Part I

Mathematical Formalism

Chapter 2

The Helmholtz Equation

The underlying physics of photonic crystals follow Maxwell's equations, and in this chapter we derive from first principles the wave equation on which the work in this thesis is based. There are many methods for solving the resulting partial differential equation (PDE), but the two most pervasive methods utilize either a time domain approach (e.g., finite difference time domain–FDTD)[2], or a frequency domain approach[3], and each method has its own merits[4].

The time domain method, as the name implies, is well suited for studying dynamical properties of the fields such as pulse propagation, transmission/reflection properties, estimate losses, etc. However, it is not as suitable for looking at resonant behavior with a single frequency of interest. While it is possible to use FDTD to look for resonant structures, one of the disadvantages is that computationally it is less efficient than the frequency domain approach, because you are looking at the system's entire response to a driving term. Another drawback of FDTD is that one must be careful that the source term is not orthogonal to the mode of interest, in other words, that there is sufficient coupling, or you risk not even finding the resonance. As the linewidth of the cavity decreases, the required precision of the frequency increases, which implies an increase in the number of timesteps required for the computation. However, the most notable advantage of FDTD is that it scales more favorably [5] than frequency domain methods.

As our group's interest in photonic crystals began with PBG cavities, it was natural to adopt the frequency domain approach. We will see in this chapter that the

frequency domain problem using a plane wave basis reduces to a Hermitian eigenvalue problem, so there are many similarities to problems encountered in elementary quantum mechanics. However, there are some known convergence issues with this technique, and it is important that we draw attention to and highlight these in section 2.4.

For a more general overview for studying the physics of photonic crystals, the standard reference is the book by Joannopoulos et al. [6], or for a more mathematical treatment, there is Sakoda's book [7] as well.

Organization

We begin this chapter by deriving from Maxwell's equations the wave equation for an inhomogeneous medium. We show in section 2.2 the plane wave expansion method for solving the Helmholtz equation to find photonic bandstructures. Adapting the method for studying PBG materials with defect inclusions is reviewed in section 2.3. Readers familiar with the method can safely omit those sections. We conclude in section 2.4 with a discussion of the general convergence issues of the method. Even though the information in that section is not new, there are too many references in more current literature that seem unaware of the issues.

2.1 Bulk Photonic Crystal

Consider a position-dependent non-magnetic medium $\epsilon(\boldsymbol{r})$ with no charge or current sources. We further restrict ourselves to consider only linear and scalar (i.e., isotropic) materials. Maxwell's equations for the various fields take on the following form:

$$\nabla \cdot \boldsymbol{D}(\boldsymbol{r},t) = 0 \tag{2.1}$$

$$\nabla \cdot \boldsymbol{B}(\boldsymbol{r},t) = 0 \tag{2.2}$$

$$\nabla \times \boldsymbol{H}(\boldsymbol{r},t) = \frac{\partial \boldsymbol{D}(\boldsymbol{r},t)}{\partial t} \tag{2.3}$$

$$\nabla \times \boldsymbol{E}(\boldsymbol{r},t) = -\frac{\partial \boldsymbol{B}(\boldsymbol{r},t)}{\partial t}. \tag{2.4}$$

Using the constitutive relations, we can eliminate $\boldsymbol{D}$ and $\boldsymbol{B}$.

$$\boldsymbol{D}(\boldsymbol{r},t) = \epsilon(\boldsymbol{r})\boldsymbol{E}(\boldsymbol{r},t) \tag{2.5}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \mu_0\boldsymbol{H}(\boldsymbol{r},t) \tag{2.6}$$

Note that unlike the standard derivation of the wave equation for a homogeneous dielectric medium, we no longer have $\nabla \cdot \boldsymbol{E} = 0$ from $\nabla \cdot \boldsymbol{D} = 0$ because in general $\nabla\epsilon(\boldsymbol{r}) \neq 0$. We can still work out the wave equations for $\boldsymbol{E}$ and $\boldsymbol{H}$ as follows.

2.1.1 Wave equation for $\boldsymbol{E}$

We take the curl of eqn. (2.4) and combine with eqn. (2.6) to obtain

$$\nabla \times \big(\nabla \times \boldsymbol{E}(\boldsymbol{r},t)\big) = -\mu_0\frac{\partial}{\partial t}\nabla \times \boldsymbol{H}(\boldsymbol{r},t) \tag{2.7}$$

We substitute in eqn. (2.3) and combine with eqn. (2.5) to arrive at the wave equation.

$$\nabla \times \big(\nabla \times \boldsymbol{E}(\boldsymbol{r},t)\big) + \mu_0\epsilon(\boldsymbol{r})\frac{\partial^2\boldsymbol{E}(\boldsymbol{r},t)}{\partial t^2} = 0 \tag{2.8}$$

The standard frequency domain method then enforces harmonic time dependence of the field to arrive at the Helmholtz equation for the electric field $\boldsymbol{E}$.

$$\boldsymbol{E}(\boldsymbol{r},t) = \boldsymbol{E}(\boldsymbol{r})e^{i\omega t} \tag{2.9}$$

$$\therefore \nabla \times \big(\nabla \times \boldsymbol{E}(\boldsymbol{r},t)\big) = \epsilon_r(\boldsymbol{r})\frac{\omega^2}{c^2}\boldsymbol{E}(\boldsymbol{r}), \tag{2.10}$$

where $\epsilon_r(\boldsymbol{r}) \equiv \frac{\epsilon(\boldsymbol{r})}{\epsilon_0}$. It turns out that the left side of the equation is not self-adjoint (i.e. not Hermitian), so we are not guaranteed a complete basis [8]. Since the $\boldsymbol{H}$ equation (to be derived below) is self-adjoint, we will work exclusively with that equation instead, but for completeness we have presented the equation for $\boldsymbol{E}$ here. Of course, once we have the $\boldsymbol{H}$ field, we can use eqn. (2.3) and (2.5) to obtain $\boldsymbol{E}$.

2.1.2 Wave equation for $\boldsymbol{H}$

The derivation for the $\boldsymbol{H}$ equation is similar to the $\boldsymbol{E}$ equation. We begin by combining eqn. (2.3 and 2.5), take the curl and then incorporate eqn. (2.4 and2.6) to obtain the following:

$$\begin{aligned}
\frac{1}{\epsilon_r(\boldsymbol{r})}\nabla \times \boldsymbol{H}(\boldsymbol{r},t) &= \epsilon_0 \frac{\partial \boldsymbol{E}(\boldsymbol{r},t)}{\partial t} && (2.11)\\
\nabla \times \left(\frac{1}{\epsilon_r(\boldsymbol{r})}\nabla \times \boldsymbol{H}(\boldsymbol{r},t)\right) &= \epsilon_0 \frac{\partial}{\partial t}\nabla \times \boldsymbol{E}(\boldsymbol{r},t) && (2.12)\\
\nabla \times \left(\frac{1}{\epsilon_r(\boldsymbol{r})}\nabla \times \boldsymbol{H}(\boldsymbol{r},t)\right) &= -\mu_0\epsilon_0 \frac{\partial^2 \boldsymbol{H}(\boldsymbol{r},t)}{\partial t^2} && (2.13)\\
\nabla \times \big(\eta(\boldsymbol{r})\nabla \times \boldsymbol{H}(\boldsymbol{r},t)\big) &= -\mu_0\epsilon_0 \frac{\partial^2 \boldsymbol{H}(\boldsymbol{r},t)}{\partial t^2}, && (2.14)
\end{aligned}$$

where $\eta(\boldsymbol{r}) \equiv (\epsilon_r \boldsymbol{r})^{-1}$ is the reciprocal of the relative permittivity function. To keep the language from being overly cumbersome, we will refer to η simply as the 'dielectric function' for the rest of this thesis.

Invoking harmonic time dependence, we arrive at our Helmholtz equation for the magnetic field $\boldsymbol{H}$.

$$\nabla \times \big(\eta(\boldsymbol{r})\nabla \times \boldsymbol{H}(\boldsymbol{r})\big) = \frac{\omega^2}{c^2}\boldsymbol{H}(\boldsymbol{r}) \qquad (2.15)$$

We also point out that in the derivation, we have not used two of Maxwell's equations (Eqn. (2.1) and (2.2)), which means that a general solution to the Helmholtz equation will not obey all of Maxwell's equations. One must check or enforce these conditions in order to have physically realizable fields.

2.2 2D Plane Wave Expansion (PWE) Method

Since equation (2.15) is self-adjoint we are therefore guaranteed a complete basis. This implies that any spatially dependent function $f(\boldsymbol{r})$ can be expressed as a linear

combination (i.e. superposition) of basis states.

$$f(\boldsymbol{r}) = \int a(\boldsymbol{k})\phi(\boldsymbol{k},\boldsymbol{r})d\boldsymbol{k} \tag{2.16}$$

$$\text{with} \qquad \int \phi(\boldsymbol{k},\boldsymbol{r})\phi(\boldsymbol{k'},\boldsymbol{r})d\boldsymbol{r} = \delta(\boldsymbol{k}-\boldsymbol{k'}), \tag{2.17}$$

where $\delta(\boldsymbol{k}-\boldsymbol{k'})$ is the Kronecker delta. The plane wave expansion method uses the set of planes waves $\{\exp(i\boldsymbol{k}\cdot\boldsymbol{r})\}$ as the complete basis. One reason for choosing the plane wave basis is that it is the eigenbasis in free space, so it is required for pseudo Q factor considerations [9]. Another significant advantage to using plane waves is that the vector fields will be divergence free (by construction), so all of Maxwell's equations (eqn. (2.1 and 2.2) in particular) are followed.

While it is certainly possible in principle to treat equation (2.15) vectorially in full 3 dimensions (3D) [10, 11], we will be making a 2D approximation for two reasons. First, current fabrication technology has limited almost all PBG devices to take on quasi-2D form, where the bandgap effect is utilized in only 2 dimensions and the transverse dimension relies on total internal reflection for optical confinement. Second, the computational resources required for a full 3D treatment of PBG devices using the plane wave method is currently prohibitively expensive[1]. A final advantage of the 2D treatment is that the TE (transverse electric) and TM (transverse magnetic) polarizations are decoupled.

To be more precise, when we refer to a 2D approximation, we mean that the dielectric function has full translational invariance in one dimension (say along the z direction), so $\frac{\partial\eta}{\partial z} = 0$. Under this configuration, the TE polarization has the electric field vectors lying in the xy plane, and the magnetic field points along the z direction. The TM polarization has the magnetic field in the xy plane and the electric field along z.

[1] See [4] summarized in section 2.4 for an exception to this statement.

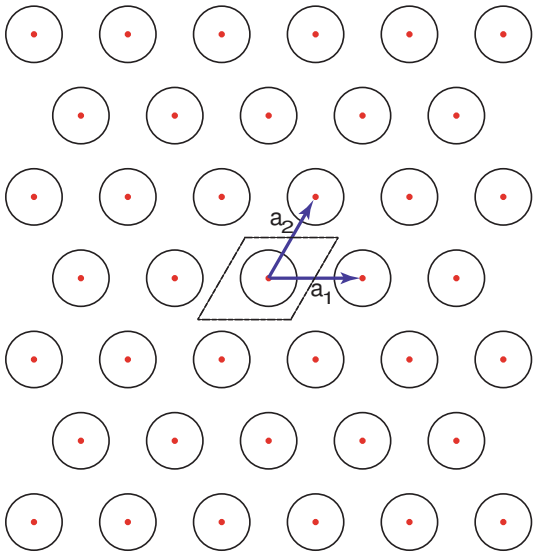

Figure 2.1: 2D hexagonal lattice in real space. The unit cell (dashed line) and the primitive vectors $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ are shown.

2.2.1 Boundary conditions

For the case of photonic crystals, the dielectric function has translational symmetry along the primitive lattice vectors $\{\boldsymbol{a_1}, \boldsymbol{a_2}\}$ in the xy plane such that $\eta(\boldsymbol{r}) = \eta(\boldsymbol{r}+\boldsymbol{R})$, $\boldsymbol{R} \equiv m_1\boldsymbol{a_1} + m_2\boldsymbol{a_2}$, where $\{m_i\}$ are integers. For primarily historical reasons, we choose as our canonical geometry a hexagonal (sometimes referred to as a triangular) lattice of cylindrical air rods in semiconductor (see figures 2.1 and 2.2). The lattice symmetry allows us to define Born von-Karman periodic boundary conditions (BCs) for our PDE, as well as define a unit cell. Again, this naturally leads us to a description in Fourier space, where because of the periodicity we also have a lattice (the reciprocal lattice). The set of reciprocal lattice vectors $\{\boldsymbol{G}\}$ is defined by the following symmetry requirement:

$$e^{i\boldsymbol{G}\cdot(\boldsymbol{r}+\boldsymbol{R})} = e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \tag{2.18}$$

$$\therefore e^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1. \tag{2.19}$$

We can leverage much of our intuition from solid state physics [12], and in fact, photonic crystals are often viewed as 'semiconductors for light.' In particular, Bloch's

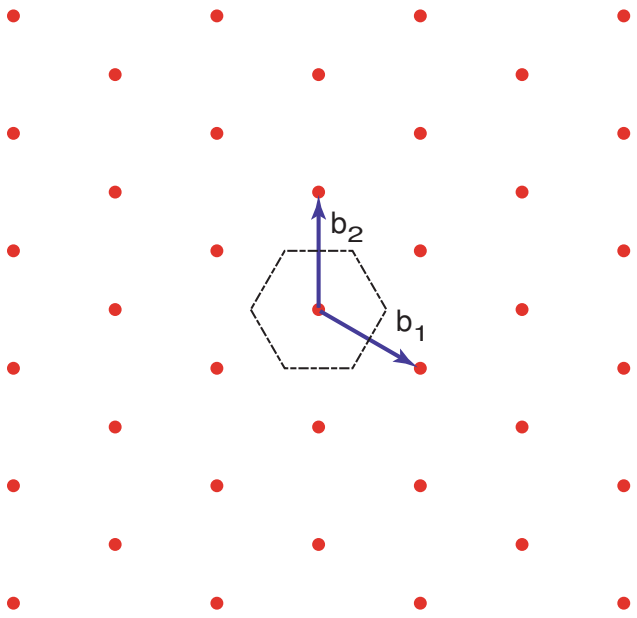

Figure 2.2: Reciprocal lattice in Fourier space of hexagonal lattice of figure 2.1

theorem applies (see chapter 8 in [12]), and we need only solve for the Bloch modes. A final warning is that in invoking the Born von-Karman BCs, we have implicitly assumed a photonic crystal of infinite extent. While this is a very standard approximation, and it is a good approximation if the physical structure is large compared to the lattice constant, it is important to be clear that the periodicity is meant to extend to infinity.

Translational symmetry implies we can express:

$$\eta(\boldsymbol{r}) = \sum_{\mathbf{G}} \eta_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \tag{2.20}$$

$$\therefore \eta_{\boldsymbol{G}} = \frac{1}{A_c}\int_{A_c} \eta(\boldsymbol{r}) e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} d^2\boldsymbol{r} \tag{2.21}$$

$$\boldsymbol{H}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} h_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\ \hat{\boldsymbol{z}} \tag{2.22}$$

$$\therefore h_{\boldsymbol{G}} = \frac{1}{A_c}\int_{A_c} \boldsymbol{H}(\boldsymbol{r})\cdot\hat{\boldsymbol{z}}\, e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} d^2\boldsymbol{r}, \tag{2.23}$$

where A_c denotes the area of the unit cell, and in eqn. (2.22) we have chosen the TE polarization. Eqn. (2.23) expresses the Fourier coefficient for some given Bloch

mode. In calculating band structures, the relationship is true though not helpful, as we are trying to solve for the unknown Bloch mode.

We substitute into eqn. (2.15) the expressions in eqn. (2.20) and (2.22).

$$\begin{aligned}
\frac{\omega^2}{c^2}\sum_{\boldsymbol{G}} h_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\ \hat{\boldsymbol{z}} &= \nabla\times\left(\left(\sum_{\boldsymbol{G'}}\eta_{\boldsymbol{G'}}e^{i\boldsymbol{G'}\cdot\boldsymbol{r}}\right)\nabla\times\left(\sum_{\boldsymbol{G}} h_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\ \hat{\boldsymbol{z}}\right)\right)\\
&= \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\nabla\times e^{i\boldsymbol{G'}\cdot\boldsymbol{r}}\nabla\times e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\ \hat{\boldsymbol{z}}\\
&= \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\nabla\times e^{i\boldsymbol{G'}\cdot\boldsymbol{r}}(i\boldsymbol{G}\times\hat{\boldsymbol{z}})e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\\
&= \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\nabla\times e^{i(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{r}}(i\boldsymbol{G}\times\hat{\boldsymbol{z}})\\
&= \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\big[i(\boldsymbol{G}+\boldsymbol{G'})\times(i\boldsymbol{G}\times\hat{\boldsymbol{z}})\big]e^{i(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{r}}\\
&= \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\big[(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{G}\big]e^{i(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{r}}\hat{\boldsymbol{z}},
\end{aligned}$$

where in the last step we made use of the triple cross product identity $a\times(b\times c) = b(a\cdot c) - c(a\cdot b)$ and the fact that $\{\mathbf{G}\}$ lie in the xy plane. We then take the inner product of both sides with a TE polarized plane wave by left multiplying and then integrating:

$$\frac{1}{A_c}\int_{A_c}\hat{\boldsymbol{z}}\,e^{-i\boldsymbol{G''}\cdot\boldsymbol{r}}\left\{\frac{\omega^2}{c^2}\sum_{\boldsymbol{G}} h_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}\ \hat{\boldsymbol{z}} = \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\big[(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{G}\big]e^{i(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{r}}\hat{\boldsymbol{z}}\right\}d^2\boldsymbol{r}$$

$$\frac{\omega^2}{c^2}\sum_{\boldsymbol{G}} h_{\boldsymbol{G}}\delta_{\boldsymbol{G},\boldsymbol{G''}} = \sum_{\boldsymbol{G},\boldsymbol{G'}}\eta_{\boldsymbol{G'}}h_{\boldsymbol{G}}\big[(\boldsymbol{G}+\boldsymbol{G'})\cdot\boldsymbol{G}\big]\delta_{\boldsymbol{G}+\boldsymbol{G'},\boldsymbol{G''}}$$

We choose to collapse the delta function on the right with $\boldsymbol{G'} = \boldsymbol{G''} - \boldsymbol{G}$. Relabeling

the indices, we arrive at the Helmholtz equation in the plane wave basis.

$$\sum_{\boldsymbol{G'}} \left[\eta_{\boldsymbol{G}-\boldsymbol{G'}}\boldsymbol{G}\cdot\boldsymbol{G'}\right] h_{\boldsymbol{G'}} = \frac{\omega^2}{c^2} h_{\boldsymbol{G}} \tag{2.24}$$

This is the starting point for bandstructure calculations of bulk (i.e. defect-free) photonic crystals. So far we have only considered **k**-points that are coupled to the $\boldsymbol{q} = 0$ component, where $\{\boldsymbol{q}\}$ denotes the set of **k**-vectors that lie in the first Brillouin zone (unit cell in reciprocal space). To complete the bandstructure to include the other **k**-points, we simply use a different expansion of the field.

$$\boldsymbol{H}_{\boldsymbol{q}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} h_{\boldsymbol{q}+\boldsymbol{G}} e^{i(\boldsymbol{q}+\boldsymbol{G})\cdot\boldsymbol{r}}\ \hat{\boldsymbol{z}} \tag{2.25}$$

Here, $\boldsymbol{q}$ labels the plane wave that is used to modulate the Bloch function. In general, the Bloch functions for different $\boldsymbol{q}$ will be different, as (following the same steps as in the above derivation) we arrive at the following:

$$\sum_{\boldsymbol{G'}} \left[\eta_{\boldsymbol{G}-\boldsymbol{G'}}(\boldsymbol{q}+\boldsymbol{G})\cdot(\boldsymbol{q}+\boldsymbol{G'})\right] h_{n,\boldsymbol{q}+\boldsymbol{G'}} = \frac{\omega_{n,q}^2}{c^2} h_{n,\boldsymbol{q}+\boldsymbol{G}} \tag{2.26}$$

Written in this form, each $\boldsymbol{q}$ within the first Brillouin zone labels a unique eigenvalue problem, and the n labels the band index of a particular mode. We do not need to separately consider **k**'s outside the first Brillouin zone because all **k**'s that differ by a translation of some **G**' in {**G**} are coupled to each other. We use the band index to keep track of the **k**'s in outer Brillouin zones.

Eigenfunctions that satisfy eqn. (2.26) are called 'bulk modes'. The eigenvalues give us the frequencies that are permitted within the material. When we solve a series of these eigenvalue problems for $\boldsymbol{q}$ along high symmetry points, we obtain the band diagram for the material. For a distribution of material with sufficient dielectric contrast (e.g. low-index air rods in high-index semiconductors), we find certain ranges where the frequencies are forbidden, which means that light at those frequencies cannot propagate in the material, hence the term 'bandgap.'

However, most useful PBG devices require that the symmetry be broken by introducing *defects* in the lattice. Defects can be an omission of an air rod, or an air rod of a different size or shape, or any other structure that causes a break in the symmetry of the system. Structures with defects are studied in the PWE method using the supercell approximation.

2.3 Supercell Treatment

The strategy here is to surround the defect(s) with enough layers of the bulk photonic crystal that the modes of interest become well localized within the defect region. We can then invoke the tight binding approximation and consider the defect plus surrounding layers as a 'unit cell,' or a supercell. This implies a lattice of defects that extend to infinity (see figures 2.3 and 2.4).

This is another standard approximation, and is valid if there is minimal interaction between the artificial neighboring defect sites. This is analogous to the tight-binding approximation in solid state physics. Note that for the line defect, the defects along the x-axis should not be considered artifacts of the supercell method, since there is a real translational symmetry in that direction (i.e. the waveguiding direction). When solving for the modes of the waveguide, we need not be alarmed or concerned if the mode is not localized in the x direction. However, that is not true in the y direction. Finally, as with the bulk photonic crystal case, the translational symmetry of the real waveguide does not extend to infinity even in the x direction.

2.3.1 Point defect: cavity

For the case of the point defect cavity, the extension to the Helmholtz equation is trivial. Instead of the set $\{\mathbf{G}\}$, we now have a new set of reciprocal lattice vectors which we will refer to as $\{\mathbf{k}\}$. We no longer have to worry about $\{\boldsymbol{q} \neq 0\}$ in the Brillouin zone since they do not form a valid expansion. As we do not actually have a supercell periodicity, physically we cannot support these longer wavelength plane

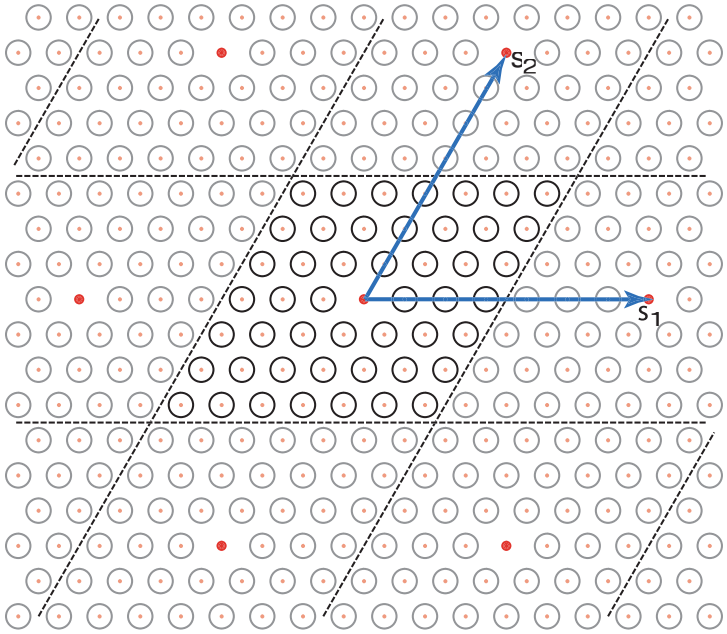

Figure 2.3: Real space dielectric function of a hexagonal lattice of air cylinders embedded in an infinitely thick block of semiconductor. The point defect is formed by omission of the central air cylinder. In the supercell treatment, the entire device makes up the 'unit cell' (outlined in the dotted line) is artificially tiled to give periodic boundary conditions. The original lattice points are shown as the faint red dots, but hold no particular special significance in the supercell treatment. The new superlattice points are the larger red dots, with the superlattice vectors $\boldsymbol{s_1}$ and $\boldsymbol{s_2}$ shown. The satellite defect structures are included but shown slightly faded out.

wave modulation of Bloch modes. The Helmholtz equation then simply sets $\boldsymbol{q} = 0$ and substitute $\mathbf{k}$'s for $\mathbf{G}$'s.

In principle, the set $\{\mathbf{k}\}$ extends to infinity, although in practice we always truncate at some finite bandwidth. There are some implications to truncation which are often overlooked. We discuss these and other subtle issues with the discrete fourier transform in appendix C. With a truncated basis, we can now obtain the Helmholtz operator in matrix form.

$$\left[\sum_{\boldsymbol{k'}} \eta_{\boldsymbol{k}-\boldsymbol{k'}}(\boldsymbol{k}\cdot\boldsymbol{k'})\right] h_{\boldsymbol{k'}} = \frac{\omega^2}{c^2} h_{\boldsymbol{k}} \tag{2.27}$$

$$\hat{\boldsymbol{\Theta}}_{\boldsymbol{k}\boldsymbol{k'}}{}^{(\eta)} h_{\boldsymbol{k'}} = \frac{\omega^2}{c^2} h_{\boldsymbol{k}} \tag{2.28}$$

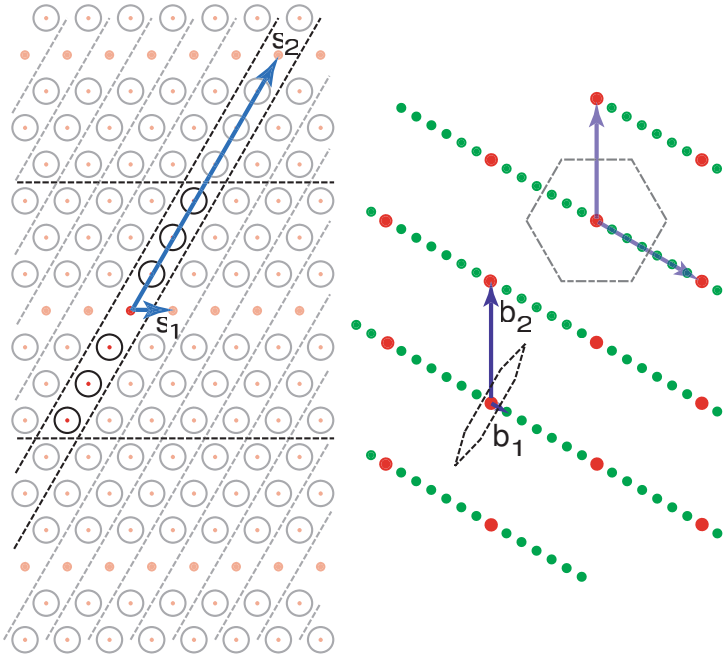

Figure 2.4: On the left side of the figure, we show the real space dielectric function of a row defect in the supercell treatment. Small red dots correspond to original lattice, and the larger red dots the 'basis vectors' for the superlattice. On the right is a picture of the reciprocal lattice. The red dots show the reciprocal lattice of the bulk photonic crystal, and the green dots show the reciprocal lattice of the new supercell. The shape of the original and supercell Brillouin zones are shown as well, with the original one shown faded.

We have omitted the band index for clarity, since in general we will only be interested in a small number of modes that lie within the bandgap. We denote $\hat{\boldsymbol{\Theta}}_{\boldsymbol{k}\boldsymbol{k}'}{}^{(\eta)}$ as the Helmholtz operator, and the problem of finding the eigenmodes $\boldsymbol{H}_{\boldsymbol{m}}(\boldsymbol{r})$ for some given distribution of dielectric will be referred to as the 'forward problem.' The superscript η makes it explicit that the operator depends on the chosen dielectric function. We can form the Helmholtz operator once we have the Fourier coefficients $\eta_{\boldsymbol{k}}$, which can be obtained analytically or numerically. It is important to note here that we set $\eta_{\boldsymbol{\kappa}} = 0$ for $\boldsymbol{\kappa} = \mathbf{k} - \mathbf{k'}$ that lie outside the truncation bandwidth. As mentioned, appendix C will explore these issues in greater detail.

2.3.2 Line defect: waveguide

For the line defect waveguide, we also replace the $\mathbf{G}$'s with $\mathbf{k}$'s, but we must now include the $\boldsymbol{q}$'s that correspond to the propagation direction. Our forward problem takes on the following form:

$$\left[\sum_{\boldsymbol{k'}} \eta_{\boldsymbol{k}-\boldsymbol{k'}}\big((\boldsymbol{q}+\boldsymbol{k})\cdot(\boldsymbol{q}+\boldsymbol{k'})\big)\right] h_{\boldsymbol{q}+\boldsymbol{k'}} = \frac{\omega_{\boldsymbol{q}}^2}{c^2} h_{\boldsymbol{q}+\boldsymbol{k}} \tag{2.29}$$

$$\hat{\boldsymbol{\Theta}}_{\boldsymbol{k}\boldsymbol{k'}}{}^{(q,\eta)} h_{\boldsymbol{q}+\boldsymbol{k'}} = \frac{\omega_{\boldsymbol{q}}^2}{c^2} h_{\boldsymbol{q}+\boldsymbol{k}} \tag{2.30}$$

again, omitting the band indices and making the dependence on η explicit. The dispersion relation $\omega(\boldsymbol{q})$ of the waveguide can be found by solving the forward problem for several $\boldsymbol{q}$'s along the propagation direction, and identifying the waveguiding mode of interest whose frequency lies within the bandgap.

2.4 Convergence Issues of the PWE Method

In the opening section of this chapter, we alluded to some known convergence issues with the PWE method. From working out the Helmholtz equation in the plane wave basis, the connection is clear between the plane wave method and Fourier analysis so it is not surprising that any difficulties in Fourier analysis (see appendix C) will lead to difficulties here. However, as it applies to solving the photonic bands problem, these issues were studied as early as 1992 by Sözüer, Haus and Inguva [13]. This was in an era when the field of photonic crystals was still in its infancy, and tremendous amount of effort went into accurate calculations of the band structure in search of true band gaps. In an effort to correct a common misconception, they wrote the following:

> *It is clear that, just because increasing N does not produce visible differences in the resulting band structure, one has not necessarily converged to the 'true' values. In this case, it is merely an indication of the slow*

convergence of the Fourier series.

Here, N refers to the number of terms in the summation. Unfortunately, most convergence analysis on the PWE method fails to address the problem, and even research published over a decade later [14, 15, 16] still fails to grasp the difference between convergence and accuracy. After all, what is the significance of 'rapid convergence' of the calculation if it does so to a wrong value? The problem is, of course, the slow convergence of the underlying dielectric function that the finite Fourier series is supposed to model. Additional terms in the series do not change the model enough so it only appears as though the calculation has converged.

In the literature, there was also some discrepancy as to how to treat the $\eta = \epsilon^{-1}$ term in the Helmholtz equation (eqn. (2.15)). While some have treated the Helmholtz operator as we have, others [11] expand $\epsilon_r(r)$ in the plane wave basis, and then invert the matrix instead. For an infinite Fourier series, the matrices $\eta_{k,k'}$ and $\epsilon_{k,k'}$ would be each other's inverse, but that is no longer true once we truncate. It appeared that the convergence was more rapid (i.e. used a fewer number of plane waves) using the matrix inversion treatment [17], thus justifying the computationally intensive matrix inversion, but bear in mind the *caveat* about convergence from above. Of course, in the early 1990's, there were more constraints on CPU and memory resources than there are today. This issue is addressed nicely by Li in 1996 [18], and we will review his work in section C.4. In the end, the point is moot, since Steven Johnson et al. found a way [4] to implement the plane wave method without explicitly forming either matrix, and have since made their software, the MIT Photonic Bands (MPB) package, freely available.

There are three key ideas to their approach. First, rather than solving the eigenvalue problem explicitly by diagonalization (computational work required $O(N^3)$), they use an iterative eigensolver. Secondly, they noticed that in the Helmholtz equation, the curl operator is diagonal in Fourier space, while the division by ϵ is diagonal in real space, so they make use of the Fast Fourier Transform (FFT) algorithm to transform to the appropriate basis so they do not actually store the entire $N \times N$ operator. They were thus able to reduce their storage requirement from $O(N^2)$ to $O(N)$.

This allowed them to use a much higher number of plane waves than otherwise practically achievable. However, representation of discontinuities in a Fourier basis still poses a problem. The final ingredient is an averaging technique that smoothes out the discretized elements that encloses a discontinuity. It makes use of effective medium theory and the grid elements containing the discontinuities are assigned a dielectric tensor instead of a scalar. Therefore, they have replaced the sharp scalar dielectric discontinuity with a smoothed dielectric tensor, making the Fourier representation less objectionable.

We have chosen to highlight some of the key contributions of that work here for two reasons. The first reason is that we were unable to take advantage of their ideas in the inverse problem, so we still face the same convergence issues described in [13]. It should become clear when we formulate the inverse Helmholtz problem in chapter 6 why we could not capitalize on their wisdom. The other reason is that though their work is often cited, there still appears to be confusion about the validity of the PWE method, especially the key components to making the method work. Other more recent articles on PWE [14, 16, 15, 19] either omit the reference entirely, or if it does cite it, the work demonstrates a complete lack of appreciation for the results by Johnson et al.. A recent article (2006) on these 'fast Fourier factorization methods' [20] still have not caught on to the fact that their method is at best equivalent if not inferior to the tensorial averaging in the Johnson reference, except that they do not even realize the $N \log N$ scaling, requiring $O(N^2)$ storage. For the reader interested in performing these computations, it is important to understand the significance of Johnson's work and to evaluate other PWE method research in light of their results.