\varphi_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''-\mathbf{G}}\Big] \quad (2.14)
\end{aligned}
$$

This vector equation is the equation from which the coupled mode equations for any lattice geometry are obtained. In particular, coupled mode equations are extracted from (2.14) by identifying on its right side terms that are solutions to the homogeneous problem corresponding to its left side, so called secular terms [34].

That the expansion (2.7) is asymptotic requires that the secular terms on the right side of (2.14) collectively vanish. Were these terms to remain, there would exist terms in the solution of $\mathbf{E}_1$ that would grow linearly with time, the result of which would be that the term $\mathbf{E}_1\xi$ would be $O(1)$ at times on the order of ξ^{-1}, destroying the asymptoticity of the expansion (2.14).

Alternatively, this treatment of secular terms can be shown to be a consequence of the existence of the first order solution $\mathbf{E}_1$. This can be seen by casting (2.14) into the form $L\mathbf{E}_1 = \mathbf{f}$, where $L = \nabla_F^2 - \nabla_F\big(\nabla_F\cdot\big) - \frac{n_0^2}{c^2}\frac{\partial^2}{\partial T_F^2}$ is a self-adjoint linear operator, and $\mathbf{f}$ a vector whose i^{th} component is given by the right side of (2.14). Now, for a solution $\mathbf{E}_1$ to exist, $\mathbf{f}$ must be in the range of L which is denoted $R(L)$. This in turn requires that $\mathbf{f}$ be orthogonal to the null space of the adjoint of L, which is simply the null space of L, $N(L)$. $N(L)$ contains the solutions to the homogeneous problem $L\mathbf{E} = \mathbf{0}$. That is, $N(L) \supset \{\mathbf{e_k}\varphi_{\mathbf{k},\omega} : ||\mathbf{k}|| = \frac{n_0\omega}{c}, \mathbf{e_k}\cdot\mathbf{k} = 0\}$. Thus, $\langle \mathbf{e_k}\varphi_\mathbf{k}, \mathbf{f}\rangle = \sum_{i=1}^3 e_i \langle \varphi_\mathbf{k}, f_i\rangle = 0$ for all $\varphi_\mathbf{k}$ such that $||\mathbf{k}|| = \frac{n_0\omega_0}{c}$, where the inner product $\langle \varphi_\mathbf{k}, f_i\rangle$ is just $\int d^3x\overline{\varphi_\mathbf{k}(\mathbf{x})}f_i(\mathbf{x})$.

In calculating the integral $\langle \varphi_\mathbf{k}, f_i\rangle$ explicitly, we treat the multiple scales as independent variables so that, for example,

$$
\begin{aligned}
&\langle \varphi_\mathbf{k}|\mathbf{k}\cdot\frac{\partial \mathbf{A_k}}{\partial X_{S,i}}|\varphi_{\mathbf{k}'}\rangle \\
=\ & \int d^3X_F\mathbf{k}\cdot\frac{\partial \mathbf{A_k}}{\partial X_{S,i}}\overline{\varphi}_\mathbf{k}(\mathbf{X_F})\varphi_{\mathbf{k}'} \\
=\ & \mathbf{k}\cdot\frac{\partial \mathbf{A_k}}{\partial X_{S,i}}\int d^3X_F\overline{\varphi}_\mathbf{k}(\mathbf{X_F})\varphi_{\mathbf{k}'} \\
=\ & \mathbf{k}\cdot\frac{\partial \mathbf{A_k}}{\partial X_{S,i}}\langle \varphi_\mathbf{k}|\varphi_{\mathbf{k}'}\rangle
\end{aligned}
$$

Explicitly writing out the inner product of $\mathbf{e_k}\varphi_\mathbf{k}$ with (2.14) yields the following equation,

$$
\begin{aligned}
&\sum_{i=1}^{3} e_{\mathbf{k},i}\Big[\Big(\mathbf{k}\cdot\nabla_S A_{\mathbf{k},i} - \frac{1}{2}\mathbf{k}\cdot\frac{\partial \mathbf{A_k}}{\partial X_{S,i}} - \frac{1}{2}k_i\nabla_S\cdot\mathbf{A_k} + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_{\mathbf{k},i}}{\partial T_S}\Big)\langle\varphi_\mathbf{k},\varphi_\mathbf{k}\rangle - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{k}'}^{(S)}\sum_{all\mathbf{G}}\Delta_\mathbf{G}^l A_{\mathbf{k},i}\langle\varphi_\mathbf{k},\varphi_{\mathbf{k}'-\mathbf{G}}\rangle - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{k}',\mathbf{k}'',\mathbf{k}'''}^{(S)}\sum_{all\mathbf{G}}\Delta_\mathbf{G}^{nl}\big(\mathbf{A}_{\mathbf{k}''}\cdot\overline{\mathbf{A}}_{\mathbf{k}'''}\big)A_{\mathbf{k}',i}\langle\varphi_\mathbf{k},\varphi_{\mathbf{k}'+\mathbf{k}''-\mathbf{k}'''-\mathbf{G}}\rangle\Big] = 0 \qquad (2.15)
\end{aligned}
$$

which, carrying out the sum over i, becomes

$$
\begin{aligned}
&i\Big(\mathbf{k}\cdot\nabla_S(\mathbf{e_k}\cdot A_\mathbf{k}) - \frac{1}{2}\mathbf{k}\cdot(\mathbf{e_k}\cdot\nabla_S)\mathbf{A_k} + \frac{n_0^2\omega_0}{c^2}\frac{\partial(\mathbf{e_k}\cdot A_\mathbf{k})}{\partial T_S}\Big)\langle\varphi_\mathbf{k},\varphi_\mathbf{k}\rangle - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{k}'}^{(S)}\sum_{all\mathbf{G}}\Delta_\mathbf{G}^l(\mathbf{e_k}\cdot\mathbf{A}_{\mathbf{k}'})\langle\varphi_\mathbf{k},\varphi_{\mathbf{k}'-\mathbf{G}}\rangle - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{k}',\mathbf{k}'',\mathbf{k}'''}^{(S)}\sum_{all\mathbf{G}}\Delta_\mathbf{G}^{nl}\big(\mathbf{A}_{\mathbf{k}''}\cdot\overline{\mathbf{A}}_{\mathbf{k}'''}\big)(\mathbf{e_k}\cdot\mathbf{A}_{\mathbf{k}'})\langle\varphi_\mathbf{k},\varphi_{\mathbf{k}'+\mathbf{k}''-\mathbf{k}'''-\mathbf{G}}\rangle = 0 \qquad (2.16)
\end{aligned}
$$

where the notation $(\mathbf{e_k}\cdot\nabla_S)\mathbf{A_k}$ refers to a vector whose j^{th} component is $\sum_{i=1}^3 e_{\mathbf{k},i}\frac{\partial}{\partial X_{S,i}}A_{\mathbf{k},j}$.

Equation (2.16) is the generalized coupled mode equation mentioned at the beginning of the chapter.

Equation (2.16) was obtained in part using the orthogonality of the normal modes $\varphi_{\mathbf{k}}$, $\langle \varphi_{\mathbf{k}}, \varphi_{\mathbf{k}'} \rangle = \delta^{(3)}(\mathbf{k} - \mathbf{k}')$. The left side of equation (2.16) is the component of $\mathbf{f}$ that lies on the homogeneous solution $\mathbf{e_k}\varphi_{\mathbf{k}}$. It is by requiring that the projection of $\mathbf{f}$ on each homogeneous solution vanish that the coupled mode equations are obtained.

Obtaining the coupled mode equations in this way may seem to present some difficulty because, in the first instance, of the existence of an infinite number of solutions $\mathbf{e_k}\varphi_{\mathbf{k}}$ for a given $\mathbf{k}$. This suggests the existence of an infinite number of coupled mode equations, one for each projection $\langle \mathbf{e_k}\varphi_{\mathbf{k}}, \mathbf{f} \rangle$. In fact however, all of these equations can be expressed as a linear combination of two linearly independent equations obtained by projecting $\mathbf{f}$ onto any two vectors $\mathbf{e_k}$ that span the plane defined by $\mathbf{k}$. This follows from the linearity of (2.16) with respect to $\mathbf{e_k}$ and guarantees that of all the projections $\langle \mathbf{e_k}\varphi_{\mathbf{k}}, \mathbf{f} \rangle$ vanish. Thus, given the condition $\mathbf{k} \cdot \mathbf{A_k} = 0$ as well as these two linearly independent differential equations, we find that there exist three independent equations for each mode envelope $\mathbf{A_k}$, the components of which can thereby be determined.

Equation (2.16) is in fact a coupled mode equation for the mode $\varphi_{\mathbf{k}}$, although not in its final form. To reduce the equation further requires finding all of the $\langle \varphi_{\mathbf{k}}, \varphi_{\mathbf{k}-\mathbf{G}} \rangle$ that are nonzero. From the orthogonality of the normal modes, this is equivalent to finding all wavevectors $\mathbf{k}'$ and all RLVs $\mathbf{G}$ such that $\mathbf{k}' - \mathbf{G} = \mathbf{k}$ as well as all $\mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ and $\mathbf{G}$ such that $\mathbf{k}' + \mathbf{k}'' - \mathbf{k}''' - \mathbf{G} = \mathbf{k}$. The modes corresponding to those wavevectors are then said to be coupled because, as will be shown shortly, the envelopes of these modes influence the evolution of one another.

Turning now to the problem of finding the nonzero terms $\langle \varphi_{\mathbf{k}}, \varphi_{\mathbf{k}-\mathbf{G}} \rangle$, we first consider the mode coupling caused by the linear index perturbation. As is evident in (2.14), it is the linear component of the index perturbation that couples modes $\varphi_{\mathbf{k}}, \varphi_{\mathbf{k}'}$ corresponding to wavevectors that satisfy $\mathbf{k} = \mathbf{k}' - \mathbf{G}$ and $||\mathbf{k}|| = ||\mathbf{k}'|| = \frac{n_0\omega_0}{c}$. Now, the condition $||\mathbf{k}|| = ||\mathbf{k}'||$ is equivalent to $||\mathbf{k}||^2 = ||\mathbf{k} + \mathbf{G}||^2$, which requires that $\mathbf{k} \cdot \mathbf{G} = -\frac{||\mathbf{G}||}{2}$. By the construction of Brillouin zones, the wavevector $\mathbf{k}$ must then necessarily lie on the face of a particular Brillouin zone and, in particular, the face corresponding to the RLV $\mathbf{G}$. Conversely, for $\mathbf{k}$ to lie on the face of a Brillouin zone, it is sufficient for $\mathbf{k} \cdot \mathbf{G} = -\frac{||\mathbf{G}||}{2}$ for some RLV $\mathbf{G}$, in which case $||\mathbf{k} + \mathbf{G}|| = ||\mathbf{k}||$ and $\mathbf{k}$ is then coupled to $\mathbf{k} + \mathbf{G}$. So, for the wavevector $\mathbf{k}$ to be coupled to another wavevector lying on the sphere S, it is necessary and sufficient for $\mathbf{k}$ to lie on the face of a Brillouin zone of the index lattice. A wavevector is coupled to more than one other wavevector when it lies on more than one face of a Brillouin zone. This occurs when the wavevector lies at the intersection of some number of faces of a Brillouin zone. This number is then the number of directions to which the given wavevector is coupled. It is coupled by the RLVs corresponding to the faces on which it lies.

This process of identifying the modes coupled by the linear perturbation to a given mode $\varphi_{\mathbf{k}}$ is exhaustive in that all of the wavevectors $\mathbf{k}'$ such that $\mathbf{k} = \mathbf{k}' - \mathbf{G}$ with $||\mathbf{k}'|| = ||\mathbf{k}||$ can be determined from the position of $\mathbf{k}$ with respect to the Brillouin zones of the index lattice. Moreover, the set of modes that a given mode is coupled to are themselves coupled only to modes in that set. To see this, suppose that a wavevector $\mathbf{k}_1$ was coupled to another, $\mathbf{k}_2$, and that $\mathbf{k}_2$ was itself coupled to $\mathbf{k}_3$. This requires that $\mathbf{k}_1 = \mathbf{k}_2$ and $||\mathbf{k}_2|| = ||\mathbf{k}_3||$ as well as the existence of RLVs $\mathbf{G}_{12}$ and $\mathbf{G}_{23}$ such that $\mathbf{k}_2 = \mathbf{k}_1 - \mathbf{G}_{12}$ and $\mathbf{k}_3 = \mathbf{k}_2 - \mathbf{G}_{23}$. This in turn implies that $||\mathbf{k}_1|| = ||\mathbf{k}_3||$ and that $\mathbf{k}_3 = \mathbf{k}_1 - (\mathbf{G}_{12} + \mathbf{G}_{23})$ so that $\mathbf{k}_1$ and $\mathbf{k}_3$ are indeed coupled inasmuch as $\mathbf{G}_{12} - \mathbf{G}_{23}$ is itself an RLV (lattices are closed under vector addition).

Having characterized the modes coupled by the linear index perturbation, we now turn to the nonlinear perturbation, Δ^{nl}. The nonlinear component of the index perturbation couples modes whose wavevectors lie on the sphere S and that satisfy $\mathbf{k}' + \mathbf{k}'' - \mathbf{k}'''$

$-\mathbf{G} = \mathbf{k}$ for some RLV $\mathbf{G}$. Now, all of the wavevectors that are coupled by the linear perturbation and that are characterized in the preceding paragraph necessarily differ by an RLV. The condition for coupling by the nonlinear perturbation can be rewritten as $(\mathbf{k} - \mathbf{k}''') + (\mathbf{k}' - \mathbf{k}'') = \mathbf{G}$. So, if those wavevectors are ones coupled by the linear perturbation, there always exists an RLV $\mathbf{G}$ satisfying that condition, insofar as $\mathbf{k} - \mathbf{k}'''$ and $\mathbf{k}' - \mathbf{k}''$ are both RLVs in this case. It is sufficient then, for modes to be coupled by the linear perturbation for them to be coupled by the nonlinear perturbation. The converse however, is not true. It is not necessary for modes to be coupled by the linear perturbation to be coupled by the nonlinear perturbation. This can be seen from the condition $(\mathbf{k} - \mathbf{k}''') + (\mathbf{k}' - \mathbf{k}'') = \mathbf{G}$ when $\mathbf{k}$ and $\mathbf{k}'''$ are linearly coupled and when $\mathbf{k}' = \mathbf{k}''$. In this case, $\mathbf{k}'$ and so $\mathbf{k}''$ can be any wavevectors on the sphere S and still satisfy the nonlinear coupling condition. In particular, the nonlinear perturbation couples together all of the modes in the expansion (2.7), a difficultly that we will address shortly.

Having determined how to identify the nonzero projections $\langle \varphi_{\mathbf{k}}, \varphi_{\mathbf{k}'} \rangle$, the coupled mode equation (2.16) can be put into its final form by replacing each of the double sums therein by the terms corresponding to those nonzero projections. The resulting equation is also referred to as an asymptotic solvability condition.

We now see that the effect of the nonlinearity coupling all modes together is to introduce an infinite number of terms into the expressions $f_{\mathbf{k}}$ via the rightmost double sum in (2.7). Also, because there exists a solvability condition for each wavevector lying on the sphere S, there exists an infinite number of coupled mode equations. However, only a finite number of equations, those corresponding to the modes that are coupled by the linear perturbation, contain terms from both the linear and nonlinear perturbations. If the leading order solution $\mathbf{E}_0$ were to be a sum only of those modes which are coupled by the linear perturbation, the number of coupled mode equations would become finite and the equations themselves would contain only a finite number of terms.

Thus, the approach we adopt in finding the CME is, given a wavevector $\mathbf{k_i}$ that is known to be present in the medium, to expand the leading order solution $\mathbf{E}_0$ only in terms of $\varphi_{\mathbf{k_i}}$ and the modes to which it is coupled. To solve for the mode envelopes $\mathbf{A_k}$, we use the projection (2.16) with two linearly independent solutions for each wavevector, $\mathbf{e}_{\mathbf{k},m}\varphi_{\mathbf{k}}$, $m = 1, 2$. This, in conjunction with the condition $\mathbf{A_k} \cdot \mathbf{k} = 0$ provides three independent equations for each mode envelope. We thus have a number of independent equations that is equal to the number of unknowns in the problem, the number of mode envelope components.

Given this approach, (2.16) can be recast into a slightly more compact form. If the set of wavevectors to which a given wavevector $\mathbf{k}$ is linearly coupled is $\Omega = \{\mathbf{k}\} \cup \{\mathbf{k}_i : i = 1, 2, \ldots N-1\}$ so that there are N coupled modes, the CME for $\mathbf{A_k}$ is,

$$\begin{aligned}
&i\Big(\widehat{\mathbf{k}} \cdot \nabla_S(\mathbf{e_k} \cdot A_\mathbf{k}) - \frac{1}{2}\widehat{\mathbf{k}} \cdot (\mathbf{e_k} \cdot \nabla_S)\mathbf{A_k} + \frac{n_0}{c}\frac{\partial(\mathbf{e_k} \cdot A_\mathbf{k})}{\partial T_S}\Big) - \\
&\frac{1}{n_0||\mathbf{k}||}\sum_{\mathbf{k}' \in \Omega} \Delta^l_{\mathbf{k}'-\mathbf{k}}(\mathbf{e_k} \cdot \mathbf{A}_{\mathbf{k}'}) - \\
&\frac{1}{n_0||\mathbf{k}||}\sum_{\mathbf{k}',\mathbf{k}'',\mathbf{k}''' \in \Omega} \Delta^{nl}_{\mathbf{k}'+\mathbf{k}''-\mathbf{k}'''-\mathbf{k}}\big(\mathbf{A}_{\mathbf{k}''} \cdot \overline{\mathbf{A}}_{\mathbf{k}'''}\big)(\mathbf{e_k} \cdot \mathbf{A}_{\mathbf{k}'}) = 0
\end{aligned} \tag{2.17}$$

where $\widehat{\mathbf{k}} = \frac{\mathbf{k}}{||\mathbf{k}||}$. Here we have used the condition $||\mathbf{k}|| = \frac{n_0\omega_0}{c}$. Equation (2.17) follow from (2.16) by recognizing that the terms in the sums over the reciprocal lattice vectors in (2.16) are nonzero only when, in the linear perturbation terms, $\mathbf{G} = \mathbf{k} - \mathbf{k}'$ and in the nonlinear terms when $\mathbf{G} = \mathbf{k}' + \mathbf{k}'' - \mathbf{k}''' - \mathbf{k}$. The sums over all reciprocal lattice vectors reduce in this way to sums only over those generated by the wavevectors in Ω. This is elucidated in section 2.2.

Given that the field is expanded using a finite number of modes, the number of terms in the CMEs originating from the nonlinear perturbation will in turn be finite. In particular, the CMEs for an expansion involving N modes will contain N^3 nonlinear terms. This is confirmed, and the nonlinear terms enumerated in section 2.2.

Before proceeding to obtain sets of CME using the method outlined in the previous section, we briefly discuss some physical aspects of (2.16), the most prominent of which are the effects of the Kerr nonlinearity. These effects are embodied in the terms $\Delta^{nl}_{\mathbf{G}}\big(\mathbf{A}_{\mathbf{k}'}\cdot\overline{\mathbf{A}}_{\mathbf{k}''}\big)A_{\mathbf{k},i}$, which correspond to cross-phase modulation for $\mathbf{k}'\neq\mathbf{k}''$ and to self-phase modulation otherwise. The strength of these terms as well as that of the linear coupling is proportional to the fourier components $\Delta^{nl}_{\mathbf{G}}$ and $\Delta^{l}_{\mathbf{G}}$. Indeed, it is the terms such as $\Delta^{nl}_{\mathbf{G}}cos(\mathbf{G}\cdot\mathbf{r})$ constituting the index perturbations that are exclusively responsible for the coupling between modes. The coupling corresponding to a given coefficient $\Delta^{l}_{\mathbf{G}}$ or $\Delta^{nl}_{\mathbf{G}}$ would continue to occur even if it were the only nonzero Fourier coefficient in the perturbations.

The linear perturbation is exactly analogous to the elastic scattering of x-rays by semiconductor crystals. The conditions for coherent scattering in that case are the same as those found above, $||\mathbf{k}|| = ||\mathbf{k}'||$ and $\mathbf{k}'-\mathbf{k}=\mathbf{G}$ for wavevectors $\mathbf{k}$ and $\mathbf{k}'$ and RLV $\mathbf{G}$. These are known as the Bragg, or Laue conditions for strong scattering. These conditions ensure that plane waves with wavevectors $\mathbf{k}$ and $\mathbf{k}'$ scattered by the family of lattice planes defined by $\mathbf{G}$ are in phase with one another and thereby interfere constructively [1].

2.2 Nonlinear Terms

In discussing the nonlinear coupling mechanism in section 2.1, we stated the CMEs describing N linearly coupled modes should contain N^3 nonlinear terms of the form $\Delta^{nl}_{\mathbf{G}}\mathbf{A}_{\mathbf{k}_j}\cdot\overline{\mathbf{A}}_{\mathbf{k}_j}(\mathbf{e}_{\mathbf{k}_n}\cdot\mathbf{A}_{\mathbf{k}_i})$. This section addresses that expectation and enumerates the nonlinear terms found in all CMEs.

If the set of wavevectors of the coupled modes is $\Omega=\{\mathbf{k}_i : i\in S=\{1,2,\ldots N\}\}$, then the CME for the mode $\mathbf{k}_n\in\Omega$, will contain a nonlinear term of the form $\Delta^{nl}_{\mathbf{G}}\mathbf{A}_{\mathbf{k}_j}\cdot$

$\overline{\mathbf{A}}_{\mathbf{k}_j}(\mathbf{e}_{\mathbf{k}_n} \cdot \mathbf{A}_{\mathbf{k}_i})$ if and only if $\mathbf{k}_n = \mathbf{k}_i + \mathbf{k}_j - \mathbf{k}_k - \mathbf{G}$. This follows from the inner product $\left\langle \varphi_{\mathbf{k}_n}, \varphi_{\mathbf{k}_i+\mathbf{k}_j-\mathbf{k}_k-\mathbf{G}} \right\rangle = \delta^{(3)}(\mathbf{k}_n - \mathbf{k}_i - \mathbf{k}_j + \mathbf{k}_k + \mathbf{G})$ that is formed in finding the CME for $\mathbf{A}_{\mathbf{k}_n}$.

The condition $\mathbf{k}_n = \mathbf{k}_i + \mathbf{k}_j - \mathbf{k}_k - \mathbf{G}$ can be recast as $\mathbf{G}_{in} + \mathbf{G}_{jk} = \mathbf{G}$ where $\mathbf{G}_{in} = \mathbf{k}_i - \mathbf{k}_n$ and $\mathbf{G}_{jk} = \mathbf{k}_j - \mathbf{k}_k$. $\mathbf{G}_{in}$ and $\mathbf{G}_{jk}$ are necessarily RLVs inasmuch as all pairwise differences between elements of Ω are RLVs.

It is clear from this alternative nonlinear coupling condition that for each triplet (i, j, k) with $i, j, k \in S = \{1, 2 \ldots N\}$, there exists an RLV $\mathbf{G}$ such that $\varphi_{\mathbf{k}_i+\mathbf{k}_j-\mathbf{k}_k-\mathbf{G}} = \varphi_{\mathbf{k}_n}$ or $\left\langle \varphi_{\mathbf{k}_n} = \varphi_{\mathbf{k}_i+\mathbf{k}_j-\mathbf{k}_k-\mathbf{G}} \right\rangle \neq 0$. Then, because the CME for $\mathbf{A}_{\mathbf{k}}$ contains a nonlinear term for each nonzero inner product $\left\langle \varphi_{\mathbf{k}_n} = \varphi_{\mathbf{k}_i+\mathbf{k}_j-\mathbf{k}_k-\mathbf{G}} \right\rangle$ and because there are N^3 triplets (i, j, k) where $i, j, k \in S$, there will be N^3 nonlinear terms in the CME for each mode in Ω.

The task of enumerating the nonlinear terms in the CME for $\mathbf{A}_{\mathbf{k}_n}$ now consists of finding all of the RLVs $\mathbf{G} = \mathbf{G}_{in} + \mathbf{G}_{jk}$ generated as i, j and k vary over S. It is worth mentioning again that the nonlinear term corresponding to a particular triplet (i, j, k) is $\Delta^{nl}_{\mathbf{G}} \mathbf{A}_{\mathbf{k}_j} \cdot \overline{\mathbf{A}}_{\mathbf{k}_j}(\mathbf{e}_{\mathbf{k}_n} \cdot \mathbf{A}_{\mathbf{k}_i})$.

To obtain an exhaustive organization of the nonlinear coupling RLVs, we partition the set $S^3 = \{(i, j, k) : i, j, k \in S\}$. That is, the set is decomposed into disjoint subsets each characterized by some relationship between n, which is fixed for each CME, and i, j, k. This is turn provides a means or organizing the RLVs $\mathbf{G} = \mathbf{G}_{in} + \mathbf{G}_{jk}$ into groups within which the triplets labelling the RLVs share some common form.

The disjoint sets constituting S^3 are,

$$S^3_{1,n} = \{(i,j,k) : (i = n, k = j) \text{ or } (i = k, j = n, k \neq n)\}$$
$$S^3_{2,n} = \{(i,j,k) : (k = j, i \neq n) \text{ or } (i = k \neq j, j \neq n)\}$$
$$S^3_{3,n} = \{(i,j,k) : i = j = n, k \neq n\}$$
$$S^3_{4,n} = \{(i,j,k) : (i \neq k, i \neq n, k \neq n, j = n) \text{ or } (i = n, j \neq k, j \neq n, k \neq n)\}$$
$$S^3_{5,n} = \{(i,j,k) : i = j \neq n, k = n\}$$
$$S^3_{6,n} = \{(i,j,k) : i \neq n, i \neq k, j \neq k, j \neq n\}$$

Note that in the definitions of each subset, only n is fixed while i, j and k can be any element of S. The subsets are labelled by n because their contents depend on n, which labels the mode whose CME is being studied.

The equalities and inequalities defining each subset of S^3 are in place to ensure that they are disjoint. Thus, to show that the union of these subsets does indeed constitute a partition of S^3, it is sufficient to show that the sum of their cardinalities is $|S^3| = N^3$. The cardinalities of the subsets, together with brief explanations of how they are found, are listed below.

$\mathbf{|S^3_{1,n}| = N + N - 1 = 2N - 1}$

In the case $i = n, k = j$, j and so k can be any element of S, while i is fixed at n, giving N triplets. In the other case defining the subset S^3_1, $i = k, j = n$ and $i \neq k$, giving $N - 1$ triplets with one for each choice of $k \in S - \{n\}$, since $k \neq n$.

$\mathbf{|S^3_{2,n}| = N(N - 1) + (N - 1)(N - 1) = (2N - 1)(N - 1)}$

The relations $k = j, i \neq n$ define $N(N - 1)$ triplets since there are $N - 1$ choices for $i(\neq n) \in S$ and N choices of $k(= j)$. In the case $i = k \neq j, j \neq n$, $i(= k)$ can be any element of $S - \{j\}$, yielding $N - 1$ choices, while j can also be any element of

$S - \{n\}$. This second condition thus defines $(N-1)(N-1)$ triplets (i, j, k).

$|\mathbf{S^3_{3,n}}| = \mathbf{N-1}$

The condition $i = j = n, k \neq n$ which defines S^3_3 defines $N-1$ triplets, one for each choice of $k \in S - \{n\}$.

$|\mathbf{S^3_{4,n}}| = \mathbf{(N-1)(N-2) + (N-1)(N-2) = 2(N-1)(N-2)}$

In the first case defining the triplets in S^3_4, $j = n, i \neq n, k \neq n, i \neq k$. Here, i can be any of the $N-1$ elements in $S - \{n\}$, leaving $N-2$ choices for k, namely any element but i or n. This defines $(N-1)(N-2)$ triplets. The same is true of the other defining condition for S^3_4, $i = n, j \neq n, k \neq n, j \neq k$.

$|\mathbf{S^3_{5,n}}| = \mathbf{N-1}$

S^3_5 is defined by $i = j \neq n, k = n$ so that there is a triplet in S^3_5 for each element of $S - \{n\}$ (for each choice of $i (= j \neq n)$.

$|\mathbf{S^3_{6,n}}| = \mathbf{(N-1)(N-1)(N-2)}$

The conditions $i \neq n, k \neq i, j \neq k, j \neq n$ allow i and k to be any elements of $S - \{n\}$ and $S - \{i\}$ respectively while j can be any element of $S - \{k, n\}$. These relations define $(N-1)(N-1)(N-2)$.

The sum of the cardinalities found above is indeed N^3 and the disjoint subsets $S^3_i, i = 1, 2, \ldots 6$ thereby constitute a partition of S^3.

The categorization of triplets in S^3 induces a grouping of the RLVs that they label, namely $\mathbf{G}_{in} + \mathbf{G}_{jk}$. To each of the subsets described above will correspond a group of RLVs all having some common form. This form will in turn be determined by the relations defining the triplets in each subset. The number of RLVs in a given group will equal the cardinality of the corresponding subset of S^3.

The following rules, which follow from the definition $\mathbf{G}_{ij} = \mathbf{k}_i - \mathbf{k}_j$, will be useful in determining the RLVs corresponding to each of the subsets S_i^3.

$$\begin{aligned}
&\mathbf{G}_{ii} = \mathbf{0}\\
&\mathbf{G}_{ij} = -\mathbf{G}_{ji}\\
&\mathbf{G}_{ji} + \mathbf{G}_{ik} = \mathbf{G}_{jk}\\
&\mathbf{G}_{ij} + \mathbf{G}_{kl} = \mathbf{G}_{il} + \mathbf{G}_{kj}
\end{aligned}$$

Using these rules, we find that the following forms of $\mathbf{G}_{in} + \mathbf{G}_{jk}$ are associated with each subset.

$(i,j,k) \in S_{1,n}^3 \Rightarrow \mathbf{G}_{in} + \mathbf{G}_{jk} = \mathbf{0}$

$$(i,j,k) \in S_{2,n}^3 \Rightarrow \mathbf{G}_{in} + \mathbf{G}_{jk} = \begin{cases} \mathbf{G}_{in} \quad (\neq \mathbf{0} \Leftarrow i \neq n) \\ or \\ \mathbf{G}_{jn} \quad (\neq \mathbf{0} \Leftarrow j \neq n) \end{cases}$$

$(i,j,k) \in S_{3,n}^3 \Rightarrow \mathbf{G}_{in} + \mathbf{G}_{jk} = \mathbf{G}_{nk} \;\; (\neq \mathbf{0} \Leftarrow k \neq n)$

$$(i,j,k) \in S_{4,n}^3 \Rightarrow \mathbf{G}_{in} + \mathbf{G}_{jk} = \begin{cases} \mathbf{G}_{ik} \quad (\neq \mathbf{0} \Leftarrow i \neq k) \\ or \\ \mathbf{G}_{jk} \quad (\neq \mathbf{0} \Leftarrow j \neq k) \end{cases}$$

$(i,j,k) \in S_{5,n}^3 \Rightarrow \mathbf{G}_{in} + \mathbf{G}_{jk} = \mathbf{2G}_{in} \;\; (\neq \mathbf{0} \Leftarrow i \neq n)$

For $(i,j,k) \in S_6^3$, the form of $\mathbf{G}_{in} + \mathbf{G}_{jk}$ cannot, a priori, be simplified.

It is important to note that although the subsets $S_i^3, i = 1,2,\ldots 6$ are disjoint, the corresponding groups of RLVs are not. As an example of this, consider the two-dimensional square lattice and a wavevector $\mathbf{k}$ lying on a corner of its first Brillouin zone. The wavevector $\mathbf{k}$ is linearly coupled to three others, $\mathbf{k}_i, i = 1,2,3$ each lying at a different corner of the Brillouin zone. In this example then, $\Omega = \{\mathbf{k}_i : i \in S =$

$\{1,2,3,4\}\}$ where, for convenience, the labels of the wavevectors proceed clockwise around the corners of the Brillouin zone. Under this labelling scheme, $\mathbf{G}_{21} = \mathbf{G}_{43}$. Thus, with $n = 1$, $\mathbf{G}_{in} + \mathbf{G}_{jk}|_{(i,j,k)=(2,3,3)} = \mathbf{G}_{in} + \mathbf{G}_{jk}|_{(i,j,k)=(1,4,3)}$ despite that $(2,3,3) \in S^3_{2,1}$ and $(1,4,3) \in S^3_{4,1}$, which are disjoint.

Before proceeding, it will be useful to demonstrate how the results developed in this section can be used to obtain the nonlinear terms in any CME. Beginning with the set of linearly coupled modes, $\Omega = \{\mathbf{k}_i : i \in S = \{1, 2, \dots N\}\}$, we will the study the CME for the envelope $\mathbf{A}_{\mathbf{k}}$. This CME will contain nonlinear terms proportional to $\Delta^{nl}_{\mathbf{0}}$. From the results of this section, there will be at least $|S^3_{1,1}| = 2N-1$ such terms, each of one of the forms $\Delta^{nl}_{\mathbf{0}}(\mathbf{A}_{\mathbf{k}_j} \cdot \overline{\mathbf{A}}_{\mathbf{k}_j})(\mathbf{e}_{\mathbf{k}_1} \cdot \mathbf{A}_{\mathbf{k}_1})$ with $j \in S$, or $\Delta^{nl}_{\mathbf{0}}(\mathbf{A}_{\mathbf{k}_1} \cdot \overline{\mathbf{A}}_{\mathbf{k}_i})(\mathbf{e}_{\mathbf{k}_1} \cdot \mathbf{A}_{\mathbf{k}_i})$ with $i \in S, i \neq 1$. As another example, consider the term proportional to $\Delta^{nl}_{\mathbf{2G}_{21}}$, corresponding to the triplet $(2,2,1)$. The CME for $\mathbf{A}_{\mathbf{k}_1}$ will contain the term $\Delta^{nl}_{\mathbf{2G}_{21}}(\mathbf{A}_{\mathbf{k}_2} \cdot \overline{\mathbf{A}}_{\mathbf{k}_1})(\mathbf{e}_{\mathbf{k}_1} \cdot \mathbf{A}_{\mathbf{k}_1})$. The remainder of the nonlinear terms can be found in a similar manner.

Finally, we note that the emphasis here has been preponderantly on the nonlinear terms because the linear terms in the CMEs are straightforward to determine. It is clear from the selection rule $\langle \varphi_{\mathbf{k}} | \varphi_{\mathbf{k}'-\mathbf{G}} \rangle$ multiplying $\Delta^{l}_{\mathbf{G}}(\mathbf{e}_{\mathbf{k}} \cdot \mathbf{A}_{\mathbf{k}'})$ that the linear terms in the CME for $\mathbf{A}_{\mathbf{k}_i}, \mathbf{k}_i \in \Omega$ will be all of the elements of $\{\Delta^{l}_{\mathbf{G}_{ij}}(\mathbf{e}_{\mathbf{k}_i} \cdot \mathbf{A}_{\mathbf{k}_j}) : j \in S = \{1, 2, \dots N\}\}$.

2.3 Symmetries

In the case that a resonant mode is incident on a dielectric lattice along one of its axes of rotational symmetry, it might be expected that the modes into which the incident one is diffracted be related by that symmetry. To confirm this intuition, we begin by comparing the solutions of the CMEs in two different lattices that are related to one another by an arbitrary rotation.

To make the investigation precise, consider one lattice described by the index perturbation $\Delta(\mathbf{r})$ and another described by $\Delta'(\mathbf{r})$. It will be assumed that Δ and Δ' are related by $\Delta(\mathbf{r}) = \Delta'(\mathbf{r}')$ where the co-ordinates $\mathbf{r}$ and $\mathbf{r}'$ are themselves related by $\mathbf{r}' = \mathbf{T}^{-1}\mathbf{r}$. Here, $\mathbf{T}$ is an orthogonal matrix, so that $\mathbf{T}^{-1} = \mathbf{T}^{\dagger}$. The co-ordinates $\mathbf{r}'$ are those of the point described by $\mathbf{r}$ but subsequent to the rotation characterized by the matrix $\mathbf{T}$. It follows that what is entailed by the relation $\Delta(\mathbf{r}) = \Delta'(\mathbf{r}')$ is that Δ' assumes the value $\Delta(\mathbf{r})$ at the point into which that with co-ordinates $\mathbf{r}$ is rotated by $\mathbf{T}$. Simply, Δ' is a rotated version of Δ.

Suppose now that a resonant mode $\varphi_{\mathbf{k}}$ is incident on the lattice $\Delta(\mathbf{r})$ and is linearly coupled to modes with wavevectors in the set $\Omega = \{\mathbf{k}_i : i \in S = \{1, 2, \ldots N-1\}\}$. The CME for the envelope $\mathbf{A}_{\mathbf{k}_i}$ will be, from (2.16),

$$i\Big(\mathbf{k}_i \cdot \nabla_S(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_{\mathbf{k}_i}) - \frac{1}{2}\mathbf{k}_i \cdot (\mathbf{e}_{\mathbf{k}_i,m} \cdot \nabla_S)\mathbf{A}_{\mathbf{k}_i} + \frac{n_0^2\omega_0}{c^2}\frac{\partial(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_{\mathbf{k}_i})}{\partial T_S}\Big) - \frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{p}\in\boldsymbol{\Omega}} \Delta^{l}_{\mathbf{p}-\mathbf{k}_i}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_{\mathbf{p}}) - \frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{q},\mathbf{r},\mathbf{s}\in\Omega} \Delta^{nl}_{\mathbf{q}+\mathbf{r}-\mathbf{s}-\mathbf{k}_i}\big(\mathbf{A}_{\mathbf{q}} \cdot \overline{\mathbf{A}}_{\mathbf{r}}\big)(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_{\mathbf{s}}) = 0 \tag{2.18}$$

Consider next a mode $\varphi_{\mathbf{k}'}$ incident on the lattice $\Delta'(\mathbf{x}) = \Delta(\mathbf{T}\mathbf{x})$ where $\mathbf{k}' = \mathbf{T}^{-1}\mathbf{k}$. The mode $\varphi_{\mathbf{k}'}$ is thus rotated with respect to $\varphi_{\mathbf{k}}$ in the same manner that $\Delta'(\mathbf{r})$ is with respect $\Delta(\mathbf{r})$. To understand how $\varphi_{\mathbf{k}'}$ is coupled, consider first that the lattice reciprocal to that formed by Δ' is simply the image of the reciprocal lattice of Δ under the rotation $\mathbf{T}$. That is, the reciprocal lattice of Δ' is $\{\mathbf{T}^{-1}\mathbf{G} : \mathbf{G} \in$ reciprocal lattice of $\Delta\}$. It then follows from the orthogonality of $\mathbf{T}$ that the set of wavevectors to which $\mathbf{k}' = \mathbf{T}^{-1}\mathbf{k}$ is coupled is $\Omega' = \{\mathbf{k}_i' = \mathbf{T}^{-1}\mathbf{k}_i : i \in S = \{1, 2, \ldots N-1\}\}$. Furthermore, the set of reciprocal lattice vectors which couple the wavevectors in Ω' is simply $\{\mathbf{T}^{-1}(\mathbf{k}_i - \mathbf{k}_j) : \mathbf{k}_i, \mathbf{k}_j \in \{\mathbf{k}\} \cup \Omega\}$

What has thus been established is that all of the coupled modes in the lattice Δ' are related to those in Δ by the rotation $\mathbf{T}$.

The CME for the envelope $\mathbf{A}^{'}_{\mathbf{k}^{'}_i}$ where $\mathbf{k}^{'}_i = \mathbf{T}^{-1}\mathbf{k}_i, \mathbf{k}_i \in \Omega$ is,

$$i\Big(\mathbf{k}^{'}_i \cdot \nabla_S(\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} \cdot \mathbf{A}^{'}_{\mathbf{k}^{'}_i}) - \frac{1}{2}\mathbf{k}^{'}_i \cdot (\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} \cdot \nabla_S)\mathbf{A}^{'}_{\mathbf{k}^{'}_i} + \frac{n_0^2\omega_0}{c^2}\frac{\partial(\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} \cdot \mathbf{A}^{'}_{\mathbf{k}^{'}_i})}{\partial T_S}\Big) -$$
$$\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{p}'\in\mathbf{\Omega}'} \Delta^{l'}_{\mathbf{p}'-\mathbf{k}^{'}_i}(\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} \cdot \mathbf{A}_{\mathbf{p}'}{}^{'}) - \frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{q}',\mathbf{r}',\mathbf{s}'\in\Omega'} \Delta^{nl'}_{\mathbf{q}'+\mathbf{r}'-\mathbf{s}'-\mathbf{k}^{'}_i}\big(\mathbf{A}_{\mathbf{q}'}{}^{'} \cdot \overline{\mathbf{A}}^{'}_{\mathbf{r}'}\big)(\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} \cdot \mathbf{A}_{\mathbf{s}}{}^{'}) = 0 \tag{2.19}$$

Here, $\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m}$ is one of two linearly independent vectors ($\mathbf{e}^{'}_{\mathbf{k}^{'}_i,1}$ and $\mathbf{e}^{'}_{\mathbf{k}^{'}_i,2}$), both orthogonal to $\mathbf{k}^{'}_i$. It is possible, though not necessary, to take $\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} = \mathbf{T}^{-1}\mathbf{e}_{\mathbf{k}_i,m}$ since $\big\langle\mathbf{T}^{-1}\mathbf{e}_{\mathbf{k}_i,m}|\mathbf{k}^{'}_i\big\rangle = \big\langle\mathbf{T}^{-1}\mathbf{e}_{\mathbf{k}_i,m}|\mathbf{T}^{-1}\mathbf{k}_i\big\rangle = \big\langle\mathbf{e}_{\mathbf{k}_i,m}|\mathbf{k}_i\big\rangle = 0$, using the orthogonality of $\mathbf{T}$.

We can now exploit the similarity between equations (2.18) and (2.19) to establish a relationship between $\mathbf{A}^{'}_{\mathbf{k}^{'}_i}$ and $\mathbf{A}_{\mathbf{k}_i}$.

Beginning with the coefficients $\Delta'^{nl}_{\mathbf{G}'}$,

$$\begin{aligned}
\Delta'^{nl}_{\mathbf{G}'} &= \Delta'^{nl}_{\mathbf{T^{-1}G}} = \frac{1}{V'_{pc}}\int d^3x\Delta'(\mathbf{x})e^{-i\mathbf{G}'\cdot\mathbf{x}} \\
&= \frac{1}{V'_{pc}}\int d^3x\Delta'(\mathbf{x})e^{-i\big\langle\mathbf{T^{-1}G}|\mathbf{x}\big\rangle} \\
&= \frac{1}{V'_{pc}}\int d^3x'\Delta'(\mathbf{T}^{-1}\mathbf{x})e^{-i\big\langle\mathbf{G}|\mathbf{Tx}\big\rangle} \\
&= \frac{1}{V_{pc}}\int d^3x'\Delta(\mathbf{x}')e^{-i\mathbf{G}\cdot\mathbf{x}'} = \Delta^{nl}_{\mathbf{G}}
\end{aligned}$$

The fourth line above follows from the orthogonality of $\mathbf{T}$ and the fifth line from the definition of Δ'. The sixth line acknowledges that $V_{pc} = V'_{pc}$ where V_{pc} and V'_{pc} are the volumes of the primitive cells associated respectively with Δ and Δ'. This must hold insofar as Δ and Δ' are related by an isometry, $\mathbf{T}$.

Thus, the coefficients $\Delta'^{nl}_{\mathbf{G}'}$ and $\Delta'^{l}_{\mathbf{G}'}$ in (2.19) can be replaced by their unprimed counterparts $\Delta^{nl}_{\mathbf{G}'}$ and $\Delta^{l}_{\mathbf{G}'}$, where $\mathbf{G} = \mathbf{TG}'$.

Turning now to the term $\mathbf{k}^{'}_i \cdot \nabla_S(\mathbf{e}^{'}_{\mathbf{k}^{'}_i,m} \cdot \mathbf{A}^{'}_{\mathbf{k}^{'}_i})$ in (2.19), we assume, without loss of generality, that the $\mathbf{e}'_{\mathbf{k}'_i,m}$ have been chosen as $\mathbf{e}'_{\mathbf{k}'_i,m} = \mathbf{T}^{-1}\mathbf{e}_{\mathbf{k}_i,m}$. Hence, $\langle \mathbf{e}'_{\mathbf{k}'_i,m}, \mathbf{A}'_{\mathbf{k}'_i} \rangle = \langle \mathbf{T}^{-1}\mathbf{e}_{\mathbf{k}_i,m}, \mathbf{A}'_{\mathbf{k}'_i} \rangle = \langle \mathbf{e}_{\mathbf{k}_i,m}, \mathbf{T}\mathbf{A}'_{\mathbf{k}'_i} \rangle$.

Secondly, we have that, for any differentiable f, $\nabla_S f(\mathbf{X}_S) = \nabla_S f(\mathbf{T}^{-1}(\mathbf{T}\mathbf{X}_S)) = \nabla_S f(\mathbf{T}^{-1}\mathbf{X}'_S) = \mathbf{T}^{-1}\nabla'_S f(\mathbf{T}^{-1}\mathbf{X}'_S)$ where $\nabla'_{S,i} = \frac{\partial}{\partial X'_{S,i}}$ and $\mathbf{X}'_S = \mathbf{T}\mathbf{X}_S$. This follows from a simple application of the chain rule, $\frac{\partial}{\partial X_{S,i}} = \sum_{i=1}^{3} \frac{\partial X'_{S,i}}{\partial X_{S,i}} \frac{\partial}{\partial X'_{S,i}}$. Thus, $\mathbf{k}'_i \cdot \nabla_S f(\mathbf{X}_S) = \langle \mathbf{k}'_i, \mathbf{T}^{-1}\nabla'_S f(\mathbf{T}^{-1}\mathbf{X}'_S) \rangle = \langle \mathbf{T}^{-1}\mathbf{k}_i, \mathbf{T}^{-1}\nabla'_S f(\mathbf{T}^{-1}\mathbf{X}'_S) \rangle = \mathbf{k}_i \cdot \nabla'_S f(\mathbf{T}^{-1}\mathbf{X}'_S)$, which follows from the orthogonality of $\mathbf{T}$.

Combining the observations made in the two preceding paragraphs, $\mathbf{k}'_i \cdot \nabla_S(\mathbf{e}'_{\mathbf{k}'_i,m} \cdot \mathbf{A}'_{\mathbf{k}'_i}(\mathbf{X}_S)) = \mathbf{k}_i \cdot \nabla'_S(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_S))$.

Lastly, we consider the term $\mathbf{k}'_i \cdot (\mathbf{e}'_{\mathbf{k}'_i} \cdot \nabla_S)\mathbf{A}'_{\mathbf{k}'_i}$. Note that $\langle \mathbf{e}'_{\mathbf{k}'_i}, \nabla_S \rangle = \langle \mathbf{e}_{\mathbf{k}_i}, \mathbf{T}^{-1}\nabla_S \rangle = \langle \mathbf{e}_{\mathbf{k}_i}, \nabla'_S \rangle$ and that $\langle \mathbf{k}'_i, \langle \mathbf{e}'_{\mathbf{k}'_i}, \nabla_S \rangle \mathbf{A}'_{\mathbf{k}'_i}(\mathbf{X}_S) \rangle = \langle \mathbf{k}_i, \mathbf{T}\langle \mathbf{e}'_{\mathbf{k}'_i}, \nabla_S \rangle \mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_S) \rangle = \langle \mathbf{k}_i, \langle \mathbf{e}'_{\mathbf{k}'_i}, \nabla_S \rangle \mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_S) \rangle$. Thus, $\mathbf{k}'_i \cdot (\mathbf{e}'_{\mathbf{k}'_i} \cdot \nabla_S)\mathbf{A}'_{\mathbf{k}'_i} = \mathbf{k}_i \cdot (\mathbf{e}_{\mathbf{k}_i} \cdot \nabla'_S)\mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_S)$.

The effect of all of the observations made above is to transform (2.19) into the following CME,

$$
\begin{aligned}
&i\Big(\mathbf{k}_i \cdot \nabla'_S(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_F)) - \frac{1}{2}\mathbf{k}_i \cdot (\mathbf{e}_{\mathbf{k}_i,m} \cdot \nabla'_S)\mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_F) + \\
&\frac{n_0^2\omega_0}{c^2}\frac{\partial(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}'_F))}{\partial T_S}\Big) - \frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{p}'\in\mathbf{\Omega}'} \Delta^l_{\mathbf{p}'-\mathbf{k}'_i}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{T}\mathbf{A}_{\mathbf{p}'}{}'(\mathbf{T}^{-1}\mathbf{X}'_F)) - \\
&\frac{n_0\omega_0^2}{c^2}\sum_{\mathbf{q}',\mathbf{r}',\mathbf{s}'\in\Omega'} \Delta^{nl}_{\mathbf{q}'+\mathbf{r}'-\mathbf{s}'-\mathbf{k}'_i}\big(\mathbf{A}_{\mathbf{q}'}{}' \cdot \overline{\mathbf{A}}_{\mathbf{r}'}{}'\big)(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{T}\mathbf{A}_{\mathbf{s}}{}'(\mathbf{T}^{-1}\mathbf{X}'_F)) = 0
\end{aligned}
\tag{2.20}
$$

Note that in (2.20), the Fourier coefficients of the perturbation and the unit vectors $e_{\mathbf{k}_i,m}$ are unprimed. Given this, and comparing (2.20) with (2.18), we find that, by the uniqueness of the solutions to these CMEs,

$$\mathbf{A}_{\mathbf{k}_i}(\mathbf{X}_F) = \mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}(\mathbf{T}^{-1}\mathbf{X}_F) \tag{2.21}$$

This identification does assume, however, that the boundary conditions on the envelopes $\mathbf{A}_{\mathbf{k}_i}$ and $\mathbf{T}\mathbf{A}'_{\mathbf{k}'_i}$ are themselves related by (2.21).

Equation (2.21) is a canonical transformation relationship for vector fields related by rotations [35]. It confirms what should have been expected from the outset, that the envelopes in the lattice Δ' are simply rotated versions of those in Δ. The significance of this results becomes patent in light of a class of special cases that we now consider in detail.

Suppose once again that a mode $\varphi_{\mathbf{k}}$ is incident on an NLPC, but is now incident along an axis of M-fold rotational symmetry of the lattice. The first observation to be made is that $\varphi_{\mathbf{k}}$ is coupled to either 1, αM or $1 + \alpha M$ other modes where $\alpha \in \mathbb{Z}^+$.

To see this, let $\Omega = \{\mathbf{k}_i : i \in \{1, 2, \ldots N\}\}$ be the set of wavevectors to which $\mathbf{k}$ is coupled. If $\mathbf{k}$ is taken to lie on a Brillouin zone of the lattice, it coupled either to exactly one other mode or it is coupled to more than one other mode. In the first case, clearly $N = 1$ and the mode to which $\varphi_{\mathbf{k}}$ is coupled is then $\varphi_{-\mathbf{k}}$. Suppose now that $\varphi_{\mathbf{k}}$ is coupled to more than one mode. At least one of these is necessarily not on the pertinent axis of symmetry, that on which $\mathbf{k}$ lies. Label this wavevector $\mathbf{k}_1$. Now, if we consider a lattice Δ' that is simply Δ rotated about $\mathbf{k}$ by an angle of $\frac{2\pi}{M}$ and if this rotation is denoted by $\mathbf{T}^{-1}$, then we find that, in Δ', $\mathbf{k}$ is linearly coupled to $\mathbf{T}^{-1}\mathbf{k}_1$. However, because $\mathbf{k}$ is an axis of M-fold rotational symmetry of the lattice, $\Delta' = \Delta$ and so $\mathbf{k}$ is coupled to $\mathbf{T}^{-1}\mathbf{k}_1$ in Δ as well.

On applying this argument M times, we find that the set of wavevectors to which $\mathbf{k}$ is coupled contains $\{(\mathbf{T}^{-1})^p\mathbf{k}_1 : p = 0, 1, \ldots M - 1\}$. It could be however that not all of the wavevectors in Ω are generated in this way. That is, there may be some wavevectors in Ω that are not related to $\mathbf{k}_1$ by any of the $(\mathbf{T}^{-1})^p$. If we suppose that there are $\alpha - 1$ such wavevectors that themselves do not rotate into one another, then, on applying the argument above to each of these wavevectors, we find $(\alpha - 1)M$ additional coupled modes. The complete set of coupled modes, apart from $-\mathbf{k}$, is then $\Pi = \{(\mathbf{T}^{-1})^p \mathbf{k}_i : i = 1, 2, \ldots \alpha, p = 0, 1, \ldots M-1\}$.

The total set of coupled modes, exclusive of $\mathbf{k}$, is then either $\{-\mathbf{k}\} \cup \Pi$, or Π. Thus, $N = 1, \alpha M$ or $\alpha M + 1$ and our observation is confirmed.

We now consider the mode envelopes $\{\mathbf{A}_{\mathbf{k}_i} : \mathbf{k}_i \in \{\mathbf{k}_i, -\mathbf{k}_i\} \cup \Pi\}$. The wavevectors in Π are taken to be labelled such that $\mathbf{k}_{iM+p} = (\mathbf{T}^{-1})^p \mathbf{k}_i$ for $p = 0, 1, \ldots M-1$ and $i = 1, \ldots \alpha$. Returning again to Δ', we find that for $i \in \{1, \ldots, \alpha\}$, from (2.21), $\mathbf{A}_{\mathbf{k}_i}(\mathbf{X}_F) = \mathbf{T}\mathbf{A}'_{\mathbf{T}^{-1}\mathbf{k}_i}(\mathbf{T}^{-1}\mathbf{X}_F) = \mathbf{T}\mathbf{A}'_{\mathbf{k}_{i+M}}(\mathbf{T}^{-1}\mathbf{X}_F)$.

The salient point to be made here is that, because $\Delta = \Delta'$, $\mathbf{A}'_{\mathbf{T}^{-1}\mathbf{k}_i} = \mathbf{A}_{\mathbf{T}^{-1}\mathbf{k}_i}$ and so $\mathbf{A}_{\mathbf{k}_i}(\mathbf{X}_F) = \mathbf{T}\mathbf{A}_{\mathbf{k}_{i+M}}(\mathbf{T}^{-1}\mathbf{X}_F)$. Indeed, after applying rotations of $\frac{2\pi}{M}p$ for $1 \leq p < M$ and applying the same argument, we find that,

$$\mathbf{A}_{\mathbf{k}_i}(\mathbf{X}_F) = \mathbf{T}^p \mathbf{A}_{\mathbf{k}_{i+pM}}((\mathbf{T}^{-1})^p \mathbf{X}_F) \tag{2.22}$$

That is, for a given $\mathbf{k}_i, 1 \leq i \leq \alpha$, the envelopes $\{\mathbf{A}_{\mathbf{k}_{i+pM}} : p = 0, 1, \ldots M-1\}$ are not independent and are related by (2.22).

In light of these developments, a number of CMEs describing the electric field in the situation being considered can be eliminated using (2.22). In particular, all of the CMEs for the modes $\mathbf{k}_{i+pM}$ with i fixed can be eliminated because their envelopes are given by $(\mathbf{T}^{-1})^p \mathbf{A}_{\mathbf{k}_i}(\mathbf{T}^p \mathbf{X}_S)$. Only the CMEs for $\mathbf{A}_{\mathbf{k}_i}, 1 \leq i \leq \alpha$ need be solved and in them, all envelopes of the form $\mathbf{A}_{\mathbf{k}_{j+pM}}, 1 \leq j \leq \alpha$ can be replaced by $(\mathbf{T}^{-1})^p \mathbf{A}_{\mathbf{k}_j}(\mathbf{T}^p \mathbf{X}_S)$ assuming the initial conditions for the envelopes $\mathbf{A}_{\mathbf{k}_{j+pM}}$ with j fixed, are the same.

2.4 Recurrence Relations

The perturbative methods discussed in section 2.1 generalize to the problem $L\mathbf{E} = \mathbf{0}$ where $\mathbf{E} \in C_3^2$, the space of twice-differentiable mappings from $\mathbb{R}^4$ to $\mathbb{R}^3$, L is a linear differential operator on that space, and both $\mathbf{E}$ and L have expansions in some parameter ξ with the following forms,

$$L = \sum_{m=0}^{\infty} \xi^m L^{(m)}$$
$$\mathbf{E} = \sum_{m=0}^{\infty} \xi^m \mathbf{E}_m$$

In section 2.1, $L^0 = \nabla_F^2 - \nabla(\nabla \cdot) - \frac{n_0^2}{c^2}\frac{\partial^2}{\partial T_F^2}$.

The k^{th} order term in $L\mathbf{E} = \mathbf{0}$ is then $\sum_{m=0}^{k} L^{(m)}\mathbf{E}_{k-m} = \mathbf{0}$, or $L^0\mathbf{E}_k = -\sum_{m=1}^{k} L^{(m)}$ $\mathbf{E}_{k-m}$. With this comes the validity of the claim concerning equation (2.12), namely that the recurrence relations that determine $\mathbf{E}_m$ are differential equations whose homogeneous solutions are those of L^0.

2.5 Summary

The method developed in this chapter provides a means of obtaining a first order approximation to the electric field in any low-contrast nonlinear periodic medium. All that is required to obtain this approximation is the form of the index perturbation, or equivalently its Fourier series representation. Having determined what modes are coupled using the Brillouin zone of the lattice, the approximation can be acquired as the solution to the system of coupled, nonlinear partial differential equations that yields the mode envelopes $\mathbf{A_k}$ (2.17). While in general this system will not have an analytic solution, it will be amenable to numerical methods.

Chapter 3

Applying The Formalism

To illustrate how the method developed in the previous chapter can generate coupled mode equations, we employ it in this chapter in two distinct settings. The first of these will involve two coupled modes in a two-dimensional lattice in which the electric field will, by assumption, be polarized in the direction orthogonal to the lattice. The second setting will consist of three modes coupled in a three-dimensional nonlinear periodic medium, in which no simplifying assumptions will, or indeed can, be made. Numerical simulations of equations obtained using the results of chapter 2 can be found in [36].

3.1 Two Modes In a Two Dimensional Lattice

We begin with a two dimensional lattice taken as lying in the xy-plane and having any lattice geometry. A mode $\varphi_{\mathbf{k}}$ whose wavevector $\mathbf{k}$ lies on an edge of a Brillouin zone of this lattice, but not at the intersection of any edges, is coupled only to one other mode $\varphi_{\mathbf{k}'}$. The wavevector $\mathbf{k}'$ is related to $\mathbf{k}$ by $\mathbf{k}' = \mathbf{k} + \mathbf{G}$ where $\mathbf{G}$ is the reciprocal lattice vector corresponding to the edge on which $\mathbf{k}$ lies. If this mode is assumed to be present in the medium, then, assuming that the leading order term $\mathbf{E}_0$ can be expressed as a sum of $\varphi_{\mathbf{k}}$ and the modes to which it is coupled by the linear perturbation,

$$\mathbf{E}_0(\mathbf{X}_\beta, T_\beta) = \mathbf{e_z}(A_1(\mathbf{X}_S, T_S)\varphi_{\mathbf{k}_1} + A_2(\mathbf{X}_S, T_S)\varphi_{\mathbf{k}_2})$$

where is has been additionally assumed that the field is polarized perpendicular to the lattice plane. This assumption allows for the use of scalar mode envelopes. Moreover, it requires projecting (2.16) only onto solutions $\mathbf{e_k}\varphi_{\mathbf{k}}$ for which $\mathbf{e_k} = \mathbf{e_z}$.

The terms in (2.16) that remain subsequent to projecting on $\mathbf{e_z}\varphi_{\mathbf{k}}$ are $-i(\mathbf{k}_1 \cdot \nabla_S A_1 + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_1}{\partial T_S})$, $\frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}}A_2$ and $\frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A_1(|A_1|^2 + 2|A_2|^2) + \Delta^{nl}_{\mathbf{G}}(A_1^2\overline{A}_2 + A_2(2|A_1|^2 + |A_2|^2)) + \Delta^{nl}_{2\mathbf{G}}\overline{A}_1 A_2^2\big]$. The solvability conditions then require that

$$\begin{aligned} &-i\big(\mathbf{k}_1 \cdot \nabla_S A_1 + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_1}{\partial T_S}\big) + \frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}}A_2 + \frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A_1(|A_1|^2 + 2|A_2|^2) \\ &+\Delta^{nl}_{\mathbf{G}}(A_1^2\overline{A}_2 + A_2(2|A_1|^2 + |A_2|^2)) + \Delta^{nl}_{2\mathbf{G}}\overline{A}_1 A_2^2\big] = 0 \end{aligned}$$

Similarly, projecting onto $\mathbf{e_z}\varphi_{\mathbf{k}'}$ and requiring that the terms vanish results in the second coupled mode equation,

$$\begin{aligned} &-i\big(\mathbf{k}_2 \cdot \nabla_S A_2 + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_2}{\partial T_S}\big) + \frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}}A_1 + \frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A_2(2|A_1|^2 + |A_2|^2) \\ &+\Delta^{nl}_{\mathbf{G}}(A_2^2\overline{A}_1 + A_1(|A_1|^2 + 2|A_2|^2)) + \Delta^{nl}_{2\mathbf{G}}\overline{A}_2 A_1^2\big] = 0 \end{aligned} \tag{3.1}$$

It has been assumed here that the index perturbation $\Delta(\mathbf{X}_F)$ has inversion symmetry so that $\Delta(\mathbf{X}_F) = \Delta(-\mathbf{X}_F)$. This requires that its Fourier components satisfy $\Delta_{\mathbf{G}} = \Delta_{-\mathbf{G}}$. It has also been assumed that the component of the linear index perturbation that is constant in space is zero, so that $\Delta^l_{\mathbf{0}} = 0$. This is equivalent to the unperturbed and perturbed media having matched linear homogeneous index components. Apart from this and the assumption that the electric field is perpendicular to the lattice plane, the equations above are the coupled mode equations for any two modes lying on a single edge of the Brillouin zone of the index lattice.

Beyond permitting the use of scalar envelopes, we see that assuming the field polarization to be orthogonal to the lattice additionally simplifies the CME (3.1) by eliminating the $\mathbf{k} \cdot (\mathbf{e_k} \cdot \nabla_S)\mathbf{A_k}$ term from (2.16). This is a consequence of the relation $\nabla \cdot \epsilon\mathbf{E} = \epsilon\nabla \cdot \mathbf{E} + \mathbf{E} \cdot \nabla\epsilon$. Our assumption about the field requires that $\mathbf{E} \cdot \nabla\epsilon = 0$, inasmuch as $\nabla\epsilon$ has no component orthogonal to the lattice. Thus,

$$\begin{aligned} \nabla \cdot \epsilon\mathbf{E} &= \nabla \cdot \mathbf{D} \\ &= 0 \\ &= \epsilon\nabla \cdot \mathbf{E} \end{aligned}$$

so that $\nabla \cdot \mathbf{E} = 0$. The absence of the terms $\mathbf{k} \cdot (\mathbf{e_k} \cdot \nabla_S)\mathbf{A_k}$ from (3.1) follows from observing that their source is the divergence of $\mathbf{E}$ in the wave equation (2.1).

In the case that $\mathbf{k} = -\frac{\mathbf{G}}{2}$, for some reciprocal lattice vector $\mathbf{G}$, $\mathbf{k}$ necessarily lies on a single edge of the Brillouin zone of the index lattice and is only coupled by $\mathbf{G}$ to $\mathbf{k} + \mathbf{G} = \frac{\mathbf{G}}{2} = -\mathbf{k}$. Labelling the envelope for $\varphi_{\mathbf{k}}$ as A_+ and that for $\varphi_{-\mathbf{k}}$ as A_- and continuing to assume that the field is orthogonal to the lattice plane, the coupled mode equations become

$$\begin{aligned} &-i\big(\mathbf{k} \cdot \nabla_S A_+ + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_+}{\partial T_S}\big) + \frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}} A_- + \frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A_-(|A_-|^2 + 2|A_+|^2) \\ &+\Delta^{nl}_{\mathbf{G}}(A_-^2\overline{A}_+ + A_+(2|A_-|^2 + |A_+|^2)) + \Delta^{nl}_{2\mathbf{G}}\overline{A}_- A_+^2\big] = 0 \end{aligned}$$

$$\begin{aligned} &-i\big(-\mathbf{k} \cdot \nabla_S A_- + \frac{n_0^2\omega_0}{c^2}\frac{\partial A_-}{\partial T_S}\big) + \frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}} A_+ + \frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A_-(2|A_+|^2 + |A_-|^2) \\ &+\Delta^{nl}_{\mathbf{G}}(A_-^2\overline{A}_+ + A_+(|A_+|^2 + 2|A_-|^2)) + \Delta^{nl}_{2\mathbf{G}}\overline{A}_- A_+^2\big] = 0 \end{aligned}$$

A new coordinate system can be defined such that its axes, say X'_S, Y'_S and Z'_S, are respectively along a unit vector in the direction of $\mathbf{k}$ and any two vectors that are orthonormal to $\mathbf{k}$ and to one another. The coordinates of a vector in the original coordinate system have the following relationships with X'_S of the new system.

$$\begin{aligned}
\frac{\partial X_S}{\partial X'_S} &= \hat{k}_x \\
\frac{\partial Y_S}{\partial X'_S} &= \hat{k}_y \\
\frac{\partial Z_S}{\partial X'_S} &= \hat{k}_z (= 0)
\end{aligned}$$

where $\hat{\mathbf{k}} = (\hat{k}_x, \hat{k}_y, \hat{k}_z)$ is a unit vector in the direction of the wavevector $\mathbf{k}$.

Now, if for each envelope a new function $A'_\lambda(\mathbf{X}_S, t)$, $(\lambda = +, -)$ is defined such that $A_\lambda(\mathbf{X}_S, T_S) = A'_\lambda(\mathbf{T^{-1}}\mathbf{X}_S, T_S)$, where $\mathbf{T}$ is the change of basis matrix from the original coordinate system to the new one, then, using the chain rule,

$$\begin{aligned}
\nabla_S A_+ &= \hat{k}_x \frac{\partial A_+}{\partial X_S} + \hat{k}_y \frac{\partial A_+}{\partial Y_S} + \hat{k}_z \frac{\partial A_+}{\partial Z_S} \\
&= \hat{k}_x \frac{\partial A_+}{\partial X_S} + \hat{k}_y \frac{\partial A_+}{\partial Y_S} \\
&= \frac{\partial A'_+}{\partial X'_S}
\end{aligned}$$

Similarly, $\nabla_S A_- = \frac{\partial A'_-}{\partial X'_S}$.

So, finally, the equations above reduce, in the new coordinate system, to

$$\begin{aligned}
&-i\big(\frac{n_0\omega_0}{c}\frac{\partial A'_+}{\partial X'_S} + \frac{n_0^2\omega_0}{c^2}\frac{\partial A'_+}{\partial T_S}\big) + \frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}} A'_- + \frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A'_-(|A'_-|^2 + 2|A'_+|^2) \\
&+\Delta^{nl}_{\mathbf{G}}({A'_-}^2\overline{A'}_+ + A'_+(2|A'_-|^2 + |A'_+|^2)) + \Delta^{nl}_{2\mathbf{G}}\overline{A'}_- {A'_+}^2\big] = 0
\end{aligned}$$

$$-i\Big(-\frac{n_0\omega_0}{c}\frac{\partial A'_+}{\partial X'_S}+\frac{n_0^2\omega_0}{c^2}\frac{\partial A'_-}{\partial T_S}\Big)+\frac{n_0\omega_0^2}{c^2}\Delta^l_{\mathbf{G}}A'_+ +\frac{n_0\omega_0^2}{c^2}\big[\Delta^{nl}_0 A'_-(2|A'_+|^2+|A'_-|^2)$$
$$+\Delta^{nl}_{\mathbf{G}}(A'^{\,2}_-\overline{A'}_+ + A'_+(|A'_+|^2+2|A'_-|^2))+\Delta^{nl}_{\mathbf{2G}}\overline{A'}_-A'^{\,2}_+\big]=0$$

which are the familiar one dimensional coupled mode equations [15].

Thus, for any wavevector $\mathbf{k}$ lying in the plane of a 2-d crystal such that $\mathbf{k}=-\frac{\mathbf{G}}{\mathbf{2}}$ for some reciprocal lattice vector $\mathbf{G}$, the electric field can be expanded using modes whose amplitudes are determined by equations identical to those for a 1-d crystal with a normally incident field. An important consequence of this is that the behaviour of pulses incident on a 2-d crystal from a direction that is parallel to some reciprocal lattice vector and that are orthogonal to the crystal plane will be the same as that of a pulse incident on a one dimensional crystal. Pulses in one dimensional periodic Kerr-nonlinear structures are studied in [37].

When $\mathbf{k}=-\frac{\mathbf{G}}{\mathbf{2}}$ for a reciprocal lattice vector $\mathbf{G}$, the incident field propagates in a direction normal to a family of lattice planes defined by $\mathbf{G}$. The set of lattice planes defined by $\mathbf{G}$ will in this case act as the boundaries between the alternating layers in a 1-d crystals. The pulse will experience what is effectively a square wave index perturbation in one dimension with Fourier components $\Delta^{nl}_{\mathbf{0}},\Delta^{l}_{\mathbf{G}},\Delta^{nl}_{\mathbf{G}}$ and so on.

3.2 Three Modes In a Three Dimensional Lattice

The use of scalar mode envelopes in the previous section was made permissible by assuming the electric field to be polarized in the unique direction in which $\mathbf{E}\cdot\nabla\epsilon=0$ In general, however, this will not be the case and the field in a two-dimensional lattice will have a nonzero component in the lattice plane. Indeed, in a three-dimensionally periodic medium, there exists no direction in which $\mathbf{E}\cdot\nabla\epsilon=0$ inasmuch as no component of $\nabla\epsilon$ is identically zero in such a medium. Moreover, the direction in which the electric field is polarized may change with time and with position in the medium. Under these more general circumstances, the use of vector envelopes in the coupled mode equations is required. In what follows, we obtain coupled mode equations in a three-dimensionally periodic medium in which three modes are linearly coupled.

What is first required in this endeavour is a set of three normal modes, $\varphi_{\mathbf{k}_i}$, that are coupled only to one another by the linear index perturbation. In this case, the corresponding wavevectors differ from another by a reciprocal lattice vector. These reciprocal lattice vectors will be labelled $\mathbf{G}_{ij}$ where $\mathbf{G}_{ij} = \mathbf{k}_j - \mathbf{k}_i$. It will be sufficient here to note that such sets of normal modes do exist and in general can be found by choosing wavevectors that lie at the intersection of exactly two faces of a Brillouin zone of the perturbation lattice. While this method will yield four coupled modes in a simple cubic lattice, it will produce only three in more general lattice structures. For example, in an FCC lattice with lattice constant a, the wavevectors $\frac{2\pi}{a}(\frac{1}{4}, \frac{1}{4}, \pm\frac{1}{4})$ and $\frac{2\pi}{a}(-\frac{7}{16}, -\frac{5}{16}, 0)$ all lie on edges of the first Brillouin zone of the lattice at which no more than two zone faces intersect. Each wavevector is coupled only to itself and the two other wavevectors in the set.

Following the formalism established in the previous chapter, two linearly independent unit vectors, $\mathbf{e}_{\mathbf{k}_j,m}, m = 1, 2$, must be found for each wavevector $\mathbf{k}_j$ such that $\mathbf{e}_{\mathbf{k}_j,m} \cdot \mathbf{k}_j = 0$. Having obtained these vectors, equation (2.14) is projected on $\mathbf{e}_{\mathbf{k}_j,m}\varphi_{\mathbf{k}_j}$ to obtain (2.16) for $j = 1, 2, 3$ and $m = 1, 2$. We label the envelope of mode $\varphi_{\mathbf{k}_j}$ as $\mathbf{A}_j$. Proceeding to find the nonzero projections $\left\langle \varphi_{\mathbf{k}_j}, \varphi_{\mathbf{k}} \right\rangle$ and $\left\langle \varphi_{\mathbf{k}_j}, \varphi_{\mathbf{k}'+\mathbf{k}''-\mathbf{k}'''-\mathbf{G}} \right\rangle$, equation (2.16) becomes

$$
\begin{aligned}
&\Big(\mathbf{k}_i \cdot \nabla_S(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_i) - \frac{1}{2}\mathbf{k}_i \cdot (\mathbf{e}_{\mathbf{k}_i,m} \cdot \nabla_S)\mathbf{A}_i + \frac{n_0^2\omega_0}{c^2}\frac{\partial(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_i)}{\partial T_S}\Big) \\
= \;& \Delta^{nl}_{\mathbf{0}}\Big[\big(\sum_{p=1}^{3}\mathbf{A}_p \cdot \overline{\mathbf{A}_p}\big)(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_i) + (\mathbf{A}_i \cdot \overline{\mathbf{A}_j})(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + (\mathbf{A}_i \cdot \overline{\mathbf{A}_k})(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k)\Big] + \\
& \Delta^{nl}_{\mathbf{G}_{ij}}\Big[(\mathbf{A}_i \cdot \overline{\mathbf{A}_j} + \mathbf{A}_j \cdot \overline{\mathbf{A}_i})(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_i) + \big(\sum_{p=1}^{3}\mathbf{A}_p \cdot \overline{\mathbf{A}_p}\big)(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + (\mathbf{A}_j \cdot \overline{\mathbf{A}_k})(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k)\Big] + \\
& \Delta^{nl}_{\mathbf{G}_{ik}}\Big[(\mathbf{A}_i \cdot \overline{\mathbf{A}_k} + \mathbf{A}_k \cdot \overline{\mathbf{A}_i})(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_i) + \mathbf{A}_k \cdot \overline{\mathbf{A}_j}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + \big(\sum_{p=1}^{3}\mathbf{A}_p \cdot \overline{\mathbf{A}_p}\big)(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k)\Big] + \\
& \Delta^{nl}_{\mathbf{G}_{jk}}\Big[(\mathbf{A}_j \cdot \overline{\mathbf{A}_k} + \mathbf{A}_k \cdot \overline{\mathbf{A}_j})(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_i) + \mathbf{A}_i \cdot \overline{\mathbf{A}_k}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + \mathbf{A}_i \cdot \overline{\mathbf{A}_j}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k)\Big] + \\
& \Delta^{nl}_{\mathbf{2G}_{ij}}\mathbf{A}_j \cdot \overline{\mathbf{A}_i}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + \Delta^{nl}_{\mathbf{2G}_{ik}}\mathbf{A}_k \cdot \overline{\mathbf{A}_i}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k) + \Delta^{nl}_{\mathbf{G}_{ij}+\mathbf{G}_{kj}}\mathbf{A}_j \cdot \overline{\mathbf{A}_k}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + \\
& \Delta^{nl}_{\mathbf{G}_{ik}+\mathbf{G}_{jk}}\mathbf{A}_k \cdot \overline{\mathbf{A}_j}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k)) + \Delta^{nl}_{\mathbf{G}_{ij}+\mathbf{G}_{ik}}\big[\mathbf{A}_k \cdot \overline{\mathbf{A}_i}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_j) + \mathbf{A}_j \cdot \overline{\mathbf{A}_i}(\mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_k)\big] \qquad (3.2)
\end{aligned}
$$

where $\{i, j, k\}$ is a cyclic permutation of $\{1, 2, 3\}$.

The nonlinear terms appearing in (3.2) can be obtained systematically in light of the developments of section 2.2. In this example, S^3 is the set of triplets over $\{1, 2, 3\}$ and $S^3_{1,i}, S^3_{2,i}, \ldots$ are, as usual, subsets of S^3.

For example, the terms involving the Fourier coefficient $\Delta^{nl}_{\mathbf{0}}$ are all characterized by the relations defining $S^3_{1,i}$, which was referred to as $S^3_{1,n}$ in 2.2.

The terms involving the coefficients $\Delta^{nl}_{\mathbf{G_{ij}}}$ and $\Delta^{nl}_{\mathbf{G_{ik}}}$ correspond to triplets in the sets $S^3_{2,i}$ *and* $S^3_{3,i}$. This coincidence of terms results from assuming that the index perturbation possesses inversion symmetry. The result of this is that the Fourier components corresponding to the RLVs characterized by $S^3_{3,i}$, which have the form $\mathbf{G}_{mi}, m \in \{1, 2, 3\}, m \neq i$, satisfy $\Delta^{nl}_{\mathbf{G_{mi}}} = \Delta^{nl}_{\mathbf{G_{im}}}$. Consequently, the terms associated with $\Delta^{nl}_{\mathbf{G_{mi}}}$ can be grouped with those associated with $\Delta^{nl}_{\mathbf{G_{im}}}$, which correspond to $S^3_{2,i}$.

A necessary condition for the nonlinear terms in (3.2) to be correct is that they satisfy the cardinality relations obtained in 2.2. The number of terms proportional to

Subset	Fourier Coefficients	Cardinality
$S_{1,i}$	$\Delta^{nl}_{\mathbf{0}}$	5
$S_{2,i}$	$\Delta^{nl}_{\mathbf{G_{ij}}}$	4
	$\Delta^{nl}_{\mathbf{G_{ik}}}$	4
$S_{3,i}$	$\Delta^{nl}_{\mathbf{G_{ji}}}$	2
	$\Delta^{nl}_{\mathbf{G_{ki}}}$	4
$S_{4,i}$	$\Delta^{nl}_{\mathbf{G_{jk}}}$	4
$S_{5,i}$	$\Delta^{nl}_{\mathbf{2G_{ij}}}$	1
	$\Delta^{nl}_{\mathbf{2G_{ik}}}$	1
$S_{6,i}$	$\Delta^{nl}_{\mathbf{G_{ij}+G_{kj}}}$	1
	$\Delta^{nl}_{\mathbf{G_{ik}+G_{jk}}}$	1
	$\Delta^{nl}_{\mathbf{G_{ij}+G_{ik}}}$	2

Table 3.1: The Fourier coefficient $\Delta^{nl}_{\mathbf{G}}$ is grouped according to the subset of S^3 into which falls the triplet over $\{1, 2, 3\}$ that defines $\mathbf{G}$ (see section 2.2). The cardinalities listed in the third column indicate how many terms in (3.2) are proportional to their corresponding Fourier coefficient.

each fourier coefficient appearing in (3.2) is given in table 3.2 wherein the coefficients are grouped according to the subset of S^3 that they are associated with.

From this table, it is clear that the total number of terms is, as expected, N^3 where $N = 3$, and that the number of terms in each of the subsets $S^3_{1,i}, S^3_{2,i} \ldots$ is equal to the cardinalities presented in section 2.2.

3.3 A Note On Numerical Simulations

As was mentioned in chapter 2, and what may be patent in their general form (2.17), the CMEs are in general analytically intractable. Indeed, the solution of the CMEs requires numerical methods.

The CMEs may nonetheless appear to cause additional difficulty in that the number of terms they contain grows as N^3. The nonlinear terms discussed in section 2.2, each of which has the form $\Delta^{nl}_{\mathbf{G}}\mathbf{A}_{\mathbf{k}_j} \cdot \overline{\mathbf{A}}_{\mathbf{k}_j}(\mathbf{e}_{\mathbf{k}_n} \cdot \mathbf{A}_{\mathbf{k}_i})$, can provide insight into the nature of the nonlinear interaction between the mode envelopes in the field expansion. In particular, they may show which Fourier coefficients in the series (2.5) give rise to cross and self phase modulation between the mode envelopes.

Obtaining this information for each Fourier coefficient will become increasingly diffcult as N increases. However, if the details of the nonlinear interactions given by these terms individually is not pertinent, the large number of terms in the CMEs will not be problematic. This is because the evaluation of the sums appearing in (2.17) for the purposes of numerical simulation need not be done explicitly.

To see this, consider that all of the terms in the first sum in (2.17) can be expressed as a dot product between two N dimensional vectors $\mathbf{u}_{\mathbf{k}_i}$ and $\mathbf{v}$ where $(\mathbf{u}_{\mathbf{k}_i})_p = \Delta^l_{\mathbf{p}-\mathbf{k_i}}$ and $(\mathbf{v})_p = \mathbf{e}_{\mathbf{k}_i,m} \cdot \mathbf{A}_p$ and where N is the number of coupled modes. So, to implementing this sum for each equation, for each $\mathbf{k}_i \in \Omega$, would simply require an N by N matrix $\mathbf{U}$ such that $(\mathbf{U})_{ij} = \Delta^l_{\mathbf{k_j}-\mathbf{k_i}}$. The sum appearing in the equation for $\mathbf{k}_i$ would then be the dot product of the i^{th} row of $\mathbf{U}$ with $\mathbf{v}$.

The second sum appearing in (2.17), that containing the nonlinear terms, can be implemented in a similar way using an N by N by N multidimensional array $\mathbf{W}$ such that $(\mathbf{W})_{qrs} = \Delta^{nl}_{\mathbf{k}_q+\mathbf{k}_r-\mathbf{k}_s-\mathbf{k}_i}$ in the equation for $\mathbf{k}_i$.

The salient point to be made here is that the potentially complicated nature of the CMEs (2.17) should make their numerical solution no more complicated than any equation of the form $L\mathbf{A}_{\mathbf{k}} = f(\mathbf{A}_{\mathbf{k}}, \mathbf{A}_{\mathbf{k_2}}, \ldots, \mathbf{A}_{\mathbf{k_N}})$ for any function $f : \mathbb{C}^N \to \mathbb{C}$ and some continuous operator $L : \mathbb{C}^3 \to \mathbb{C}$.

3.4 Summary

The formalism developed in chapter two was deployed in this chapter in the first instance to obtain some general results concerning two coupled modes in two dimensional lattices, and in the second case to obtain equations governing any three coupled modes in a three dimensional lattice. Equations, and in particular nonlinear terms, for the most general circumstances can be derived following precisely the same method detailed in section 3.2, which itself relied on the details of section 2.2.

Chapter 4

Discussion and Conclusion

We have developed a coupled mode theory for low-contrast nonlinear photonic crystals, the central result of which is equation (2.17).

Though the formalism offers greater generality than previous attempts to construct coupled mode theory, it is not completely general. In addition to ways in which this work could be augmented, we discuss below future work that would lend to its completeness.

4.1 Future Work

Though this work is suited to the analysis of very general low-contrast nonlinear photonic crystals, it is not applicable to the most general periodic nonlinear media. This is patent, indeed by definition, in that it is explicitly intended for photonic crystals in which large index contrasts are absent. It is more subtly evident, however, in some approximations made in the derivation of Chapter 2.

For example, the wave equation from which the CMEs are obtained is itself an approximation, as is our truncation of the square of the index of refraction in (2.4). These approximations, however, are reasonable in light of our assumptions of slowly varying fields in low-contrast media. In any case, the validity of our assumptions must be tested against experiment.

Perhaps the most significant approximation is that the electric field is accurately represented by a zeroth-order expansion in ξ. This awaits rigourous demonstration, perhaps beginning with section 2.4.

The limit at which this assumption is no longer true is that at which the unperturbed medium index contrast to too large for normal modes therein to be plane waves. In this limit, they must be Bloch modes, as in [24, 27, 28] and more recently, [38].

In light of this, and to determine if the coupled mode theory developed here agrees with work aimed at high-contrast media, it would be fruitful to determine whether works such as [24, 27, 28] reduce in the limit of low-contrast to a theory resembling ours.

Our theory could additionally be supplemented by considering anisotropic media, as is done for high-contrast media in [39, 40]

4.2 Conclusion

In summary, we have developed a method sufficient for obtaining CMEs in any Kerr-nonlinear, low-contrast photonic crystal. The full generality of the method awaits to be exploited by, for example, its application to 3-d crystals, 1-d or 2-d crystals with arbitrary field polarizations, or to time dependent problems.

Bibliography

[1] C. Kittel. *Introduction to Solid State Physics.* Wiley, New York, 1996.

[2] S. John. *E 315 - NATO ASI SERIES Photonic Band Gap Materials, edited by C.M. Soukoulis.* Kluwer Academic Publishers, The Netherlands, 1996.

[3] E. Yablonovitch. Inhibited spontaneous emission in solid-state physics and electronics. *Physical Review Letters*, 58:2059–2062, 1987.

[4] S. John. Strong localization of photons in certain disordered dielectric superlattices. *Physical Review Letters*, 58:2486–2489, 1987.

[5] E. R. Brown and O. B. McMahon. High zenithal directivity from a dipole antenna on a photonic crystal. *Applied Physics Letters*, 68:1300–1302, 1996.

[6] S. Y. Lin, E. Chow, V. Hietala, P. R. Villeneuve, and J. D. Joannopoulos. Experimental demonstration of guiding and bending of electromagnetic waves in a photonic crystal. *Science*, 282:274–276, 1998.

[7] M.Charbonneau-Lefort, E. Istrate, M. Allard, J. Poon, and E. H. Sargent. Photonic crystal heterostructures: Waveguiding phenomena and methods of solution in an envelope function picture. *Physical Review B*, 65:125318, 2002.

[8] K. Busch and S. John. Liquid-crystal photonic-band-gap materials: The tunable electromagnetic vacuum. *Physical Review Letters*, 83:967–970, 1999.

[9] R. Menzel. *Photonics : linear and nonlinear interactions of laser light and matter.* Springer, New York, 2001.

[10] Chai Yeh. *Applied Photonics.* Academic Press, 1994.

[11] G. P. Agrawal. *Nonlinear Fiber Optics.* Academic Press, San Diego, 1989.

[12] B. J. Eggleton, R. E. Slusher, C. M. de Sterke, P. A. Krug, and J. E. Sipe. Bragg grating solitons. *Physical Review Letters*, 76:1627–1630, 1996.

[13] W. Chen and D. L. Mills. Gap solitons and the nonlinear optical response of superlattices. *Physical Review Letters*, 58:160–163, 1987.

[14] D. L. Mills and S. E. Trullinger. Gap solitons in nonlinear periodic structures. *Physical Review B*, 36:947–952, 1987.

[15] D. Pelinvosky, J. Sears, L. Brzozowski, and E. H. Sargent. Stable all-optical limiting in nonlinear periodic structures. i. analysis. *Journal of the Optical Society of America B*, 19:43–53, 2002.

[16] C. J. Herbert, W. S. Capinsky, and M. S. Malcuit. Optical power limiting with nonlinear periodic structures. *Optics Letters*, 17:1037–1039, 1992.

[17] L. Brzozowsky and E. H. Sargent. Optical signal processing using nonlinear distributed feedback structures. *IEEE Journal of Quantum Electronics*, 36:550–555, 2001.

[18] P. Tran. All-optical switching with a nonlinear chiral photonic bandgap structure. *Journal of the Optical Society of America B*, 16:70–73, 1999.

[19] N. G. R. Broderick, D. Taverner, and D. J. Richardson. Nonlinear switching in fiber bragg gratings. *Optics Express*, 3:447–453, 1998.

[20] N. D. Sankey, D. F. Prelewitz, and T. G. Brown. All-optical switching in a nonlinear periodic-waveguide structure. *Applied Physics Letters*, 60:1427–1429, 1992.

[21] G. Assanto and G. I. Stegeman. Optical bistability in nonlocally nonlinear periodic structures. *Applied Physics Letters*, 56:2285–2287, 1990.

[22] J. He and H. E. Cada. Optical bistability in semiconductor periodic structures. *IEEE Journal of Quantum Electronics*, 27:1182–1188, 1991.

[23] L. Brzozowsky and E. H. Sargent. All-optical analog-to-digital converters, hardlimiters and logic gates. *Journal of Lightwave Technology*, 19:114–119, 2001.

[24] C. M. de Sterke and J. E. Sipe. Envelope function approach for the electrodynamics of nonlinear periodic structures. *Physical Review A*, 38:5149–5165, 1988.

[25] A Babin and A. Figotin. Nonlinear photonic crystals: I. quadratic nonlinearity. *Waves in Random Media*, 11:R31–R102, 2001.

[26] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics*. Wiley, New York, 1999.

[27] S. Pereira and J. E. Sipe. Canonical hamiltonian formulation of the nonlinear schrodinger equation in a one-dimensional, periodic kerr medium. *Physical Review E*, 65:046601, 2002.

[28] S. Pereira and J. E. Sipe. Hamiltonian formulation of the nonlinear coupled mode equations. *Physical Review E*, 66:026606, 2002.

[29] N. A. R. Bhat and J. E. Sipe. Optical pulse propagation in nonlinear photonic crystals. *Physical Review E*, 64:056604, 2001.

[30] L. Brzozowski, V. Sukhovatkin, E. H. Sargent, A. J. SpringThorpe, and M. Extavour. Intensity-dependent reflectance and transmittance of semiconductor periodic structures. *IEEE Journal of Quantum Electronics*, 39:924–930, 2003.

[31] H. Lee and G. P. Agrawal. Nonlinear switching of optical pulses in fiber bragg gratings. *IEEE Journal of Quantum Electronics*, 39:508–515, 2003.

[32] N. Akozbek and S. John. Self-induced transparency solitary waves in a doped nonlinear photonic band gap material. *Physical Review E*, 58:3876–3895, 1998.

[33] C. M. de Sterke and J.E. Sipe. Switching dynamics of finite periodic nonlinear media : A numerical study. *Physical Review A*, 42:2858–2872, 1990.

[34] A. H. Nayfeh. *Perturbation Methods.* Wiley, New York, 1973.

[35] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory.* Westview Press, 1995.

[36] J. L. Sheriff, I. A. Goldthorpe, and E. H. Sargent. Optical limiting and intensity-dependent diffraction from low-contrast nonlinear periodic media: Coupled-mode analysis. *Physical Review E*, (Submitted,2004).

[37] Winnie N. Ye, L. Brzozowski, D. Pelinvosky, and E. H. Sargent. Stable all-optical limiting in nonlinear periodic structures. iii. nonsolitonic pulse propagation. *Journal of the Optical Society of America B*, 20:695–705, 2003.

[38] J.E. Sipe, N.A. Bhat, P. Chak, and S. Pereira. Effective field theory for the nonlinear optical properties of photonic crystals. *Physical Review E*, 69:016604, 2004.

[39] S. Pereira and J.E. Sipe. Light propagation through birefringent, nonlinear media with deep gratings. *Physical Review E*, 62:5745–5757, 2000.

[40] S. Pereira and J.E. Sipe. Nonlinear pulse propagation in birefringent fiber bragg gratings. *Optics Express*, 3:418–432, 1998.