Plane-wave methods for modelling photonic crystal fibre

Pearce, Gregory John

Award date:
2006

Awarding institution:
University of Bath

PLANE-WAVE METHODS FOR MODELLING PHOTONIC CRYSTAL FIBRE

Gregory John Pearce

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
Department of Physics

August 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Contents

Acknowledgements 7

Abstract 9

1 Introduction 11

1.1 Photonic crystal fibres 12

1.1.1 Guidance in PCF 12

1.1.2 Types of PCF and their applications 16

1.1.3 Computational modelling 19

1.2 Thesis outline 22

2 Maxwell's equations and their computational solution in PCF 25

2.1 Maxwell's equations in PCF 25

2.2 Computational solution of the PCF equations 27

2.2.1 Finite-difference time-domain 28

2.2.2 Frequency-domain methods 28

2.3 Comparison of computational methods 32

2.3.1 Structural generality of input 32

2.3.2 Output and its relevance to experiment 34

2.3.3 Accuracy, speed and storage efficiency 35

2.4 Summary 37

3 The fixed-frequency plane-wave method **39**

3.1 Plane wave expansion 40

3.2 Solving the reciprocal-space matrix equation 42

3.2.1 Iterative eigensolver 43

3.2.2 Iterative linear solver 46

3.2.3 Use of the fast Fourier transform 47

3.3 Issues 51

3.3.1 Dielectric discontinuities 51

3.3.2 Linear equation preconditioning 53

3.4 Summary 54

4 One-dimensional 'PCF' **57**

4.1 Motivation for the 1D model 57

4.2 The 1D model and its analytical solutions 59

4.2.1 Boundary conditions 60

4.2.2 Analytical solutions 61

4.3 The fixed-frequency plane-wave method in one dimension 63

4.3.1 Plane-wave expansion 63

4.3.2 Smoothing in the 1D method 65

4.4 Results 66

4.4.1 Systematic error introduced by smoothing 66

4.4.2 Plane-wave convergence 69

4.5 Conclusions 73

5 Adaptive curvilinear coordinates in the fixed-frequency plane-wave method 77

5.1 Introduction 78

5.2 The fixed-frequency plane-wave method in generalised curvilinear coordinates 79

5.3 Implementation 83

5.3.1 Generation of coordinate-related fixed quantities 83

5.3.2 Generation of structure-dependent fixed quantities 84

5.3.3 Matrix-vector operations by FFT 86

5.4 Adaptive coordinate generation 89

5.4.1 Fictitious energy method 90

5.4.2 Implementation of the fictitious energy method 92

5.4.3 Characterisation of adaptive grids 97

5.5 Conclusions 98

6 Applications of the plane-wave method in curvilinear coordinates 101

6.1 Application to PCF cladding structure 101

6.1.1 Adaptive grid generation 102

6.1.2 Dielectric smoothing in the GCC method 105

6.1.3 Plane-wave calculations 106

6.2 Application to supercell calculations 109

6.2.1 Adaptive grid generation 109

6.2.2 Plane-wave calculations 114
6.2.3 Computational cost 117
6.3 Conclusions 122

7 Large index contrast bandgap-guiding PCF 125

7.1 Introduction 126
7.2 Fibre design 127
7.2.1 Core wall design 129
7.3 Modelling large index contrast fibres 131
7.3.1 Restarting and preconditioning 133
7.4 Results 136
7.4.1 Numerical behaviour 136
7.4.2 Computational results 138
7.5 Conclusions 142

8 Bend loss in low contrast bandgap-guiding PCF 145

8.1 Introduction 146
8.2 Modelling an all-solid low contrast fibre 147
8.2.1 Fibre structure 148
8.2.2 Plane-wave calculations 148
8.2.3 Fundamental mode decay rate 150
8.3 Bend loss 151
8.3.1 Quantifying bend loss 153
8.4 Comparison with experiment 155

8.5 Understanding the bandstructure 158

8.6 Bend loss in air-silica PCF 161

8.7 Conclusions 162

9 Conclusions **165**

A Fields of the 1D 'PCF' **169**

A.1 General structure 169

A.2 TE modes 170

A.3 TM modes 171

B Reciprocal-space matrix equation in curvilinear coordinates **173**

B.1 Structure of the matrix equation 173

B.2 Matrix elements 175

B.2.1 Laplacian operator 175

B.2.2 Gradient operator 176

B.2.3 'Potential' operator 177

C Gradient of the fictitious energy functional **179**

C.1 Derivative of the metric tensor 179

C.2 Derivative of the elastic energy 181

C.3 Derivative of the total fictitious energy 182

D Symmetry in reciprocal space **183**

D.1 Determination of field symmetry 183

D.2 Enforcing symmetry of adaptive grids 185

E Density of states **189**

References **191**

Acknowledgements

I must first of all thank my supervisor, Professor David Bird, primarily for his guidance and assistance over the past three years, but also for providing much-needed optimism and encouragement even at times when seeing the way out of problems was difficult. This thesis has also benefited greatly from his critical eye and patient reading of early drafts.

I would also like to thank all the other PhD students and members of staff with whom I have worked, especially those whose work has motivated or supported my own work. In particular I am grateful to Dr John Roberts, Dr John Pottage and Dr Feng Luan, whose work is included in this thesis (and acknowledged separately where it appears).

Stuart Bell has generously offered his time to proof-read this thesis, and his help in improving its readability and pointing out errors is very much appreciated. Any remaining errors are, of course, entirely my own responsibility.

Going back a little further in time, I'd like to thank Dr Thompson and Mr Coope in particular for making science at secondary school fascinating and enjoyable; Mrs Whiting for persuading me to apply to Cambridge as an undergraduate; Mrs Susan Stobbs for making physics sufficiently interesting at first-year degree level for me to want to learn more; and Professor Mike Payne for giving me the opportunity to get a taste for research while still an undergraduate.

Finally, I'd like to thank my parents for being unendingly supportive and always welcoming at home, both while I've been a student and also long before. My grandparents, especially Granny Seabridge, have also helped me very much financially over the years. I am very grateful to all my family for their support.

Abstract

The work described in this thesis aims to develop and apply the fixed-frequency plane-wave method as a means of solving Maxwell's equations in photonic crystal fibre (PCF). It is first shown, by applying the method to a simple one-dimensional periodic dielectric structure, that the convergence of the method is dependent on the frequency of discrete sampling at the sharp interfaces present in the structure. This motivates the reformulation of the plane-wave method in generalised curvilinear coordinates, which provides the ability to achieve position-dependent sampling frequencies and hence enhance the convergence of the method. The improved convergence behaviour of the plane-wave method in generalised curvilinear coordinates is demonstrated by its application to realistic PCF structures.

The fixed-frequency plane-wave method is also applied to PCF structures that comprise materials with a large refractive index contrast. The difficulties associated with modelling such structures are discussed, including in particular the slow convergence of the linear solver used in eigenmode determination. An improved method of preconditioning based on an exact inverse of the linear problem corresponding to a smoothed structure is described. The fixed-frequency plane-wave method is then used to determine an appropriate hollow-core fibre structure for the guidance of light in the mid- to far-infrared wavelength region.

Finally the application of the fixed-frequency plane-wave method to fibre structures with a low index contrast is discussed, and a fibre is modelled that has been fabricated experimentally. A method to estimate the susceptibility of a fibre to bend loss is developed, and comparisons are drawn between the theoretical estimates and experimental results. The method also correctly reproduces the experimental finding that air-silica PCFs are much less susceptible to bend loss than low contrast fibres.

The conclusions drawn from this work, together with possible directions for future work, are summarised in the final chapter.

Chapter 1

Introduction

The concept of a 'photonic crystal', as opposed to crystals familiar from solid-state physics, dates back to ideas proposed by Yablonovitch [1] and John [2] in 1987. In these papers, it was suggested that photons could be controlled in a material by introducing periodic variations in its refractive index, in an approximately equivalent way to that in which the behaviour of electrons can be manipulated by the periodicity of the atomic lattice in a semiconductor crystal. In the same way that semiconductors can have electronic bandgaps (energy ranges over which no allowed electronic states exist), photonic crystals can also have 'photonic' bandgaps, frequency ranges over which light propagation through the crystal is forbidden.

Photonic crystals can be fabricated with periodicity in one (1D), two (2D) or three (3D) dimensions. In order to exhibit photonic band gaps, the length scales over which the refractive index in a photonic crystal varies must generally be of the same order as the wavelength of light ($\approx 0.5\,\mu$m for visible light) [3]. While planar 1D and 2D structures are comparatively easy to fabricate at these length scales, the manufacture of 3D structures is a particular challenge and still an area of much research [4]. Nevertheless, 1D, 2D and 3D photonic crystals have all found a wide range of applications, both in understanding the fundamental behaviour of light and in developing new optical devices [3, 5–7].

Photonic crystal fibre (PCF) is a type of photonic crystal that has periodicity in two dimensions but is uniform (and generally much greater in extent) in the

third [8, 9]. The work presented in this thesis focuses on the development of computational modelling of PCF, and therefore the remainder of this introduction concentrates on PCF rather than photonic crystals more generally. The basic concepts behind PCF and its potential applications are discussed, together with the reasons for carrying out computational modelling. The outline of the thesis is then described.

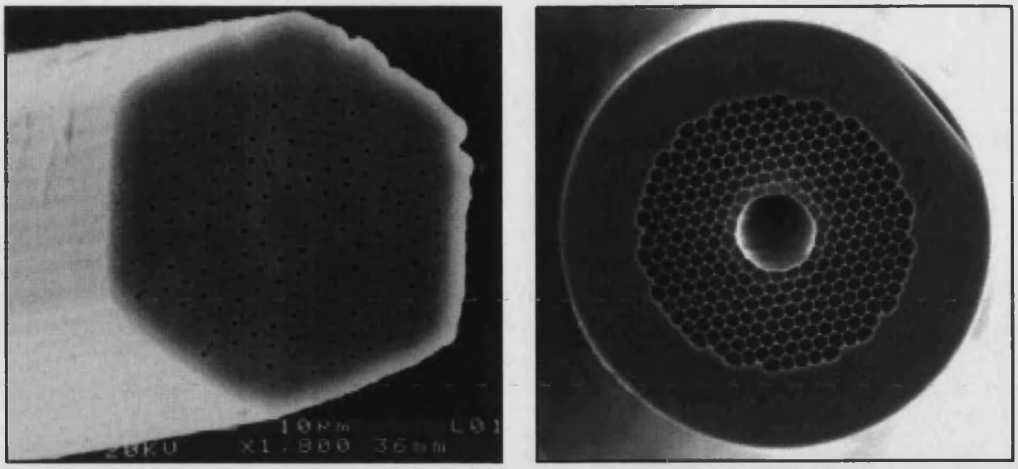
1.1 Photonic crystal fibres

'Conventional' optical fibres, in their simplest form a single strand of glass surrounded by air, have been used since the 1970s for a range of applications including telecommunications, imaging and power delivery [9–11]. Although there have been improvements in fabrication techniques used to create fibres, together with changes in the devices made from them, all such changes have been incremental [9]. There have been few changes to the basic structure of a glass 'core' surrounded by a 'cladding' of lower refractive index, which together confine light to the core by total internal reflection [11]. This is discussed below in Sec. 1.1.1.

The first photonic crystal fibres were fabricated in the 1990s [12]. Instead of being formed from a single strand of solid glass, PCFs are characterised by periodic microstructure in two dimensions perpendicular to the fibre axis. This can guide light in two very different ways. If the average refractive index of the cladding is lower than that of the core (for example, because the cladding has air holes whereas the core is solid silica), the fibre can guide by total internal reflection, the same mechanism as that in a conventional optical fibre. This is termed 'index guidance', and an example of an index-guiding fibre is shown in Fig. 1.1a. If, however, the core has a *lower* index than the average refractive index of the cladding (for example, because the core is a large air hole surrounded by a silica cladding with smaller air holes), a fundamentally different guidance mechanism—bandgap guidance—can occur. A bandgap-guiding PCF is shown in Fig. 1.1b. The two guidance mechanisms are discussed below.

1.1.1 Guidance in PCF

It is well-known that when light encounters any interface between materials the component of its wavevector parallel to the interface is conserved [11]. If the



(a) Index-guiding PCF, diameter $\approx 40\,\mu$m (b) Bandgap-guiding PCF, diameter $\approx 120\,\mu$m

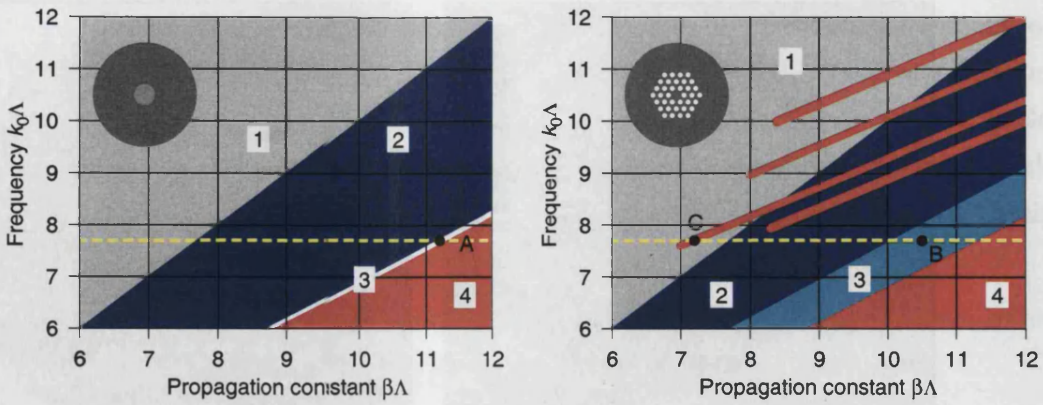
Figure 1.1: Index-guiding and bandgap-guiding PCF structures. (a) is from Ref. 13; (b) courtesy of BlazePhotonics.

structure of an optical fibre is invariant along its length (defined to be the z-axis, with unit vector $\hat{\mathbf{z}}$), all interfaces must be parallel to the z-direction and hence the component of the wavevector along $\hat{\mathbf{z}}$ is conserved. This conserved quantity is known as the 'propagation constant', β. The largest value of β that is allowed in a homogeneous medium of refractive index n is $\beta = nk_0$, where $k_0 = 2\pi/\lambda$ is the 'free-space wavevector'. At this cutoff value of β, light is propagating entirely in the z-direction; its wavevector is $\mathbf{k} = \beta\hat{\mathbf{z}}$. For a given value of β, the propagation of light is forbidden in regions with $n < \beta/k_0$ and total internal reflection prevents light from passing into such regions. As a result, light tends to be trapped in regions of higher refractive index.

A conventional optical fibre consists of a core region surrounded by cladding with a slightly *lower* refractive index, as shown in the schematic in Fig. 1.2a. There are four propagation regimes (marked 1-4) in the figure:

Region 1: In this region, $\beta < nk_0$ for refractive indices of $n = 1$ (air), $n = 1.45$ (cladding) and $n = 1.47$ (core). Light may propagate both within and outside the fibre.

Region 2: $\beta < nk_0$ only for $n = 1.45$ (cladding) and $n = 1.47$ (core). Propagation in air is forbidden (evanescent) so light is confined to the fibre, but not specifically to the core.



(a) Conventional optical fibre (b) PCF

Figure 1.2: Schematic propagation diagram for light in conventional optical fibre and PCF; axes are normalised to the pitch (spacing between PCF cladding holes) Λ. The conventional fibre has a core of germanium-doped silica ($n \approx 1.47$) and pure silica cladding ($n \approx 1.45$). The example PCF is drawn from silica and has a triangular array of air holes with an air-filling fraction of $\approx 45\%$. Points A-C on the yellow dotted lines show values of β for which light can be guided in the fibre core at a given frequency, and are described in the text. The regions 1-4 are also described in the text. Figure based on that given in Ref. 8.

Region 3: The thin region 3 has $\beta < nk_0$ only for $n = 1.47$ (core). It is in this regime at points such as A that the fibre is used to guide light; propagation is possible only in the core.

Region 4: No propagation: $\beta > nk_0$ for all n present.

It is important to note from Fig. 1.2a that β/k_0 for any mode guided in the core is always less than the core refractive index but greater than the cladding refractive index. Using index guidance it is not possible to confine light in a region of space with a lower refractive index than that of the cladding.

An example PCF is given in Fig. 1.2b, in this case consisting of cladding with a triangular array of air holes in a large silica rod, with a central 'core' created by a missing hole. The regions of interest in the diagram are:

Region 1: Light is free to propagate within and outside the fibre, as in the conventional fibre.

Region 2: Propagation is forbidden in air, since $\beta > k_0$, but allowed in core and cladding.

Region 3: $\beta < nk_0$ for $n = 1.45$ (silica) but the *average* refractive index of the microstructured cladding is reduced by the presence of the holes, and so propagation is forbidden there. Propagation is allowed only in the pure silica, i.e. either in the central defect or in the region outside the microstructured cladding.

Region 4: No propagation in this region, where $\beta > nk_0$ for silica.

Like in the conventional fibre, it is possible to guide light by total internal reflection in region 3, such as at point B. Here, β/k_0 is lower than the refractive index of the silica but the presence of the holes depresses the average refractive index of the cladding, so light is trapped in the core with β/k_0 greater than the average index of the cladding.

There is another possible mechanism for the guidance of light, unique to PCF. The periodic holes in the cladding cause the appearance of photonic bandgaps, which are regions where propagation is forbidden at certain values of β. The bandgaps are represented in Fig. 1.2b as thin 'fingers'. For certain designs of cladding, the bandgaps can extend into region 1 (i.e. they cross the 'air-line'). Points such as C can then exist, where propagation is possible in air but *not* in the cladding. In this case, if the fibre is fabricated with a hollow (air) core, light launched into the core may be confined to that region by the bandgap in the cladding[1]. This is not possible in conventional fibre, because total internal reflection can only restrict light to regions of *higher* refractive index.

The first bandgap-guiding PCF, with a solid core containing a small central air hole (such that its average refractive index was lower than that of the cladding), was reported in 1998 [14]. The first true 'hollow-core' PCF based on bandgap guidance was then fabricated in 1999 [15], opening up a new range of potential uses for PCF. These are discussed in more detail in Sec. 1.1.2.

There also exists another class of bandgap-guiding fibres, known as Bragg fibres, first proposed in the 1970s [16, 17]. These fibres have a coaxial and circularly-symmetric refractive index distribution, and their photonic crystal cladding is

[1]Note that the bandgaps considered here are 'out-of-plane' bandgaps, i.e. bandgaps forbidding propagation along the fibre axis for a given frequency. Planar 2D photonic crystals, in which only in-plane propagation is possible, also display bandgaps but generally a greater refractive index contrast is needed to create them [3].

therefore effectively one-dimensional [17]. Bragg fibres with hollow cores for use in CO_2 laser transmission have been fabricated [18], as have solid-core fibres designed to have particular properties when illuminated from outside the cladding [19]. Some of the methods discussed in Chapter 2 are applicable to Bragg fibres, but in general this thesis will consider only PCF structures with fully two-dimensional periodic claddings.

1.1.2 Types of PCF and their applications

Although it is not possible here to describe all types of PCF and their applications (and potential applications), this section aims to provide an outline of the range of PCF structures that have been developed and the uses to which they have been applied.

Endlessly single-mode and large mode-area PCF

One of the first PCF structures to be fabricated had a cladding consisting of silica containing small air holes (with a ratio of hole diameter to lattice pitch of ≈ 0.15) in a triangular lattice, with a solid core formed by the omission of one hole [12]. It was later shown that this fibre guided a single mode and did not become multi-moded as the wavelength is reduced [13]. This type of fibre with relatively small air holes and a solid core became known as 'endlessly single-mode' PCF[2]. Conventional optical fibres, conversely, are always multi-moded at sufficiently short wavelength [11].

Because the single-mode nature of an endlessly single-mode fibre depends only on the ratio of its hole diameter to pitch, endlessly single-mode fibres can be made arbitrarily large while still guiding a single mode (although, ultimately, increased susceptibility to bend loss places an upper limit on the size) [13, 20]. Such 'large mode-area' fibres are useful in delivering high power without nonlinear effects or material damage (because the area over which the power is distributed in the core can be large), and also for lasers and amplifiers [8, 20, 21].

[2]Fig. 1.1a shows an example of an endlessly single-mode fibre.

Dispersion control and non-linear effects in PCF

In contrast to large mode-area fibres, PCF structures with *small* glass cores are ideal for engineering dispersion. Group velocity dispersion (GVD), the tendency of a short pulse to lengthen during propagation, is often a major problem in telecommunications [8, 10]. By careful control of the core dimensions and the air-filling fraction of the fibre cladding, it is possible to use PCF to give precise control over the GVD. This can include shifting the zero-GVD point of silica (usually at $\lambda \approx 1.3\,\mu$m) into the visible region [22], or achieving nearly-zero dispersion over a wavelength range of hundreds of nanometres [23,24]. Control over dispersion could find uses in telecommunications [25,26], but it also has important consequences for non-linear optics because fine control over dispersion is necessary in order to achieve phase-matching (and hence provide efficiency) in non-linear processes [8].

PCFs with very small glass cores and high air-filling fractions provide the opportunity to attain very high intensities of light in glass, and this is ideal for investigating and using non-linear effects. One particularly valuable application of non-linearity in PCF is supercontinuum generation, which relies upon the combination of a range of non-linear effects to create giant spectral broadening of short, high-energy pulses [27, 28]. This has an obvious application in providing a high-power white-light source for optical characterisation, but also in more diverse applications such as optical coherence tomography [29] and spectroscopy [30]. PCF-generated optical supercontinua have also found an important use in frequency metrology. Because the supercontinuum spectrum comprises a 'comb' of precisely-separated frequency peaks, it can be used for accurate measurement of optical frequencies. Accuracies of up to a few parts in 10^{15} have been reported [31,32].

Light-atom interactions in hollow-core PCF

One long-standing difficulty in non-linear optics is how to maximise the interaction between laser light and gases, which have a low density and interact weakly with light [8]. Hollow-core PCF provides an efficient solution to this problem, because light guided in a hollow core can interact with gas atoms introduced into the core over very long interaction lengths. This has allowed the study of effects such as low-threshold stimulated Raman scattering (SRS), in which gas molecules

interact with laser light and cause a concurrent up- and down-shift in the light frequency [33]. SRS may have potential applications in laser frequency-shifting, extending the wavelengths of solid-state lasers into new frequency regions without requiring high powers to induce non-linearity [8, 33].

A phenomenon recently observed in hollow-core PCF is that of electromagnetically-induced transparency (EIT). EIT is an effect related to coherent population trapping [34], in which a narrow window of transparency can be induced in the centre of the absorption profile of a medium when two lasers are used simultaneously to excite a three-level system [34,35]. Although EIT has been observed previously in atomic vapours [36], the long interaction lengths in hollow-core PCF have allowed it to be observed in molecules (which interact much more weakly with the laser field than do atoms) [35, 37]. The narrow linewidths obtained by EIT offer possibilities in improving light storage, laser cooling and atomic clocks [35].

Particle guidance in hollow-core PCF

The use of 'optical tweezers' to manipulate small dielectric particles is now a well-known phenomenon [38], but hollow-core PCF provides an opportunity to use the same dipole forces within a fibre core to guide particles over long distances. The guided mode in the core is well-confined, and so lower laser powers can be used than would be necessary without the fibre. The guidance of 5-μm polystyrene spheres along 15 cm of PCF with a core diameter of 20 μm has been demonstrated [39], and it is hoped that the technique may eventually be applied to atoms and molecules.

Power delivery in hollow-core PCF

One obvious advantage of hollow-core PCF is that, if light is guided in air, low-loss transmission—potentially lower than in conventional optical fibres, in which light must be guided in glass rather than air—should be possible [9]. In fact, it has been shown that thermodynamic effects during the fibre drawing process cause surface roughness, and scattering from this roughness limits the reduction in loss that can be achieved using hollow-core PCF [40]. Nevertheless, hollow-core PCF does offer the possibility of guiding high-power beams in air without damage to the fibre, particularly over relatively short distances [41]. This has

many potential applications, including in laser dentistry [42], astronomy [43] and laser machining of metals [44]. It should also be possible to guide light in the mid- to far-infrared region using hollow-core PCF (see Ref. 45 and Chapter 7).

The discussion of dispersion control in PCF above referred to index-guiding PCF with solid silica cores, but it is also possible to make use of dispersion control in hollow-core PCF to enable the propagation of high-power pulses. The combined effects of non-linearity, dispersion and fibre damage make it impossible to transmit high-power pulses through conventional fibres [9]. However, this can be avoided in hollow-core PCF. The GVD of a hollow-core fibre can be engineered to be 'anomalous' (such that higher-frequency components of a pulse travel faster than lower-frequency components), and this can cancel the effect of self-phase modulation, a tendency of pulses to broaden as they propagate. As a result, it is possible for pulses of specific shapes known as solitons to propagate without spectral distortion [46, 47].

Low contrast bandgap-guiding PCF

The first bandgap-guiding PCF structures comprised silica with air holes. However, bandgap guidance does not rely on the large index contrast between air and silica, and it is therefore also possible to observe bandgap guidance in fibres that do not contain air holes. Guidance was first observed in such fibres when the air holes of an index-guiding PCF were filled with a high-index liquid, ensuring that the silica core was of lower refractive index than the average index of the cladding [48]. Later, bandgap-guiding fibres were fabricated from glass alone [49,50][3]. A more detailed discussion of low contrast bandgap-guiding fibres, including their applications, is given in Chapter 8.

1.1.3 Computational modelling

Computational modelling of optical fibres is useful for two broad reasons. Firstly, it enables fibres to be designed in advance of their manufacture, thereby providing economic savings and reducing the time spent fabricating non-optimal designs. Secondly, it enables understanding of the fundamental behaviour of light which can be more difficult to obtain experimentally. This section describes briefly

[3] A typical all-solid bandgap-guiding fibre is shown in Fig. 8.1 on page 146.

the information available from computational modelling and the way in which experimental work benefits from numerical models.

There are many ways to model conventional optical fibres computationally. Often the small refractive index contrast between the core and cladding of a fibre enables the use of approximations, but in symmetrical cases (e.g. circular fibres) it is possible to solve the fundamental equations of light propagation even without assuming low index contrast, using only one-dimensional root-finding techniques [11]. Progress in modelling elliptical fibres, structural non-uniformity and non-linear dispersion can usually be made using exact results together with perturbation theory [11]. Photonic crystal fibres, conversely, often do not have low index contrast or circular symmetry and cannot be treated as a perturbation of an exactly-soluble conventional fibre geometry. To model PCF, it is instead necessary to carry out full numerical calculations [8]. The requirements of a good computational method are discussed in detail in Chapter 2.

Fig. 1.2 shows 'propagation diagrams' illustrating the mechanism of guidance in conventional optical fibres and PCF. Fig. 1.2a refers to conventional fibres, and the information plotted in the figure requires only knowledge of the refractive indices of the materials comprising the fibre. Fig. 1.2b, however, cannot be constructed without complete knowledge of the structure of the fibre cladding. Detailed numerical modelling is the only way in which information relating to the bandgap 'fingers'—the extent to which they cross the air-line or, indeed, whether they exist at all—can be determined without fabricating the structure. Extensive modelling is therefore often used to determine optimal air-filling fractions and geometries for PCF claddings to obtain the desired bandgap properties. If a fibre is to be made to guide light at a particular range of wavelengths, the choice of cladding pitch (to ensure that a bandgap is present at the required wavelengths) is also necessary and relies on knowledge of the bandgap positions. Computer models are therefore an essential tool in designing claddings for bandgap-guiding PCF[4].

A more useful way to present the positions of bandgaps in PCF cladding structures is to note that the regions of interest in the propagation diagram are where bandgaps cross the air-line (or other low-index line, if the fibre core is not hol-

[4]As a result, the existence of photonic bandgaps in air-silica PCF was first demonstrated on the basis of numerical modelling before being verified experimentally [51].

low). Quantities such as β/k_0 ('effective index') or $(\beta - k_0)$ are therefore more useful than β, because the positions of interest can then be arranged more centrally on the plot. The photonic density of states, used in Chapters 7 and 8, and described in detail in Appendix E, is a convenient way to present this information; in addition to showing the position of bandgaps, it also provides information about the underlying bandstructure. As shown in Chapter 8, the density of states can be useful in understanding the origin of bandgaps and their relationship to fibre geometry. Although some of this information can be obtained or verified experimentally [52], it is much more easily obtained by computation.

Conventional optical fibres can, depending on geometry and refractive index, support either one ('single-mode') or more than one ('multi-mode') guided mode in the core. The same is true of PCF, and so in addition to the positions of bandgaps it is useful to know in advance how many modes are guided by the fibre[5]. It is also helpful to know the dispersion of modes (the variation of β with k_0). This information can be obtained from numerical models, and is conveniently presented as a 'mode trajectory' plot, in which the locus of individual modes in k_0-β space can be followed (see, e.g., Chapters 6 and 7).

A wide range of other fibre properties such as the fraction of modal power in air or glass, the spatial distribution of fields within the fibre, and the transmission loss can all be determined by computation. These properties are often used to influence the design of PCF; for example, to maximise its ability to deliver high-power beams, to reduce transmission losses, or to engineer the fibre dispersion [8,9].

In conclusion, much information can be gained from numerical modelling. This can assist in the designing of PCF for specific applications, but also in analysing its behaviour and making progress in developing a complete understanding of how light is guided in fibres. Numerical modelling is therefore an essential tool in all aspects of research into PCF.

[5]Information on the guided modes is relevant to index-guiding as well as to bandgap-guiding PCF.

1.2 Thesis outline

Chapter 2 sets out the basic problem to be solved, i.e. the equations (Maxwell's equations and others derived from them) governing light propagation in PCF. It then describes the requirements for a good computational method and the extent to which these requirements are satisfied by the range of methods currently in use. This is followed in Chapter 3 by an outline of the fixed-frequency plane-wave method, which is one method to solve Maxwell's equations in PCF. The remainder of this thesis aims to develop and apply the fixed-frequency plane-wave method.

Chapter 4 describes the fixed-frequency plane-wave method in one dimension as a powerful tool to investigate the fundamental behaviour of the method more generally. The results of this chapter suggest that the sampling of dielectric interfaces in a structure is particularly important when using the plane-wave method, because the convergence of solutions depends on the accuracy with which these interfaces are described.

The findings of Chapter 4 motivate the reformulation of the fixed-frequency plane-wave method in generalised curvilinear coordinates (GCCs), which is a way to obtain position-dependent sampling frequencies and therefore ensure that dielectric interfaces in a structure are properly described. The theory and implementation of this method are explained in detail in Chapter 5. The GCC method is then tested, first by application to a structure representing a PCF cladding and then to a supercell geometry in which the guided modes of an example fibre are calculated. These results are presented in Chapter 6.

Chapter 7 concerns the application of the plane-wave method to PCF structures that are suitable for the guidance of light at mid- to far-infrared wavelengths. The large refractive index contrast in these structures causes slow convergence of the linear solver used by the fixed-frequency plane-wave method, and as a result an improved method of preconditioning has been developed to overcome this problem. The plane-wave method is then used to investigate realistic designs for hollow-core PCF, and a suitable design that could be fabricated is presented.

Chapter 8 describes the modelling of an all-solid bandgap-guiding fibre and the development of methods using results from the fixed-frequency method that can

be used to estimate the susceptibility of fibres to bend loss. The thesis is concluded in Chapter 9 with a summary of its findings and suggestions for future work.

The work described in Chapters 5-8 has been published in the following papers:

- G.J. Pearce, T.D. Hedley and D.M. Bird, "Adaptive curvilinear coordinates in a plane-wave solution of Maxwell's equations in photonic crystals", Phys. Rev. B **71**, 195108 (2005).
- G.J. Pearce, J.M. Pottage, D.M. Bird, P.J. Roberts, J.C. Knight and P.St.J. Russell, "Hollow-core PCF for guidance in the mid to far infra-red", Opt. Express **13**, 6937 (2005).
- T.A. Birks, F. Luan, G.J. Pearce, A. Wang, J.C. Knight and D.M. Bird, "Bend loss in all-solid bandgap fibres", Opt. Express **14**, 5688 (2006).

In addition, the author has applied the method of Chapter 3 to a range of other systems and has contributed to the following publications:

- F. Luan, A.K. George, T.D. Hedley, G.J. Pearce, D.M. Bird, J.C. Knight and P.St.J. Russell, "All-solid photonic bandgap fiber", Opt. Lett. **29**, 2369 (2004).
- J.M. Stone, G.J. Pearce, F. Luan, T.A. Birks, J.C. Knight, A.K. George and D.M. Bird, "An improved photonic bandgap fiber based on an array of rings", Opt. Express **14**, 6291 (2006).
- J.C. Knight, F. Luan, G.J. Pearce, A. Wang, T.A. Birks and D.M. Bird, "Solid photonic bandgap fibres and applications", Jpn. J. Appl. Phys. **45**, 6059 (2006).
- T.A. Birks, G.J. Pearce and D.M. Bird, "Approximate band structure calculation for photonic bandgap fibres", Opt. Express **14**, 9483 (2006).
- A. Wang, G.J. Pearce, F. Luan, D.M. Bird, T.A. Birks and J.C. Knight, "All solid photonic bandgap fiber based on an array of oriented rectangular high index rods", Opt. Express **14**, 10844 (2006).

Chapter 2

Maxwell's equations and their computational solution in PCF

In order to consider the modelling of PCF, we first state the fundamental equations of electromagnetism (Maxwell's equations), and develop from them the governing equations that are relevant to the propagation of light in PCF. We then consider the range of methods for solving these equations that are currently in use, and discuss in general terms the requirements for a good computational method in terms of the structures it can model, the information it is able to provide, and the speed, efficiency and accuracy with which it can perform calculations.

2.1 Maxwell's equations in PCF

The fundamental equations of electromagnetism are due to Maxwell [53]. In SI units and modern vector notation, they are:

$$\nabla \cdot \mathbf{B} = 0 \tag{2.1}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{2.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \tag{2.4}$$

where $\mathbf{H}$ and $\mathbf{E}$ are the magnetic and electric fields respectively, $\mathbf{D}$ is the electric displacement field, $\mathbf{B}$ the magnetic flux density, $\mathbf{J}$ the electric current density and ρ the electric charge density. In PCF we need consider only the *source-free* Maxwell equations, which have $\mathbf{J} = \mathbf{0}$ and $\rho = 0$. We also consider fibres to be uniform along their length (conventionally defined to be the z-direction) and linear in their response to fields, which gives rise to the following constitutive relations:

$$\mathbf{D} = n^2(x,y)\varepsilon_0\mathbf{E} \tag{2.5}$$

$$\mathbf{B} = \mu_0\mathbf{H}, \tag{2.6}$$

where we introduce the refractive index $n(x,y)$, and take the magnetic permeability to be constant and equal to its free space value μ_0 everywhere, an assumption valid for non-magnetic materials. A complete solution of Maxwell's equations in the steady state is then the determination of $\mathbf{E}$ and $\mathbf{H}$ for a given refractive index (or dielectric function) profile. However, by exploiting the invariance of PCF along its length we may reduce the complexity of the problem further.

In isotropic media the solutions to Maxwell's equations take the form of plane waves. The translational invariance of PCF in the z-direction therefore implies that the z-dependence of the fields may be separated:

$$\mathbf{E}(x,y,z) = \mathbf{e}(x,y)\exp(i\beta z); \qquad \mathbf{H}(x,y,z) = \mathbf{h}(x,y)\exp(i\beta z), \tag{2.7}$$

where β is the propagation constant previously defined in Sec. 1.1.1. We may also decompose the fields into longitudinal (h^z and e^z)[1] and transverse ($\mathbf{h}_t$ and $\mathbf{e}_t$) components which are respectively parallel and perpendicular to the fibre axis, and assume an implicit time dependence[2] of the form $\exp(-i\omega t)$ to give:

$$\mathbf{E} = (\mathbf{e}_t + e^z\hat{\mathbf{z}})\exp(i(\beta z - \omega t)); \qquad \mathbf{H} = (\mathbf{h}_t + h^z\hat{\mathbf{z}})\exp(i(\beta z - \omega t)). \tag{2.8}$$

The set of equations (2.1)–(2.4) can be simplified by substituting the separated fields of Eq. (2.8) and the constitutive relations (2.5) and (2.6). If either the magnetic or electric field is eliminated, we obtain homogeneous equations for the

[1] Although non-standard, superscript indices referring to the Cartesian components of fields are introduced here to ensure consistency with Chapter 5.

[2] Note that this implicit time dependence implies the use of a frequency-domain method. Time- and frequency-domain approaches to solving Maxwell's equations are discussed in Sec. 2.2.

fields [11]:

$$(\nabla_t^2 + n^2 k_0^2 - \beta^2)\mathbf{e} = -(\nabla_t + i\beta\hat{\mathbf{z}})\mathbf{e}_t \cdot \nabla_t \ln n^2 \quad (2.9)$$

$$(\nabla_t^2 + n^2 k_0^2 - \beta^2)\mathbf{h} = \{(\nabla_t + i\beta\hat{\mathbf{z}}) \times \mathbf{h}\} \times \nabla_t \ln n^2, \quad (2.10)$$

where $k_0 = \omega/c$ is the free-space wavevector and ∇_t is the 'transverse gradient'[3]. Equating the transverse components of each side of Eqs. (2.9) and (2.10) yields two equations in the transverse fields only:

$$(\nabla_t^2 + n^2 k_0^2 - \beta^2)\mathbf{e}_t = -\nabla_t(\mathbf{e}_t \cdot \nabla_t \ln n^2) \quad (2.11)$$

$$(\nabla_t^2 + n^2 k_0^2 - \beta^2)\mathbf{h}_t = (\nabla_t \times \mathbf{h}_t) \times \nabla_t \ln n^2. \quad (2.12)$$

Solution of either one of these equations is sufficient to provide *all* field components, because the remaining components can be obtained from Maxwell's equations. For example, if $\mathbf{h}_t$ is known, then h^z, e^z and $\mathbf{e}_t$ are given by [11]:

$$h^z = \frac{i}{\beta}\nabla_t \cdot \mathbf{h}_t \quad (2.13)$$

$$e^z = i\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2} \frac{1}{k_0 n^2}\hat{\mathbf{z}} \cdot \nabla_t \times \mathbf{h}_t \quad (2.14)$$

$$\mathbf{e}_t = -\left(\frac{\mu_0}{\varepsilon_0}\right)^{1/2} \frac{1}{k_0 n^2}\hat{\mathbf{z}} \times \{\beta\mathbf{h}_t + i\nabla_t h^z\}. \quad (2.15)$$

Therefore, by solution of either Eq. (2.11) or Eq. (2.12), a complete description of the electromagnetic fields in PCF may be obtained. It is much more common computationally to solve Eq. (2.12) because, unlike the electric field, the magnetic field is a continuous quantity across all interfaces in a material.

2.2 Computational solution of the PCF equations

There is a wide range of different computational methods in current use for solving Maxwell's equations in photonic crystals. In general these fall under the broad class of either time-domain or frequency-domain methods. In time-domain methods, fields are represented on a real-space grid and, by using Maxwell's equations, are evolved in time. Frequency-domain methods instead rely on expanding

[3]The transverse gradient is the two-dimensional projection of the gradient operator, defined here by $\nabla\psi = \nabla_t\psi + i\beta\psi\hat{\mathbf{z}}$.

fields in basis states of a definite frequency and solving a resulting eigenproblem usually of the form of Eq. (2.12). The most commonly-used methods of modelling PCF are plane-wave methods (either fixed-frequency or fixed-wavevector), the multipole method and the finite element method; these are outlined here and then discussed in greater detail in the remainder of this chapter.

2.2.1 Finite-difference time-domain

The only commonly-used time-domain method for solving Maxwell's equations is finite-difference time domain (FDTD) [54, 55]. As with all finite-difference methods, fields are represented at discrete points on a real-space grid. Maxwell's equations are then written in a way such that the differential operators are replaced by approximations involving only the fields at the discrete grid points. The fields are then advanced in time, again on a discrete grid. Commonly, Yee's algorithm is used which ensures numerical stability by storing and relating $\mathbf{E}$ and $\mathbf{H}$ components in a particular way, and 'leapfrogging' the fields in time [54]. However, as with many FDTD algorithms, this places a restriction on the temporal resolution, which must be sufficiently fine at any given spatial resolution. In 2D, this implies that the total calculation time scales with spatial resolution as $O((\Delta x)^3)$ [54].

2.2.2 Frequency-domain methods

Frequency-domain methods are more common in solving Maxwell's equations in PCF, and there are many different methods available. Instead of advancing fields forward in time, frequency-domain methods assume an implicit time dependence of the form $e^{-i\omega t}$. There is then a choice of variable for which to solve. One method is to fix the wavevector of light $\mathbf{k}$ and compute a set of modes characterised by frequency ω (the 'fixed-wavevector' method). Another possibility is instead to fix ω and—for a d-dimensional system—$(d-1)$ components of $\mathbf{k}$, and compute a set of the remaining components (the 'fixed-frequency' method). Although not all computational approaches solve Eq. (2.12) directly, it is helpful to use it as an example to illustrate the difference between fixed-wavevector and fixed-frequency methods. In the fixed-wavevector approach, the propagation

constant β is fixed and Eq. (2.12) becomes:

$$\left\{\frac{1}{n^2}(\nabla_t^2 - \beta^2) + \nabla_t \frac{1}{n^2} \times \nabla_t \times\right\} \mathbf{h}_t = -\frac{\omega^2}{c^2}\mathbf{h}_t, \tag{2.16}$$

or, as it is more compactly written (re-introducing the full magnetic field vector $\mathbf{H}$),

$$\nabla \times \frac{1}{n^2}\nabla \times \mathbf{H} = \frac{\omega^2}{c^2}\mathbf{H}. \tag{2.17}$$

The allowed values of ω^2 are then calculated as eigenvalues and the magnetic field vectors $\mathbf{H}$ are eigenvectors.

In the fixed-frequency method, it is $k_0 = \omega/c$ that is fixed and instead the eigenvalues are a set of allowed β^2:

$$\left\{\nabla_t^2 + n^2 k_0^2 + \nabla_t \ln n^2 \times \nabla_t \times\right\} \mathbf{h}_t = \beta^2 \mathbf{h}_t. \tag{2.18}$$

The main mathematical difference between Eqs. (2.17) and (2.18) is that only Eq. (2.17) is Hermitian. As a result, different approaches tend to be used to solve Maxwell's equations in the fixed-wavevector and fixed-frequency formulation, with the fixed-frequency method requiring a generalised complex eigensolver rather than an eigensolver that takes advantage of Hermiticity. However, the use of a generalised eigensolver does have the benefit that, once implemented, complex dielectrics can be handled with no additional effort. It is also clear that obtaining solutions at a given frequency from Eq. (2.17) requires an iterative method because ω is a variable. The advantages and disadvantages of the two formulations are given in more detail in Sec. 2.3 later in this chapter.

Several different frequency-domain methods for solving Maxwell's equations in PCF are outlined below.

Beam propagation method

The beam propagation method (BPM) [56] is a general term used to describe methods by which the effect of an optical waveguide on a chosen input field at a specified frequency can be determined by 'propagating' the beam forwards along the axis of the waveguide. The specific details of how this is done vary, but there exist methods using finite-difference and finite-element schemes, and also

methods based on the fast Fourier transform; a discussion of these is given in Ref. 57.

A recent application of the BPM to PCF is the 'imaginary distance' BPM [57], in which the z-axis is transformed into an imaginary axis so that each mode of the fibre has a different rate of attenuation. By propagating a random input field through a fibre and examining the output at different lengths, the modes can be extracted sequentially in order of increasing propagation constant.

Finite-difference, finite-element and boundary-element methods

The frequency-domain finite-difference method [58, 59] is formulated in a similar way to FDTD (described in Sec. 2.2.1). The principal difference between the methods is that there is a time variable and temporal evolution of fields in FDTD, whereas in frequency-domain finite-difference the fields have a fixed frequency. Solution of the finite-difference equations is via a matrix eigenproblem.

The finite-element method (FEM) [60] is a powerful tool that reformulates differential equations (in this case Maxwell's equations) as a minimisation problem. In order to do this, fields are represented on a grid of nodes which define the vertices of 'elements', and via this an approximation to the total electromagnetic energy may be calculated and minimised (usually as a matrix problem). Implementations specific to PCF have been developed [61–63].

The boundary-element method (BEM) uses elements along the dielectric boundaries (interfaces), rather than discretising the entire computational domain [64, 65]. The fields are obtained via Green's second theorem, which relates fields within the boundaries to integrals over the boundary elements. Because elements are only needed at the boundaries, the BEM is easier to discretise and has more compact storage than a method in which all space is discretised [65].

Multipole method

The multipole method for solving Maxwell's equations uses e^z and h^z as variables and expands the modal fields of the structure in localised basis functions about holes/inclusions at a fixed frequency [66, 67]. In order to be efficient, this expansion is done using a basis set matched to the structure which, in the case of PCF,

is usually cylindrical harmonics—although it is possible to formulate the method using other basis sets. The field e^z inside the l-th cylinder with refractive index n_c is written in *local* polar coordinates (r_l, ϕ_l) as

$$e^z(r_l, \phi_l) = \sum_{m=-\infty}^{\infty} A_m^l J_m\left(r_l\sqrt{k_0^2 n_c^2 - \beta^2}\right) \exp(im\phi_l), \tag{2.19}$$

and that outside the cylinder in the region of refractive index n_0 as:

$$e^z(r_l, \phi_l) = \sum_{m=-\infty}^{\infty} \left[B_m^l J_m\left(r_l\sqrt{k_0^2 n_0^2 - \beta^2}\right) + C_m^l H_m^{(1)}\left(r_l\sqrt{k_0^2 n_0^2 - \beta^2}\right)\right] \exp(im\phi_l), \tag{2.20}$$

where A_m^l, B_m^l and C_m^l are sets of coefficients to be determined, and J_ν and $H_\nu^{(1)}$ are respectively the Bessel and Hankel functions of the first kind, of order ν. Similar expressions are used to describe the magnetic field h^z. In Eq. (2.20), the J_m terms represent the incident part of the field e^z for the l-th cylinder, and the $H_m^{(1)}$ terms the outgoing part associated with a source inside the cylinder. These are related by the requirement that the incident part of the field close to cylinder l must be due to the sum of the outgoing parts of the fields from all of the other cylinders.

Boundary conditions on the edges of the cylinders and the edges of the computational domain lead to a set of equations relating A_m^l, B_m^l, and C_m^l (and their magnetic field counterparts) to each other and to β. The expansions shown above are truncated in m, and the resulting finite set of equations can be cast as a matrix problem and solved for the set of allowed β.

Plane-wave methods

Plane-wave methods use expansions of fields and the dielectric function in the form of plane waves, i.e. basis functions of the form $\exp(i\mathbf{G}\cdot\mathbf{x})$ where $\mathbf{G}$ are reciprocal lattice vectors of the periodic structure under consideration. They require the structure to be perfectly periodic, but place no restrictions on the geometry of the contents of the repeating unit cell.

Unlike most other frequency-domain methods, both a formulation with fixed

wavevector [68] and with fixed frequency (described in this thesis) are in current use. In both approaches, solution is by substituting fields expressed as a superposition of plane waves into either Eq. (2.17) or Eq. (2.18) and solving a resulting matrix eigenproblem iteratively, although in each case the solution must proceed by a different method owing to the structure of the eigenproblems. This is described in more detail in Sec. 3.2.1.

2.3 Comparison of computational methods

In choosing a means to solve Maxwell's equations in PCF it is necessary to consider the requirements for a good solver and how these are met by different computational methods. The requirements can be divided roughly into three categories: the generality of possible inputs to the method; the output of the method and its relevance to experiments; and the storage efficiency (i.e. memory requirement), speed and accuracy of the method itself. In this section the requirements for a good method are set out and the available methods are compared.

2.3.1 Structural generality of input

Ideally a computational method should place as few restrictions as possible on its input, the dielectric function. The ideal method would be able to handle either periodically-repeating or finite structures; real or complex dielectric functions (so as to model intrinsic material loss if desired); dispersive dielectrics whose dielectric function depends on frequency; and should place no restrictions on the actual physical form of the refractive index, so that an arbitrary profile can be used. In practice all methods have restrictions of some sort on their input, and these limit their range of usefulness in modelling PCF in different ways.

The most general methods are FDTD, frequency-domain finite-difference and the FEM. They can be formulated with periodic boundary conditions or a finite computational area, and can generally handle any dielectric functions in any physical shape, provided the spatial (and temporal, in the case of FDTD) mesh on which calculations are performed is sufficiently fine. However, this comes with a cost: although very general, finite-difference and finite-element methods are often slow compared with basis-set methods.

A further difficulty with FDTD is that it is difficult to include material dispersion of dielectrics. Material dispersion is a variation of dielectric function with frequency [69]. It is trivial to include material dispersion in fixed-frequency methods, as the structure is simply entered with a dielectric function appropriate to the chosen fixed frequency. However, it is more difficult to include it in FDTD because there is no fixed frequency: although methods do exist to include material dispersion, they increase the complexity of calculations [54]. More generally, it is difficult to include material dispersion in *any* method that does not have a fixed frequency. In the fixed-wavevector plane-wave method, for example, solutions in dispersive media must be obtained via an iterative or perturbative approach [70].

In using basis-set methods for frequency-domain calculations (multipole and plane-wave methods being the most common), the choice of basis set itself places restrictions on the geometry of the structure under consideration. The multipole method is usually formulated using cylindrical harmonics to describe the fields (see, e.g., Eqs. (2.19) and (2.20)). It is therefore only applicable to structures comprising inclusions that can be circumscribed by non-overlapping circles in a homogeneous background material[4]. All real (fabricated) hollow-core PCF structures, and many other types of PCF, have a complex shape of core including a core wall, and most have a rounded-hexagon cladding structure which cannot be described as isolated inclusions contained within non-overlapping circles. This makes multipole methods unsuitable, in general, for modelling hollow-core PCF. However, the multipole method is a particularly efficient way to model structures with circular inclusions, and it has the further advantage that it can be formulated to handle either periodic or finite structures.

Plane-wave methods can handle arbitrary profiles of dielectric functions with any geometry—they are entirely generic except for one constraint: the dielectric function must be periodic. This implicit periodicity makes plane-wave methods ideal to study the properties of PCF cladding, the periodic structure surrounding the central region of the fibre. As discussed in Sec. 1.1.3, it is important to understand the cladding as it is fundamental in determining the positions of bandgaps and hence the guidance properties of the fibre. However, real PCFs have a different structure at their centre (such as an air core or a missing high-index inclusion), which can be thought of as a defect in the periodic cladding. To

[4]It is usually, and most efficiently, applied to purely circular inclusions, but has also been extended to treat elliptical inclusions [71].

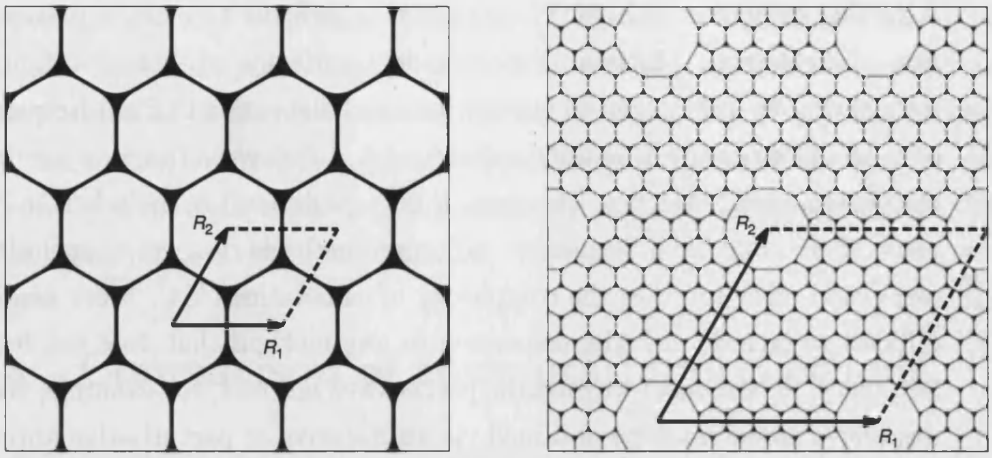


Figure 2.1: Typical geometries for a single cladding unit cell (left) and an 8×8 supercell containing repeated 7-cell 'cores' and many cladding unit cells (right). The primitive lattice vectors $\mathbf{R}_1$ and $\mathbf{R}_2$ (see description in Sec. 3.1) are shown and the dotted lines complete the computational unit cell.

model the effects of defects in the structure, it is still possible to use the plane-wave method by employing a supercell approximation. A supercell is a large 'unit cell' containing a single defect and many repeated unit cells of cladding; the exponential decay of fields in the cladding region ensures that the interaction between adjacent defect images (introduced by the periodicity of the supercell lattice) is minimal[5]. Fig. 2.1 demonstrates an example supercell geometry.

2.3.2 Output and its relevance to experiment

In the case of modelling PCF, and often more generally in the study of photonic crystals [3], it is the eigenmodes of light propagation in a given structure that are needed in order to determine its guidance properties. In an experiment, light is typically launched into a fibre at a given frequency and it is of interest which modes (each characterised by a propagation constant, β) may propagate. This makes fixed-frequency methods the most natural approach, because they are most easily compared directly to experiments.

Time-domain methods such as FDTD are well suited to dynamical problems such as the calculation of transmission losses and quality factors of resonators.

[5]The supercell approximation is also common in electronic structure calculations when a plane-wave expansion is used to describe the wavefunctions of individual molecules or surfaces. Large unit cells consisting mostly of vacuum are used to minimise the interaction between adjacent molecules or surfaces [72].

However, they are less well suited to eigenmode determination. Determining the mode frequencies of a system using FDTD typically requires launching a random input field into the simulated fibre, taking a Fourier transform of the output after evolving the system for a suitable time, and filtering out peaks that correspond to the modes [73]. Calculation of the corresponding fields is difficult and requires re-running the simulation with a narrow band filter for each mode [68].

The finite nature of PCF cladding in any real fibre has the effect of causing radiation loss, usually known as 'confinement loss' [66]. This loss arises from coupling to the continuum of radiation modes that exist outside the PCF cladding, thereby creating 'leaky' modes characterised by $\Im(\beta) \neq 0$ even in non-absorbing dielectrics. The calculation of confinement losses (as opposed to material losses, which can be calculated by any method that can handle complex dielectric functions) can be carried out only using methods with a finite real-space computational domain. It is therefore not possible to calculate confinement losses with plane-wave methods, because the implicit periodicity of a supercell calculation makes all modes lossless. The multipole method is better suited to the calculation of radiation losses [66,67], and the finite-difference, finite-element and boundary-element methods can also handle losses.

2.3.3 Accuracy, speed and storage efficiency

Finding computational solutions always involves a trade-off between accuracy and speed/memory requirements, with more accurate results invariably requiring either a longer time to calculate, a greater amount of storage, or both. All methods provide some level of control over this trade-off. Finite-difference and finite-element methods use a real-space (and temporal, in the case of FDTD) grid, and increasing the density of grid points generally leads to more accurate results. Basis-set methods such as the multipole and plane-wave methods control accuracy by the size of the basis set, which in practice must always be truncated; larger basis sets result in greater accuracy. In the case of all of these methods, increasing the spacing of grid points or the basis set size makes the method slower and require more memory.

The choice of basis in all basis-set methods has a critical effect on storage efficiency and also on calculation time. For the greatest compactness of storage (i.e. the least number of basis functions needed to obtain a sufficiently accurate

result), the basis functions should be chosen to reflect the geometry of the underlying structure. As a result, the multipole expansion is very efficient in calculating the modes of PCF structures with circular inclusions: its calculations can be carried out using a small number of basis functions, allowing it to be both fast and memory-efficient. However, this imposes unwanted symmetry restrictions. The most general choice of basis set is the plane-wave basis, as it imposes no symmetry restrictions and requires only that the dielectric structure under consideration is periodic. However, it generally does not provide a compact representation of fields or dielectric functions, particularly when the required structure has sharp discontinuities which can be described only with a very large basis set. The implications of this effect and a discussion of how it can be overcome are given in Sec. 3.3.1.

The generality of plane-wave methods comes at a cost of being memory-inefficient, but it does not necessarily make them slow. This is because it is possible to perform calculations on a set of plane-wave coefficients (reciprocal-space quantities) in real space as well as reciprocal space, with the conversion from reciprocal to real space and back carried out by the fast Fourier transform (FFT). The FFT can perform the convolution of two reciprocal-space quantities of size N in $O(N \log N)$ time rather than, as is needed in real space, $O(N^2)$; this can represent a very large saving in time when carrying out calculations. It is for this reason that, despite apparently seeming inefficient, plane-wave methods are in fact a very versatile and fast way to perform calculations.

Another important consideration when choosing a numerical method for PCF is how its efficiency is affected by whether a structure is index- or bandgap-guiding. The crucial difference between the two guidance mechanisms is that the β-values of interest for index-guiding structures are the largest in the eigenvalue spectrum, whereas those for bandgap-guiding structures lie in the interior of the spectrum. Time-domain methods are easily able to extract interior eigenvalues [68] but variational methods, often used in the fixed-wavevector formulation of the plane-wave method (both in photonics [68] and in electronic structure [72]), are unable to locate interior eigenvalues without employing transformations of the eigensystem. These tend to worsen the convergence of the method [68]. The fixed-frequency plane-wave method, however, does not employ variational methods and can easily extract interior eigenvalues iteratively (see Sec. 3.2.1).

2.4 Summary

There are many methods available to solve Maxwell's equations in PCF, and all place restrictions on the structures that can be modelled and the usefulness of the information they are able to provide. The most generic methods often come at a cost of being slow, and similarly the fastest methods impose restrictions on their input and output that make them unsuitable for many applications. In terms of being generic *and* reasonably fast and efficient, the fixed-frequency plane-wave method is ideal for many types of calculation: although it requires relatively large basis sets, the only restriction it imposes on the dielectric function is periodicity. The use of FFTs in the method enables calculations to be performed rapidly. Additionally, working at fixed frequency has several advantages: information from modelling may be readily compared to experimental results, and material dispersion and loss are included easily.

Chapter 3

The fixed-frequency plane-wave method

The overall aim of the work described in this thesis is to make improvements to the fixed-frequency plane-wave method and also to apply it to several new designs of PCF. The applications of the method include both practical designs for specific applications (e.g. PCF for guidance in the infrared, described in Chapter 7), and also more fundamental studies of the mechanism of light guidance in PCF (Chapter 8). However, in order to understand the development and applications of the method, it is first helpful to summarise the work of Hedley in developing it in its original form. The remainder of this chapter outlines the fixed-frequency plane-wave method as developed by Hedley and described in detail (including mathematical derivations) in Ref. 74.

A description is first given of the fixed-frequency plane-wave expansion and how it allows Maxwell's equations to be formulated as a matrix eigenproblem. Methods for the solution of this equation are then discussed, which include the use of an iterative eigensolver and an iterative linear solver in combination. The use of fast Fourier transforms to make the method rapid and efficient is then described. Also discussed are some of the difficulties involved in using the method, and attempts that have been made to solve them. The work described in Chapters 5-7 represents new attempts to solve these problems.

3.1 Plane wave expansion

A dielectric function that is periodic in the x-y plane with primitive lattice vectors $\mathbf{R}_1$ and $\mathbf{R}_2$ (which define any repeating unit, either a cladding unit cell or a supercell as shown in Fig. 2.1 on page 34) satisfies the condition

$$n^2(\mathbf{x}) = n^2(\mathbf{x}+\mathbf{R}), \tag{3.1}$$

where $\mathbf{R} = l_1\mathbf{R}_1 + l_2\mathbf{R}_2$ is a lattice vector ($l_1, l_2 \in \mathbb{Z}$) and $\mathbf{x} = (x, y)$ is the 2D position vector. In the case of PCF the lattice is usually hexagonal and so $|\mathbf{R}_1| = |\mathbf{R}_2|$. Throughout this thesis, *cladding* lattice pitch is always denoted Λ, giving $|\mathbf{R}_1| = |\mathbf{R}_2| = \Lambda$ for a cladding unit cell, whereas an $N_s \times N_s$ supercell satisfies $|\mathbf{R}_1| = |\mathbf{R}_2| = N_s\Lambda$. However, the method presented here can be applied to any other 2D lattice.

It is also possible to define a reciprocal lattice using the primitive reciprocal lattice vectors $\mathbf{G}_1$ and $\mathbf{G}_2$, which are related to the real-space lattice vectors by [75]:

$$\mathbf{G}_\alpha \cdot \mathbf{R}_\beta = 2\pi\delta_{\alpha\beta}, \tag{3.2}$$

where $\alpha, \beta = 1, 2$ and $\delta_{\alpha\beta}$ is the Kronecker delta. A set of reciprocal lattice vectors is then $\mathbf{G} = m_1\mathbf{G}_1 + m_2\mathbf{G}_2$ for $m_1, m_2 \in \mathbb{Z}$. Using these reciprocal lattice vectors, the dielectric function may be expanded in a plane-wave basis as

$$n^2(\mathbf{x}) = \sum_{\mathbf{G}} n^2_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{x}}, \tag{3.3}$$

where the coefficients $n^2_{\mathbf{G}}$ are defined for a unit cell of area Ω by[1]

$$n^2_{\mathbf{G}} = \frac{1}{\Omega}\int_{\substack{\text{unit}\\\text{cell}}} n^2(\mathbf{x})\, e^{-i\mathbf{G}\cdot\mathbf{x}}\, \mathrm{d}^2x. \tag{3.4}$$

All integrals in the remainder of this thesis are over the area of a unit cell unless otherwise stated.

A similar set of coefficients is also required (see Eqs. (3.7a) to (3.7d) below) to

[1] In practice, these coefficients are obtained by fast Fourier transform; the generation of plane-wave coefficients is described in detail in Sec. 3.2.3.

describe $\ln n^2$, and $[\ln n^2]_{\mathbf{G}}$ is defined through an integral in an analogous way to $n^2_{\mathbf{G}}$.

Bloch's theorem [75] states that a field resulting from a periodic 'potential', in this case the transverse magnetic field $\mathbf{h}_t = (h^x, h^y)$, can also be written as a sum over plane waves with the inclusion of a Bloch vector $\mathbf{k}$:

$$h^i(\mathbf{x}) = \sum_{\mathbf{G}} h^i_{\mathbf{k},\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}}. \tag{3.5}$$

The sets of coefficients $n^2_{\mathbf{G}}$ and $h^i_{\mathbf{k},\mathbf{G}}$ $(i \in x, y)$ represent respectively n^2 and $\mathbf{h}_t$ in reciprocal space and it is important to note that these sets of coefficients completely specify the dielectric function and transverse magnetic field. If they are required, all other fields can be obtained from Eqs. (2.13) to (2.15); other quantities such as the Poynting vector can then be calculated once all field components are known.

Substitution of Eqs. (3.3) and (3.5) into Eq. (2.18) and integrating over the area of the unit cell yields the vector wave equation in reciprocal-space form:

$$\sum_{\mathbf{G}'} \begin{pmatrix} M_{xx} & M_{yx} \\ M_{xy} & M_{yy} \end{pmatrix} \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}'} \\ h^y_{\mathbf{k},\mathbf{G}'} \end{pmatrix} = \beta^2 \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}} \\ h^y_{\mathbf{k},\mathbf{G}} \end{pmatrix}, \tag{3.6}$$

where the elements M_{ij} are:

$$M_{xx} = -|\mathbf{k}+\mathbf{G}'|^2\delta_{\mathbf{G},\mathbf{G}'} + k_0^2 n^2_{\mathbf{G}-\mathbf{G}'} + (G_y - G'_y)(k_y + G'_y)[\ln n^2]_{\mathbf{G}-\mathbf{G}'} \tag{3.7a}$$

$$M_{yy} = -|\mathbf{k}+\mathbf{G}'|^2\delta_{\mathbf{G},\mathbf{G}'} + k_0^2 n^2_{\mathbf{G}-\mathbf{G}'} + (G_x - G'_x)(k_x + G'_x)[\ln n^2]_{\mathbf{G}-\mathbf{G}'} \tag{3.7b}$$

$$M_{yx} = -(G_y - G'_y)(k_x + G'_x)[\ln n^2]_{\mathbf{G}-\mathbf{G}'} \tag{3.7c}$$

$$M_{xy} = -(G_x - G'_x)(k_y + G'_y)[\ln n^2]_{\mathbf{G}-\mathbf{G}'}. \tag{3.7d}$$

The terms of Eqs. (3.7a) to (3.7d) in $|\mathbf{k}+\mathbf{G}|^2$ arise from the ∇_t^2 operation in Eq. (2.18), and those involving $n^2_{\mathbf{G}-\mathbf{G}'}$ arise from $n^2 k_0^2$. The remaining terms are a result of the $\nabla_t \ln n^2 \times \nabla_t \times$ operation; the elements M_{xy} and M_{yx} show manifestly the way in which this term couples h^x and h^y.

It is clear from Eqs. (3.7a) to (3.7d) that the eigenproblem under consideration is not Hermitian. This is a consequence of using the fixed-frequency method, as previously discussed in Sec. 2.2.2; the fixed-wavevector formulation based on

Eq. (2.17) results instead in a Hermitian matrix problem. The non-Hermiticity of the eigenproblem and its consequences are discussed further in Sec. 3.2.1.

The reciprocal-space representation of Eq. (3.6) is exact, but it can not be solved computationally without truncating the plane-wave expansion of Eq. (3.5) and consequently making the sum over **G**-vectors finite. To do this without breaking symmetry, a circular cutoff $G_{\max}$ is introduced such that only components of h^x and h^y lying within the reciprocal-space circle defined by $|\mathbf{G}| < G_{\max}$ are included. For future reference let the number of included components of each of h^x and h^y be N_{PW}. It follows that if a vector $\mathbf{v}$ is constructed such that it contains reciprocal-space components of both h^x and h^y, the resulting eigenproblem is

$$\mathbf{M}\mathbf{v} = \beta^2\mathbf{v}, \tag{3.8}$$

where $\mathbf{M}$ is a $2N_{\mathrm{PW}} \times 2N_{\mathrm{PW}}$ matrix composed of the relevant M_{ij} components. However, as described below in Sec. 3.2, this full matrix is never constructed explicitly.

For the purpose of applying a preconditioner (described in Sec. 3.3.2), it is convenient to choose a particular ordering of the components of vector $\mathbf{v}$ and the related vectors subsequently defined in Sec. 3.2. We choose to order $\mathbf{v}$ such that $v_{2g(\mathbf{G})-1} = h^x_{\mathbf{k},\mathbf{G}}$ and $v_{2g(\mathbf{G})} = h^y_{\mathbf{k},\mathbf{G}}$, where $g(\mathbf{G})$ is an one-to-one mapping of all the **G**-vectors in ascending order of magnitude onto a set of consecutive integers beginning at 1, i.e. $g(\mathbf{G}_k) < g(\mathbf{G}_l)$ if $|\mathbf{G}_k| < |\mathbf{G}_l|$. Components with **G**-vectors of equal length appear contiguously in $\mathbf{v}$.

3.2 Solving the reciprocal-space matrix equation

For sufficiently small $G_{\max}$, the matrix $\mathbf{M}$ of Eq. (3.8) can be diagonalised directly using standard techniques [76]. Direct diagonalisation requires storage of $O(N^2_{\mathrm{PW}})$ and time $O(N^3_{\mathrm{PW}})$. Typically N_{PW} is of the order of 1,000 for single cell calculations and 100,000 for supercells. Single cell calculations are therefore feasible using direct diagonalisation, although it is not the fastest method. Supercell calculations are made impractical in terms of both storage and time: as a rough estimate, storing $\mathbf{M}$ for N_{PW}=100,000 requires $\approx$ 300 GB if double pre-

cision is used, and diagonalisation would take over a month on a typical desktop PC. However, it is possible to circumvent both the time and storage problems by using iterative methods and FFTs rather than direct diagonalisation. These methods are discussed below.

3.2.1 Iterative eigensolver

Direct diagonalisation of an $N \times N$ matrix provides N eigenvalues, some of which may be repeated. In the case of the matrix $\mathbf{M}$ of Eq. (3.8) nearly all of these eigenvalues are not needed—in fact, only a small number (often as few as 3-4, and rarely more than 50) are required, a small fraction of the total number. This suggests that a different method—one which could determine a small number of eigenvalues—would be much quicker than direct diagonalisation. If it did not require the storage of the full matrix, it would simultaneously circumvent the problem that $\mathbf{M}$ is too large to fit into memory.

In the case of hollow-core PCF, the eigenvalues of interest are those with $\beta \approx k_0$, which correspond to modes free to propagate in the hollow (air) core but prevented from propagating in the cladding by the existence of a photonic bandgap. Modes with smaller β comprise the continuum of unbound modes free to propagate in the air holes of the cladding, and larger-β modes are index-guided and are associated with the high-index regions of the cladding. The eigenvalues of interest in HC-PCF are therefore *interior* eigenvalues.

Many iterative methods exist for the determination of *extremal* eigenvalues without the need to calculate the entire eigenvalue spectrum, and these are generally very efficient [77]. They do not require storage of the complete matrix $\mathbf{M}$, and instead require only the action of $\mathbf{M}$ on arbitrary vectors. This solves the problem of being unable to store the complete matrix. However, it is computationally expensive to determine interior eigenvalues by 'working in' from one end of the spectrum: the extremal eigenvalues so obtained are not useful, and in the worst case this method requires calculating half of the total number of eigenvalues.

In order to use an iterative method to determine a few interior eigenvalues, it is necessary to use a method of shifting to create a transformed eigensystem whose extremal eigenvalues correspond to the interior eigenvalues of the system under consideration. In the remainder of this section, the way this is done in the

fixed-wavevector plane-wave method is briefly described, and then a complete description is given of the technique used in the fixed-frequency method.

Iterative eigensolver in the fixed-wavevector method

The smallest eigenvalues of a Hermitian matrix $\mathbf{A}$ (a different matrix from the non-Hermitian $\mathbf{M}$ previously considered) can be determined using the Rayleigh-Ritz algorithm [78], which states that the smallest eigenvector $\mathbf{x}_0$ and its eigenvalue λ_0 satisfy

$$\lambda_0 = \min \frac{\mathbf{x}^\dagger \mathbf{A} \mathbf{x}}{\mathbf{x}^\dagger \mathbf{x}}, \tag{3.9}$$

where the minimisation is with respect to variations in $\mathbf{x}$, and $\mathbf{x} = \mathbf{x}_0$ at the minimum. Functional minimisation (e.g., by a conjugate gradient method) yields λ_0, and sequentially larger eigenvalues can be found by maintaining orthogonality to $\mathbf{x}_0$ and previous eigenvectors. To transform the system such that the interior eigenvalues of interest become the smallest eigenvalues of $\mathbf{A}$, consider the following two equations (where $\mathbf{B}$ represents the matrix whose interior eigenvalues are desired):

$$\mathbf{B}\mathbf{x} = \lambda \mathbf{x} \tag{3.10}$$

$$(\mathbf{B} - \sigma \mathbf{I})^2 \mathbf{x} = (\lambda - \sigma)^2 \mathbf{x}. \tag{3.11}$$

The two systems shown have the same eigenvectors, but the eigenvalues are shifted: the eigenvalues of $\mathbf{B}$ closest to σ are the smallest eigenvalues of $(\mathbf{B} - \sigma \mathbf{I})^2$. Therefore if $\mathbf{A} = (\mathbf{B} - \sigma \mathbf{I})^2$ and σ is chosen to be close to the desired eigenvalues, then finding the smallest eigenvalues of $\mathbf{A}$ has the desired result of finding the interior eigenvalues of $\mathbf{B}$. The fixed-wavevector plane-wave method uses this method or variants of it [68]. However, squaring the matrix or, in practice, performing an FFT-based multiplication operation (see Sec. 3.2.3 below) twice has the undesired effect of making the method less well-conditioned, i.e. it tends to converge slowly [68].

Iterative eigensolver in the fixed-frequency method

The fixed-frequency plane-wave method does not create a Hermitian eigenproblem and a different approach is required. Instead of using a variational method to search for the smallest eigenvalues of a spectrum, it is more convenient to use the 'shift-invert' procedure to make the eigenvalues of interest the *largest* eigenval-

ues of a transformed system [74, 77]. Consider the following two eigenproblems, where now $\mathbf{M}$ is the matrix of Eq. (3.8) discussed previously:

$$\mathbf{M}\mathbf{x} = \lambda\mathbf{x} \tag{3.12}$$

$$(\mathbf{M} - \sigma\mathbf{I})^{-1}\mathbf{y} = \mu\mathbf{y}. \tag{3.13}$$

Multiplying both sides of Eq. (3.13) by $(\mathbf{M} - \sigma\mathbf{I})$ gives:

$$\mathbf{y} = \mu(\mathbf{M}\mathbf{y} - \sigma\mathbf{y}), \tag{3.14}$$

and hence

$$\mathbf{M}\mathbf{y} = \left(\frac{1}{\mu} + \sigma\right)\mathbf{y}, \tag{3.15}$$

which is manifestly of the same form as Eq. (3.12). Therefore, the eigenvectors $\mathbf{x}$ and $\mathbf{y}$ are identical and their corresponding eigenvectors are related by $\mu_i = 1/(\lambda_i - \sigma)$. Thus, by choosing σ close to the required eigenvalues ($\sigma \approx k_0^2$ in hollow-core PCF) and locating the largest eigenvalues of the transformed system $\tilde{\mathbf{M}} = (\mathbf{M} - \sigma\mathbf{I})^{-1}$, the desired interior eigenvalues of the original matrix $\mathbf{M}$ can be found. Although the shift-invert method does not require squaring $\mathbf{M}$, it does require some means of inverting $(\mathbf{M} - \sigma\mathbf{I})$, which will be discussed in Sec. 3.2.2.

A range of iterative eigensolvers have been investigated specifically for use as part of the fixed-frequency plane-wave method for PCF by Hedley [74]. The finding of this work was that the implicitly restarted Arnoldi method as implemented in the software package ARPACK [79] was the most efficient eigensolver for this application. A similar method has also been used by other authors to solve the eigenvalue equation created by the fixed-frequency finite-difference method [59].

The ARPACK eigensolver requires as input only the result of the matrix-vector operation $(\mathbf{M} - \sigma\mathbf{I})^{-1}\mathbf{u}$ for arbitrary vectors $\mathbf{u}$, which implies that—provided some means of carrying out this operation without storing $\mathbf{M}$ can be found—there is no requirement ever to create the matrix whose eigenvalues are being determined. To find the eigenvalues, ARPACK repeatedly multiplies a random starting vector by $(\mathbf{M} - \sigma\mathbf{I})^{-1}$ in order to build up a Krylov subspace from which the eigenvalues of interest are extracted. This is a fast method because, at each iteration, an optimal step towards the solution is computed taking into account all

previous search directions[2]. Typically, Hedley found that the algorithm requires only 4-10 multiplications per eigenvalue [74].

3.2.2 Iterative linear solver

The iterative eigensolver requires the result of multiplying a given vector by $(\mathbf{M} - \sigma\mathbf{I})^{-1}$. To provide this without inverting—or even storing—the matrix $\mathbf{M}$, a different iterative method can be used. The result of $(\mathbf{M} - \sigma\mathbf{I})^{-1}\mathbf{u}$ is equivalent to determining $\mathbf{w}$ in the following equation:

$$(\mathbf{M} - \sigma\mathbf{I})\mathbf{w} = \mathbf{u}, \tag{3.16}$$

where $\mathbf{u}$ is known. This is simply the solution to a set of linear equations, and many iterative solvers for this problem exist [80,81]. A range of such solvers has been tested [74] and the 'generalised minimum residual' method (GMRES) [81,82] as implemented in the PCG package [83] has been found to give particularly rapid convergence.

Like ARPACK, GMRES requires only matrix-vector multiplications, in this case of the form $(\mathbf{M} - \sigma\mathbf{I})\mathbf{y}$ for arbitrary $\mathbf{y}$, in order to determine the unknown vector $\mathbf{w}$ of Eq. (3.16). Provided an efficient preconditioner is used (see Sec. 3.3.2), a typical number of multiplications needed for each $(\mathbf{M} - \sigma\mathbf{I})^{-1}\mathbf{u}$ evaluation is 10–40. As this must be done for each matrix-vector multiplication needed by the eigensolver, it follows that of the order of 100 matrix-vector multiplications are needed for each eigenvalue of interest. However, the convergence of GMRES can be very structure-dependent, and the typical values reported by Hedley are not necessarily representative of all structures. Chapter 7 discusses a situation in which the linear solver convergence is slow, and describes an attempt to rectify this problem. Preconditioning difficulties are also encountered when using the fixed-frequency method in generalised curvilinear coordinates (see Chapters 5 and 6).

In summary, if the 'forward' operation $(\mathbf{M} - \sigma\mathbf{I})\mathbf{y}$ can be carried out on any given vector $\mathbf{y}$, we can efficiently locate an interior subset of eigenvalues of $\mathbf{M}$ without needing to store the matrix explicitly or perform a complete diagonalisation. In

[2]Only a summary is given here, but more specific details of the method can be found in Refs. 74,77,79.

the following section a fast method for carrying out $(\mathbf{M} - \sigma\mathbf{I})$ multiplications using FFTs is described.

3.2.3 Use of the fast Fourier transform

The action of $(\mathbf{M} - \sigma\mathbf{I})$ on a vector could be supplied by direct multiplication, but this requires storage of the matrix $\mathbf{M}$ and also a time of $O(N_{\mathrm{PW}}^2)$. To do this more quickly and without storing $\mathbf{M}$ it is possible to make use of the FFT, which requires instead a time of $O(N_{\mathrm{PW}} \log N_{\mathrm{PW}})$. The way in which this is done is discussed in this section.

The fast Fourier transform

The FFT is an fast implementation of the discrete Fourier transform (DFT), which creates a real-space representation of a periodic function f on a uniformly-sampled $N_1 \times N_2$ grid from a set of reciprocal-space components F and vice versa. If the real- and reciprocal-space representations are defined respectively by:

$$f(n_1, n_2) = f(\mathbf{x} = n_1\mathbf{R}_1/N_1 + n_2\mathbf{R}_2/N_2) \tag{3.17a}$$

$$F(m_1, m_2) = F(\mathbf{G} = m_1\mathbf{G}_1 + m_2\mathbf{G}_2), \tag{3.17b}$$

then the forward and inverse transforms are given by:

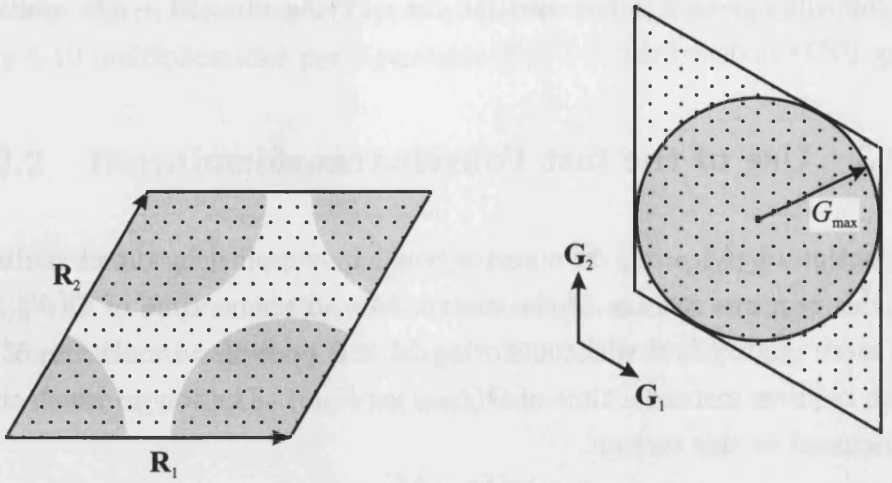
$$f(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{m_1=0}^{N_1-1} \sum_{m_2=0}^{N_2-1} F(m_1, m_2)\, e^{i\mathbf{G}\cdot\mathbf{x}} \tag{3.18}$$

$$F(m_1, m_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2)\, e^{-i\mathbf{G}\cdot\mathbf{x}}. \tag{3.19}$$

In the case of PCF, the function f may be any of n^2, $\ln n^2$, h^x or h^y.

In practice, PCFs tend to be based on hexagonal lattices and it is therefore natural to choose a grid such that $N_1 = N_2 = N$; Fig. 3.1 shows example real- and reciprocal-space grids.

As described in Sec. 3.1, a circular cutoff is used to make the total number



(a) Real-space $N = 16$ FFT grid of the form described in Eq. (3.17a). The shaded areas represent an example PCF geometry of circular air holes in glass.

(b) Corresponding reciprocal-space FFT grid, with the circular cutoff and $G_{\max}$ shown (only $\mathbf{G}$-vectors in the shaded region are included). The reciprocal lattice vectors labelled $\mathbf{G}_1$, $\mathbf{G}_2$ are shown magnified by a factor of 4 for ease of viewing.

Figure 3.1: Real- and reciprocal-space FFT grids for PCF

of reciprocal-space components of h^i finite[3]. To use the FFT it is convenient to extend the cutoff to the nearest edge of the FFT grid and set any unfilled elements (i.e., those with $|\mathbf{G}| > G_{\max}$ but included in the sum of Eq. (3.18), which are outside the shaded circle in Fig. 3.1b) to zero. The result of this choice of cutoff is that the linear FFT grid size N and the number of basis states are related by $N^2 \propto N_{\mathrm{PW}}$. As the FFT can be evaluated in $O(N^2 \log N)$ time, this implies that computation time is $O(N_{\mathrm{PW}} \log N_{\mathrm{PW}})$—a great improvement over the $O(N_{\mathrm{PW}}^2)$ of direct matrix-vector multiplication.

Implementation of matrix-vector multiplication by FFT

To demonstrate the use of FFTs in performing matrix-vector operations, consider the h^x component of the LHS of Eq. (3.6) which, written out in full, is:

$$\sum_{\mathbf{G}'}(M_1 + M_2 + M_3)h^x_{\mathbf{k},\mathbf{G}'} + M_4 h^y_{\mathbf{k},\mathbf{G}'}, \tag{3.20}$$

[3]Note that the reciprocal-space grid stored by the FFT is ordered differently for compactness, but Fig. 3.1b shows the 'actual' grid demonstrating the circular cutoff more clearly.

where

$$M_1 = -|\mathbf{k}+\mathbf{G}'|^2\delta_{\mathbf{G},\mathbf{G}'} \tag{3.21a}$$

$$M_2 = k_0^2 n^2_{\mathbf{G}-\mathbf{G}'} \tag{3.21b}$$

$$M_3 = (G_y - G'_y)(k_y + G'_y)[\ln n^2]_{\mathbf{G}-\mathbf{G}'} \tag{3.21c}$$

$$M_4 = -(G_y - G'_y)(k_x + G'_x)[\ln n^2]_{\mathbf{G}-\mathbf{G}'}. \tag{3.21d}$$

The term in M_1 can clearly be carried out in reciprocal space:

$$\sum_{\mathbf{G}'} M_1 h^x_{\mathbf{k},\mathbf{G}'} = -|\mathbf{k}+\mathbf{G}|^2 h^x_{\mathbf{k},\mathbf{G}}, \tag{3.22}$$

but the term in M_2 is best performed as follows:

$$\sum_{\mathbf{G}'} M_2 h^x_{\mathbf{k},\mathbf{G}'} = \sum_{\mathbf{G}'} k_0^2 n^2_{\mathbf{G}-\mathbf{G}'} h^x_{\mathbf{k},\mathbf{G}'} \tag{3.23a}$$

$$= k_0^2 \left[n^2(n_1,n_2) h^x_{\mathbf{k}}(n_1,n_2)\right]_{\mathbf{G}}, \tag{3.23b}$$

where Eq. (3.23b) follows from Eq. (3.23a) by the convolution theorem. The real-space dielectric function $n^2(n_1,n_2)$ is calculated once before using the eigensolver, but $h^x_{\mathbf{k}}(n_1,n_2)$ (and also $h^y_{\mathbf{k}}(n_1,n_2)$) must be created by FFT during each matrix-vector multiplication operation. The quantities labelled (n_1,n_2) in Eq. (3.23b) are multiplied together pointwise on discrete grids of the form of Eq. (3.17a) in real space.

The term in M_3 (and similarly that in M_4) is evaluated as:

$$\sum_{\mathbf{G}'} M_3 h^x_{\mathbf{k},\mathbf{G}'} = \sum_{\mathbf{G}'} (G_y - G'_y)(k_y + G'_y)[\ln n^2]_{\mathbf{G}-\mathbf{G}'} h^x_{\mathbf{k},\mathbf{G}'} \tag{3.24a}$$

$$= \left[\left\{(k_y + G'_y) h^x_{\mathbf{k},\mathbf{G}'}\right\}(n_1,n_2) \times \left\{(G_y - G'_y)[\ln n^2]_{\mathbf{G}-\mathbf{G}'}\right\}(n_1,n_2)\right]_{\mathbf{G}}, \tag{3.24b}$$

where the quantities in { } are evaluated in reciprocal space and then transformed to real space, and the quantity labelled $[\,]_{\mathbf{G}}$ is transformed to reciprocal space following the pointwise real space multiplication. As above for the M_2 term, note that the term involving $\ln n^2$ can be calculated once before using the eigensolver and stored on a real-space grid, but the term in h^x must be transformed to real space during each matrix-vector multiplication.

The total number of FFTs and inverse FFTs required to carry out a single matrix-vector operation is of interest for later comparison with the reformulation of the method in Chapter 5. These are as follows:

1. $h^x_{\mathbf{k},\mathbf{G}} \rightarrow$ real space (for $k_0^2 n^2 h^x$ multiplication)
2. $h^y_{\mathbf{k},\mathbf{G}} \rightarrow$ real space (for $k_0^2 n^2 h^y$ multiplication)
3. $\left[(k_x + G_x)h^y_{\mathbf{k},\mathbf{G}} - (k_y + G_y)h^x_{\mathbf{k},\mathbf{G}}\right] \rightarrow$ real space (for $\nabla_\mathrm{t} \ln n^2 \times \nabla_\mathrm{t} \times \mathbf{h}_\mathrm{t}$)
4. After combining terms in real space, $h^x \rightarrow$ reciprocal space
5. After combining terms in real space, $h^y \rightarrow$ reciprocal space.

Thus the matrix-vector operation corresponding to multiplication by $\mathbf{M}$ is carried out by 5 FFTs.

Determination of dielectric function plane-wave coefficients by FFT

From Eqs. (3.23a) and (3.24b) it is clear that to obtain an exact representation of the matrix-vector multiplication in Eq. (3.6) we require components of n^2 and $\ln n^2$ at $\mathbf{G}$-vector magnitudes up to a maximum of $2G_{\mathrm{max}}$. In practice, however, convergence tends to be improved by extending the circular cutoffs of both n^2 and $\mathbf{h}_\mathrm{t}$ to the edge of the same FFT grid [74].

The plane-wave coefficients $n^2_{\mathbf{G}}$ and $[\ln n^2]_{\mathbf{G}}$ can be determined from the input dielectric function by FFT, but this need be done only once before using the eigensolver because they remain fixed throughout the course of the calculation. For a calculation to be carried out ultimately with an $N \times N$ FFT grid, we sample n^2 on a *fine* real-space grid (of size $\approx 8N \times 8N$), then Fourier transform to reciprocal space. Smoothing is then applied by multiplication of the Fourier components by a Gaussian envelope, for the reasons discussed below in Sec. 3.3.1. The resulting smoothed components are transformed back to real space and the logarithm is taken to create the function $\ln n^2$ on the fine real-space grid, which is then transformed to reciprocal space. Components of both n^2 and $\ln n^2$ satisfying $|\mathbf{G}| < G_{\mathrm{max}}$ (where G_{max} corresponds to the smaller $N \times N$ FFT grid) are extracted for use in the mode-solving calculation. By using this oversampling procedure, the components of n^2 and $\ln n^2$ remain consistent, i.e. they describe the same dielectric function in each case.

The quantities appearing in Eqs. (3.23b) and (3.24b) are the real-space representations of the dielectric function and the gradient function $\nabla_t \ln n^2$. As described above, these are calculated once *from the plane-wave coefficients* ($n^2_{\mathbf{G}}$ and $[\ln n^2]_{\mathbf{G}}$). The functions must be generated on real-space grids from the truncated set of plane-wave coefficients rather than directly from the input dielectric function in order to maintain the equivalence with matrix-vector multiplication of the form of Eq. (3.8).

3.3 Issues

Although the fixed-frequency plane-wave method has many features that make it ideally suited to studying the propagation of light in PCF, it also suffers from a number of drawbacks. These problems, together with some means to overcome them considered by Hedley [74], are discussed below. Other ways to solve them are the subject of Chapters 5 and 7 of this thesis.

3.3.1 Dielectric discontinuities

The microstructure of PCF is such that, in most cases, its dielectric function is piecewise constant but discontinuous at the boundaries between materials, e.g. between air and glass in a typical hollow-core fibre. The consequences of this discontinuity for the fixed-frequency plane-wave method are discussed in this section.

Maxwell's equations require that $\mathbf{H}$, $\hat{\mathbf{n}} \times \mathbf{E}$ and $\hat{\mathbf{n}} \cdot \varepsilon\mathbf{E}$ are conserved across any boundary with normal vector $\hat{\mathbf{n}}$ [69]. In PCF, where translational invariance requires that there is no z-component to $\hat{\mathbf{n}}$, continuity of $\hat{\mathbf{n}} \times \mathbf{E}$ implies that e^z is also continuous. It then follows from Eq. (2.14) that $\varepsilon^{-1}\hat{\mathbf{z}} \cdot \nabla_t \times \mathbf{h}_t$ is also conserved across any interface, and this in turn implies that any discontinuity in the dielectric function is accompanied by a discontinuity in the gradient of $\mathbf{h}_t$.

The plane-wave expansions of Eqs. (3.3) and (3.5) provide an exact representation[4] of any function in the limit of an infinite number of plane-wave coefficients, as the Fourier basis set is complete [78]. However, the Fourier representation of

[4]We define 'exact' to mean that the function $f_N(\mathbf{x})$ expanded in N coefficients and the true function $f(\mathbf{x})$ satisfy $\int |f_N(\mathbf{x}) - f(\mathbf{x})|^2 \, \mathrm{d}^2x \to 0$ in the limit $N \to \infty$.

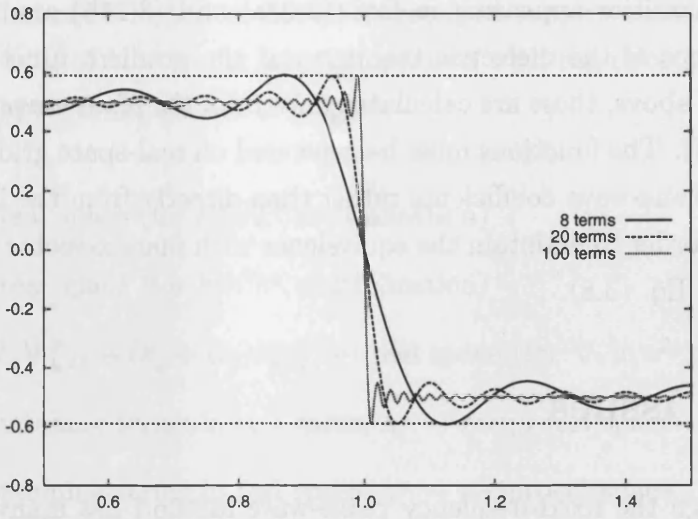


Figure 3.2: Convergence of the Fourier expansion for a periodic unit step function (i.e. a square wave). Half of the period of the function is shown. The horizontal dashed lines show the expected overshoot of ≈ 0.0895. Note that addition of terms in the Fourier expansion moves the location of the overshoot towards the discontinuity but does not reduce its magnitude.

a function that is discontinuous (such as n^2 or $\nabla_t h^i$) suffers the appearance of Gibbs' phenomenon, which is an 'overshoot' as the discontinuity is approached. The location of the overshoot maximum/minimum relative to the discontinuity is inversely proportional to the number of terms in a truncated Fourier series, but the size of the overshoot does not decrease as the number of terms increases; it tends to a finite value of $\approx 0.0895D$ for a discontinuity of magnitude D [78,84], as illustrated in Fig. 3.2. This is a well-known cause of difficulties in the convergence and accuracy of numerical methods that use Fourier expansions [68,74,78,84,85].

The difficulties associated with discontinuous dielectric functions can be prevented by artificially smoothing the interfaces between dielectrics. A convenient way to do this is to use Gaussian smoothing, convolving the dielectric function with a Gaussian of chosen width (the 'smoothing width'). Other methods are in use such as dielectric averaging [68, 86]. When using Gaussian smoothing, the choice of smoothing width is crucial. A large smoothing width gives results that are perturbed from the correct results, but allows faster convergence with respect to the plane-wave cutoff. Narrower smoothing widths give greater accuracy but at the expense of the size of basis needed for convergence. This subject is discussed and investigated in greater detail for a 1D model in Chapter 4, and leads to the work presented in Chapter 5.

3.3.2 Linear equation preconditioning

The linear equation (3.16) may be solved as described in Sec. 3.2.2 by successive application of $(\mathbf{M} - \sigma\mathbf{I})$ to vectors provided by GMRES, but it is generally the case that convergence when using this method is either slow or fails entirely. This failure of an iterative numerical method is not uncommon: it is also found in variational schemes for the fixed-wavevector plane-wave method [68] and the plane-wave method in condensed matter [72]. It is common in such cases to use *preconditioning* [80]. The simplest method is to multiply by a 'left preconditioner' $\mathbf{P}$:

$$\mathbf{P}(\mathbf{M} - \sigma\mathbf{I})\mathbf{w} = \mathbf{P}\mathbf{u}, \tag{3.25}$$

where $\mathbf{P}$ is chosen to be an approximate inverse of $(\mathbf{M} - \sigma\mathbf{I})$. For each iteration, preconditioned GMRES may supply an arbitrary vector $\mathbf{y}$ and require either the result of a matrix-vector operation $(\mathbf{M} - \sigma\mathbf{I})\mathbf{y}$, or the result of a preconditioning operation $\mathbf{P}\mathbf{y}$.

Even with preconditioning, however, the convergence of iterative methods tends to be highly sensitive to the exact choice of preconditioner: often preconditioning schemes are not easily transferable either between numerical methods or even between specific applications of a single method [87].

A good preconditioner for the fixed-frequency plane-wave method has been found to be a combination of an exact inverse (which can be determined by LU decomposition) of the sub-matrix of $(\mathbf{M} - \sigma\mathbf{I})$ defined by $|\mathbf{G}|, |\mathbf{G}'| < G^{\mathrm{P}}_{\mathrm{max}}$, and the Jacobi preconditioner [80] given by $P_{ij} = \delta_{ij}/(M_{ii} - \sigma)$ for larger $\mathbf{G}$-vectors [74]. Unlike the full matrix $\mathbf{M}$, the preconditioning sub-matrix is sufficiently small to be formed and stored explicitly in memory. The choice of $G^{\mathrm{P}}_{\mathrm{max}}$ determines the size N_{P} of the exactly-inverted square, and in general it is much smaller than G_{max} (typically resulting in $N_{\mathrm{P}} \lesssim 0.1 N_{\mathrm{PW}}$ [74]).

The ordering of the elements of $\mathbf{y}$ (the same as that described for $\mathbf{v}$ in Sec. 3.1) allows the preconditioner to be applied easily. The elements of $\mathbf{M}$ satisfying $|\mathbf{G}|, |\mathbf{G}'| < G^{\mathrm{P}}_{\mathrm{max}}$ are contiguous and simply make up a square sub-matrix which, if $\mathbf{M}$ were ever formed, would be in its upper left corner. This sub-matrix is created explicitly and inverted, and the preconditioner can then be applied to elements y_k with $k \leq N_{\mathrm{P}}$ by direct matrix-vector multiplication. The remaining

elements of $\mathbf{y}$ with $k > N_\mathrm{P}$ are preconditioned with the Jacobi preconditioner, which simply involves multiplying each y_k by $(M_{kk} - \sigma)^{-1}$.

The preconditioner cutoff $G^\mathrm{P}_\mathrm{max}$ gives control over the trade-off between the quality of the preconditioner (which determines the number of GMRES iterations needed) and the time taken to create and apply it. The total time taken to perform the matrix-vector multiplications needed by GMRES is proportional to the number of iterations needed[5]. However, there is no simple relationship between N_P and the number of iterations other than that, in general, increasing N_P usually gives an improvement.

The drawbacks to using a large preconditioning sub-matrix are the time taken to perform the exact inversion to create the preconditioner, which is $O(N_\mathrm{P}^3)$ (although this need be done only once for any given structure and frequency), and the time taken to apply the preconditioner by matrix-vector multiplication, which is $O(N_\mathrm{P}^2)$. For very large preconditioners, storage can also become a problem. These difficulties have motivated the study of a new type of preconditioner, which is described in Chapter 7.

In some cases the preconditioner described above is inadequate, and the result is slow or 'stagnated' (failed) convergence of GMRES. This difficulty has been encountered in high-contrast PCF structures such as those for guiding far infrared light (as discussed in Chapter 7), and also in using the fixed-frequency plane-wave method in curvilinear coordinates (discussed in Chapter 5). Further discussion of preconditioning and methods to overcome the shortcomings of existing preconditioners are given in those chapters.

3.4 Summary

By making use of two nested iterative methods—an eigensolver and a linear solver—together with the fast Fourier transform, the fixed-frequency plane-wave method provides a fast method to solve Maxwell's equations that is ideally suited to the calculation of PCF bandstructures. However, it suffers from two main disadvantages. Firstly, the sharp interfaces in PCF cause difficulties for the plane-

[5]There is also a GMRES overhead, which is time taken by the GMRES algorithm itself. It is not usually significant, but can become so if the preconditioner is ineffective and the number of GMRES iterations becomes large. The overhead is discussed further in Chapter 7.

wave expansion and this makes it necessary to use smoothing, which perturbs the results. Secondly, the linear solver converges slowly unless it is preconditioned. Although previous attempts have been made to design a good preconditioner, this preconditioner is not always effective.

The work presented in this thesis describes work to improve the fixed-frequency plane-wave method by overcoming these disadvantages. Chapters 5 and 6 describe and analyse a method that improves the convergence behaviour of the plane-wave method when sharp dielectric interfaces are used, by allowing a plane-wave cutoff that can vary in space. In Chapter 7, an improved method of preconditioning is presented, which has enabled the fixed-frequency plane-wave method to be used to investigate dielectric structures with large refractive index contrast.

Chapter 4

One-dimensional 'PCF'

In this chapter, a one-dimensional (1D) model of 'PCF' is developed and then investigated both analytically and with the fixed-frequency plane-wave method. This has two benefits: firstly, because a given spatial resolution can be achieved in 1D with much less computational effort than in 2D, it allows the plane-wave method to be studied in much greater detail than in 2D. Secondly, it allows the accuracy of the method and the effect of smoothing to be tested rigorously against analytical solutions. Valid and useful comparisons can then be made with the corresponding 2D calculations because the 1D model does not differ in complexity from the equivalent in 2D: the same convergence difficulties caused in 2D by sharp interfaces are reproduced by the 1D model.

There follows a discussion of the motivation for studying a 1D model, and a description of the model itself and its analytical solutions. The use of the fixed-frequency plane-wave method to solve the 1D model is then described. Comparisons are made between the plane-wave and analytical solutions in order to analyse the effect of smoothing and the convergence behaviour of the plane-wave method.

4.1 Motivation for the 1D model

In its usual 2D formulation described in Chapter 3, the fixed-frequency plane-wave method for PCF stores dielectric functions and the magnetic field on FFT

grids of dimension $N \times N$. Consequently, the time taken for Fourier transforms is $O(N^2 \log N)$ and the storage required for the FFT grids (or, equivalently, the set of plane-wave coefficients) is $O(N^2)$. The corresponding real-space sampling frequency of the dielectric function in any given direction, which is related to the accuracy of the plane-wave description of the function (and discussed more quantitatively below in Sec. 4.4.2), is proportional to N. In 1D, however, a sampling frequency of N has storage requirement $O(N)$ and the FFT requires a time of $O(N \log N)$.

The greatly reduced storage and time requirements of the plane-wave method in 1D make it an ideal way of studying the fundamental convergence properties of the method, because much finer spatial resolutions can be obtained with a given amount of memory and time than in 2D. This makes the 1D model ideal to study air/glass systems with a fine spatial resolution that would not be possible in 2D, but it can also be used to investigate structures containing metals. Metals have a dielectric function with a negative real part and non-zero imaginary part (e.g. gold at visible wavelength $\lambda \approx 0.6\,\mu$m has $n^2 \approx -10 + i$ [88]). As a result of the large refractive index step between air and metal, and the consequential change of sign of magnetic field gradient (see below in Sec. 4.2.1), it is likely that metals prove to be an especially difficult case for the fixed-frequency plane-wave method. Although the study of metals is not an overall aim of this thesis, it is interesting, having developed a 1D plane-wave code, to use the code to investigate briefly whether carrying out calculations involving metals is realistic in 2D.

A further advantage of investigating 1D 'PCF' is that the simple model can be solved analytically, providing exact results which can be compared to those of the plane-wave method. If exact results are known for infinitely-sharp interfaces, it is possible to quantify the systematic error introduced by smoothing the dielectric interfaces when using the plane-wave method (as described in Sec. 3.3.1). The 'smoothing error' is important because the level of smoothing must always be chosen such that its associated error is within tolerable limits. The 1D plane-wave method can also be used to investigate a second issue: the size of basis set needed to obtain converged results at a given level of smoothing. This sets an ultimate limit on which calculations are feasible in a given amount of time, or with a given amount of memory.

As will be shown in Sec. 4.2, the basic structure of the governing equations—

including the 'vector term'—are present in 1D as well as in 2D, and therefore it is possible to make useful comparisons between the behaviour of the 1D model and its 2D counterpart.

4.2 The 1D model and its analytical solutions

The 1D model considered here is an array of thin dielectric slabs in air, periodic in the x-direction and extending infinitely in the y- and z-directions, with a dielectric function given by

$$n^2(x) = \begin{cases} n^2_{\text{dielec}} & -d/2 < x \leq d/2 \\ n^2_{\text{air}} & \text{elsewhere} \end{cases}, \tag{4.1}$$

with periodicity such that $n^2(x+\Lambda) = n^2(x)$. This gives a unit cell of length Λ containing a dielectric slab of width d. Fig. 4.1 on page 62 (and later figures) show the structure of the system.

When n^2 is a function of x only, the real-space vector wave equation (2.12) separates into two decoupled equations:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}x^2} + n^2k_0^2\right\} h^x = \beta^2 h^x \tag{4.2}$$

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}x^2} + n^2k_0^2 - \frac{\mathrm{d}\ln n^2}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}x}\right\} h^y = \beta^2 h^y. \tag{4.3}$$

In this particular case—where the h^x and h^y components are decoupled—we are free to choose either $h^y = 0$ or $h^x = 0$ to form two linearly-independent sets of modes. Note that this is not possible in general for PCF, which has 2D structure and is infinite in the third dimension; it is, however, possible for 2D planar photonic crystals when considering in-plane propagation [3].

The mode with $h^y = 0$ must have $e^z = 0$ (from Eq. (2.14)) and therefore $\mathbf{E}$ is normal to the direction of propagation. This is denoted a 'TE' (transverse electric) mode, and is described by Eq. (4.2). Similarly the mode with $h^x = 0$ has $h^z = 0$ (from Eq. (2.13)) and $\mathbf{H}$ is normal to the direction of propagation, a 'TM' (transverse magnetic) mode, described by Eq. (4.3).

It is interesting to note that Eq. (4.2) is exactly equivalent to the time-indepen-

dent Schrödinger equation in 1D (taking $\hbar^2/2m = 1$):

$$\left\{-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right\}\psi = E\psi, \tag{4.4}$$

if the identifications $-k_0^2 n^2(x) \equiv V(x)$ and $-\beta^2 \equiv E$ are made with the potential function and the total energy, and the wavefunction ψ is considered to be analogous to h^x. In fact, the system described is well-known in quantum mechanics as the Kronig-Penney model of electrons in a periodic potential [89].

The TM mode equation (4.3) shares more characteristics with the 2D vector wave equation than does the TE mode equation, as it has an additional term $\frac{\mathrm{d}}{\mathrm{d}x}\ln n^2 \cdot \frac{\mathrm{d}}{\mathrm{d}x}h^y$ corresponding to $\nabla_\mathrm{t} \ln n^2 \times \nabla_\mathrm{t} \times \mathbf{h}_\mathrm{t}$ in 2D. As a result, the boundary conditions on the fields for 1D TM polarisation are also more representative of those on the full solutions in 2D. The TM boundary conditions also make the TM modes more difficult to converge when using the plane-wave method; if a particular size of basis set gives converged TM results, the TE results are always also converged. For these reasons it is more useful to consider the TM modes in 1D. The numerical results presented in this chapter will concentrate on TM modes.

4.2.1 Boundary conditions

Maxwell's equations require that $\mathbf{H}$ is continuous across all interfaces and so TE and TM modes must both satisfy continuity of h^x and h^y at the edges of the dielectric slab. The conditions on the x-derivative of the field, however, are different for the two polarisations.

For the TE mode, integrating Eq. (4.2) once yields:

$$\frac{\mathrm{d}h^x}{\mathrm{d}x} + k_0^2 \int n^2 h^x \,\mathrm{d}x = \beta^2 \int h^x \,\mathrm{d}x. \tag{4.5}$$

Given that the integral of a function containing a finite number of discontinuities is itself continuous, Eq. (4.5) can only be satisfied if $\frac{\mathrm{d}h^x}{\mathrm{d}x} \equiv (h^x)'$ is also continuous. Continuity of h^x and $(h^x)'$ is analogous to the boundary conditions on ψ in quantum mechanics, where ψ and ψ' are continuous.

To determine the boundary condition on $(h^y)'$ for the TM mode, we rearrange

Eq. (4.3) and divide by n^2 to give:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{n^2}\frac{\mathrm{d}h^y}{\mathrm{d}x}\right)+\frac{n^2k_0^2-\beta^2}{n^2}h^y=0, \tag{4.6}$$

then integrate:

$$\frac{1}{n^2}\frac{\mathrm{d}h^y}{\mathrm{d}x}+\int\frac{n^2k_0^2-\beta^2}{n^2}h^y\,\mathrm{d}x=0. \tag{4.7}$$

In this case again the integrated quantity must be continuous, and it is therefore $(h^y)'/n^2$ that is conserved at the interfaces. (Note that this result may also be derived from the more general 2D boundary condition, that $\varepsilon^{-1}\hat{\mathbf{z}}\cdot\nabla_\mathrm{t}\times\mathbf{h}_\mathrm{t}$ is conserved.) For a metal/air interface, where the real part of n^2 changes sign, this implies that there must be also be a change of sign of the gradient of the magnetic field.

4.2.2 Analytical solutions

In any region of *constant* n^2, the Helmholtz equation $h''+(n^2k_0^2-\beta^2)h=0$ must be solved; the difference between TE and TM modes is evident here only in the boundary conditions at interfaces. Analytic solution can be made easily at the Brillouin zone centre (Bloch wavevector $k\Lambda=0$; no phase change between unit cells) and at the Brillouin zone boundary ($k\Lambda=\pi$; phase change of π between unit cells)[1]. In either case the modes are of definite symmetry, either even or odd about $x=0$; this allows the equation to be solved by piecewise sine and cosine (or exponential) functions in the dielectric and air regions. Numerical solution is needed to determine the values of β that satisfy the boundary conditions. (For TE modes, this is exactly equivalent to the method used to solve 'finite potential well' problems in quantum mechanics.) Full details of the field construction and conditions on β are given in Appendix A.

Modes can be described as *propagating* if $\Re(\beta\Lambda)>0$ because in some (or all) regions of space they are not evanescent. It is useful to categorise them further as *guided* if $\Re(\beta\Lambda)>n_\mathrm{air}k_0\Lambda$, because they are evanescent in air and hence localised to the dielectric. A simple example of the mode structure of propagating modes, calculated as described above, is provided by glass slabs of $n=1.5$ (approximately

[1]It is also possible to solve the model analytically for arbitrary Bloch wavevector, but this is more complex as it requires recourse to transfer matrix methods [89]. Solutions at the centre and edge of the Brillouin zone are sufficient here.

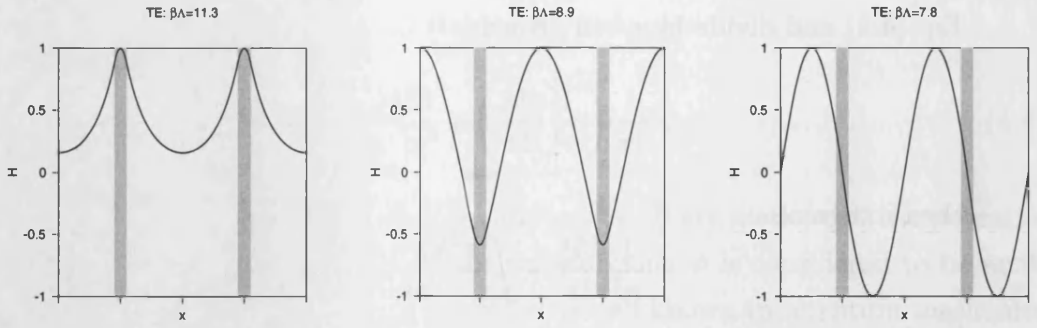


Figure 4.1: The three propagating TE modes of the glass slab waveguide with $n^2_{\text{dielec}} = 2.25$, $n^2_{\text{air}} = 1$, $k_0\Lambda = 10$, $d = 0.1\Lambda$ and at the Brillouin zone centre. The leftmost (highest β) mode has $\beta\Lambda > n_{\text{air}}k_0\Lambda$ and is guided; it is evanescent in air but not in glass. The other two modes propagate both in air and glass. There are no propagating modes at lower β.

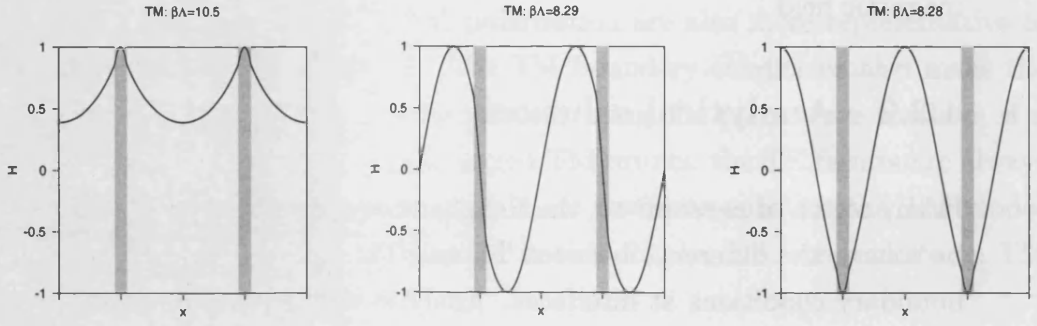


Figure 4.2: The three propagating TM modes of the glass slab waveguide described in Fig. 4.1. The leftmost (highest β) has $\beta\Lambda > n_{\text{air}}k_0\Lambda$ and is guided. There are no propagating modes at lower β. Note the discontinuous field gradient at the interfaces distinguishing these modes from those in Fig. 4.1.

the refractive index of silica) and shown in Figs. 4.1 and 4.2. We choose $k = 0$ (the Brillouin zone centre) for simplicity.

A more 'extreme' case is provided by metallic slabs in air, which can be solved exactly as described above. The general structure of modes of this waveguide is demonstrated by the system with $n^2_{\text{dielec}} = -10 + i$, $n^2_{\text{air}} = 1$, $d = 0.1\Lambda$, again with $k = 0$ for simplicity. There exist TE and TM modes (Figs. 4.3 and 4.4) with $\Re(\beta\Lambda) < n_{\text{air}}k_0\Lambda$ and also a set of two TM 'surface' modes (Fig. 4.5) with $\Re(\beta\Lambda) > n_{\text{air}}k_0\Lambda$, which must be evanescent in *both* the metal and air and are therefore bound to the metal-air interface. The existence of these modes is a direct consequence of the boundary condition on $(h^y)'$ in the TM case: the change of sign of dielectric function causes a change of sign of $(h^y)'$, thereby allowing modes

that decay exponentially on both edges of the interface. There are no TE surface modes.

4.3 The fixed-frequency plane-wave method in one dimension

This section outlines how the fixed-frequency plane-wave method can be used to solve the 1D model previously described, and begins with a brief description of the method in 1D. Only a summary is given, as the 1D method shares many characteristics with its 2D counterpart already outlined in Chapter 3: like the 2D method, it uses an iterative eigensolver, a preconditioned iterative linear solver and fast Fourier transforms to find solutions efficiently.

4.3.1 Plane-wave expansion

A dielectric function in 1D that is periodic in the x-direction with a lattice constant (or unit cell length) R_0 satisfies

$$n^2(x) = n^2(x + R), \tag{4.8}$$

where $R = lR_0$ is a lattice 'vector' and $l \in \mathbb{Z}$. As in 2D, it is possible to define a primitive reciprocal lattice vector G_0 by the relationship $G_0 R_0 = 2\pi$, and this enables the construction of a set of reciprocal lattice vectors $G = mG_0$ for $m \in \mathbb{Z}$. Expansion of the dielectric function and magnetic fields in a plane-wave basis then takes the form:

$$h^i(x) = \sum_G h^i_{k,G} e^{i(k+G)x} \tag{4.9}$$

$$n^2(x) = \sum_G n^2_G e^{iGx}, \tag{4.10}$$

where k is the 1D Bloch wavevector. The plane-wave coefficients n^2_G (and h^i_G in an analogous way) are defined by:

$$n^2_G = \frac{1}{R_0} \int_{\substack{\text{unit} \\ \text{cell}}} n^2(x) e^{-iGx} \, \mathrm{d}x. \tag{4.11}$$

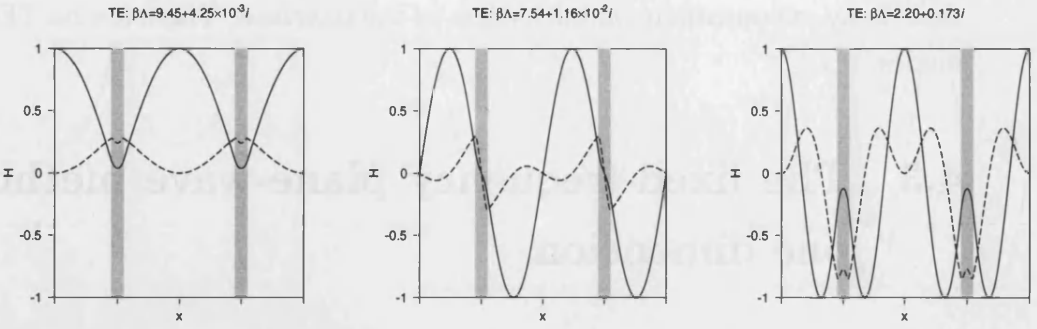


Figure 4.3: The three propagating TE modes of the slab waveguide with $n^2_{\text{dielec}} = -10+i$, $n^2_{\text{air}} = 1$, $d = 0.1\Lambda$, $k_0\Lambda = 10$ and at the Brillouin zone centre. The real (solid line) and imaginary (dashed line) parts of the field are shown with the imaginary part multiplied by 50 for clarity. None of these modes is bound to the metal, since for all three $\Re(\beta\Lambda) < n_{\text{air}} k_0 \Lambda$.

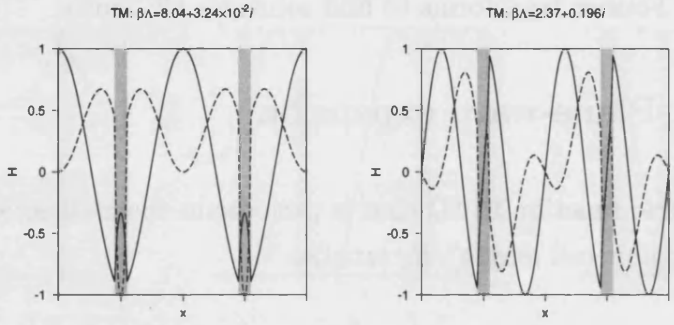


Figure 4.4: The two propagating TM modes of the slab waveguide described in Fig. 4.3. The real (solid line) and imaginary (dashed line) parts of the field are shown with the imaginary part multiplied by 50 for clarity. As in the TE case, neither is bound to the metal.

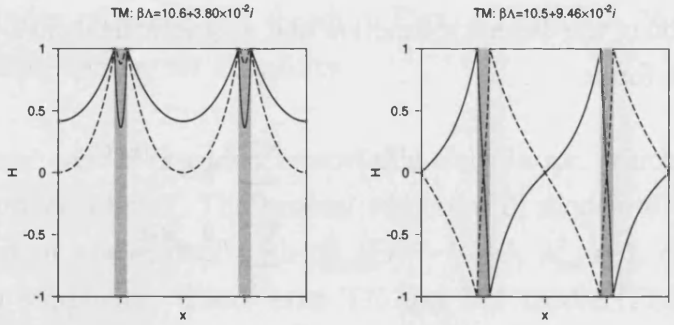


Figure 4.5: The two surface TM modes of the slab waveguide described in Fig. 4.3. The real (solid line) and imaginary (dashed line) parts of the field are shown scaled separately for clarity: the real part of the field of the even mode is greater than the imaginary part by a factor of ≈ 20, and the real part of the field of the odd mode is smaller than the imaginary part by a factor of ≈ 6. These modes have $\beta\Lambda > n_{\text{air}} k_0 \Lambda$ and are therefore bound to the metal. Note the discontinuous field gradient at the metal-air interfaces, both here and in Fig. 4.4.

The reciprocal-space representations of Eqs. (4.2) and (4.3) using these expansions are then given by:

$$\sum_{G'} \left\{ -(k+G')^2 \delta_{G,G'} + k_0^2 n^2_{G-G'} \right\} h^x_{k,G'} = \beta^2 h^x_{k,G} \tag{4.12}$$

$$\sum_{G'} \left\{ -(k+G')^2 \delta_{G,G'} + k_0^2 n^2_{G-G'} + (G-G')(k+G')[\ln n^2]_{G-G'} \right\} h^y_{k,G'} = \beta^2 h^y_{k,G}. \tag{4.13}$$

The solution of either one of these 1D equations can then be carried out in the same way as that of the 2D equations described in Chapter 3. However, as noted in Sec. 4.1, large FFT grids can be used, which correspond to much finer spatial resolutions than in 2D.

4.3.2 Smoothing in the 1D method

None of the discussion of the implementation of the fixed-frequency plane-wave method in 2D, given in Sec. 3.2, is substantially different in 1D. Similarly, the same issues as those discussed in Sec. 3.3 are present, although owing to the reduced storage requirements in 1D it is generally possible to overcome any preconditioning problems by using a sufficiently large exact preconditioner sub-matrix. However, it is worth emphasising how smoothing is carried out in order to understand the numerical results presented in the remainder of this chapter.

The intuitive choice of smoothing method, discussed previously, is Gaussian smoothing: convolution of the dielectric function with a Gaussian in real space. This can be performed rapidly in reciprocal space by multiplication of the Fourier components n^2_G by another Gaussian:

$$n^2_G \to n^2_G \exp\left(-\frac{G^2 \Gamma^2}{16 \ln 2} \right), \tag{4.14}$$

where Γ is the full-width half-maximum (FWHM) of the real-space convolving Gaussian. In the case of TM polarisation, where components of $\ln n^2$ as well as of n^2 are required, it is important that the components of $\ln n^2$ remain consistent with those of n^2. For a calculation on an FFT grid of size N, we generate the unsmoothed n^2 on a fine real-space grid (with $\approx 8N$ points in the unit cell), then take the Fourier transform (by FFT) to reciprocal space and perform the smooth-

ing (using Fourier components with magnitude up to $G \approx 8G_{\max}$), and then invert to real space and take the logarithm to create the function $\ln n^2$. Both n^2 and $\ln n^2$ are then transformed to reciprocal space once more, and only components with $|G| < G_{\max}$ are selected. The combination of Gaussian smoothing and an oversampled fine FFT grid ensures that the amplitude of Fourier components at the edge of the fine FFT grid is negligible. This prevents any aliasing of the components of either n^2 or $\ln n^2$ (i.e. the appearance of spurious components at large G as a result of the sampling of the FFT [76]), and ensures that the same dielectric function is described by both sets of coefficients.

4.4 Results

The results presented in this section compare calculations carried out using the fixed-frequency plane-wave method to the analytical solutions of Sec. 4.2.2. In particular two effects are important: firstly, the error introduced by smoothing relative to the analytical result obtained for a sharp interface and, secondly, the way in which smoothing affects the convergence of calculations at different basis set sizes.

4.4.1 Systematic error introduced by smoothing

In order to study the systematic error introduced by smoothing, *fully converged* plane-wave calculations were carried out and the values of β so obtained were compared with those from analytical solution of the 1D model waveguide. Plane-wave convergence is essential, so that any deviation of β from its true (sharp interface) value β_0 is due only to smoothing and is not an issue of numerical convergence. To ensure this, all calculations in this section were carried out on fine grids of dimension $2^{19} = 524,288$. Equivalently, the spatial resolution of the real-space FFT grid was $\Delta x \approx 2 \times 10^{-6}\Lambda$, which is two orders of magnitude less than the size of any feature in the dielectric function[2]. There is further discussion of the criterion for plane-wave convergence and how it is related to spatial resolution below in Sec. 4.4.2.

Fig. 4.6 shows the error in propagation constant introduced by smoothing for an

[2]Note that it impossible to obtain such a resolution in 2D using a desktop PC. Storage of the FFT grid requires $\approx$ 8 MB but in 2D an equivalent spatial resolution would require $\approx$ 4 TB.

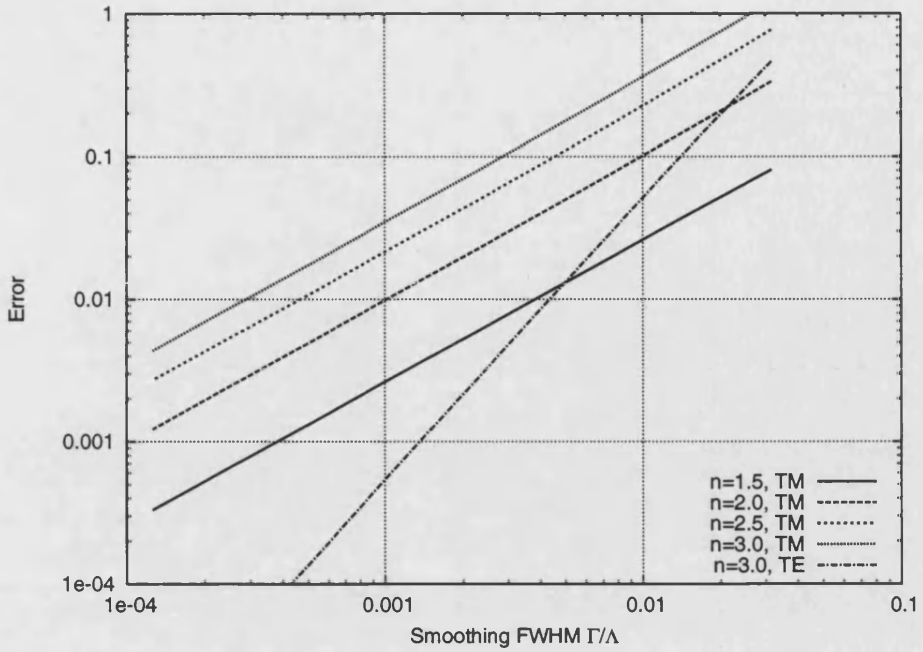


Figure 4.6: Variation of systematic error $|\beta - \beta_0|\Lambda$ with smoothing width Γ for the bound even TM mode of the 1D slab waveguide with $d = 0.1\Lambda$. The line with steeper slope shows the error for the bound even TE mode of the waveguide with $n = 3.0$ and is shown for comparison.

example system of glass slabs of $d = 0.1\Lambda$ in air. A range of glass-like refractive indices are used from $n = 1.5$ (approximately that of silica) to $n = 3.0$ (a little over that of chalcogenide glass [45]); most glasses used in PCF have refractive indices in this range. The mode under consideration is the bound even TM mode (see, for example, Fig. 4.2, leftmost mode), but the results do not differ significantly for any other TM mode. Also shown, for comparison, is the smoothing error for the bound even TE mode (see, for example, Fig. 4.1, leftmost mode).

Much useful information can be obtained from Fig. 4.6. Firstly, it shows that the error introduced is itself smoothly-varying and monotonic, i.e. that decreasing the smoothing width always decreases the error. This is in contrast to the plane-wave convergence error, discussed below. Secondly, it shows that the error introduced by smoothing is dependent on refractive index, becoming more significant at larger refractive indices. This is important to note when modelling PCF comprising high-index glass, as in Chapter 7.

More quantitatively, Fig. 4.6 provides a quick way to estimate an appropriate level of smoothing (at least to the correct order of magnitude) to obtain a given

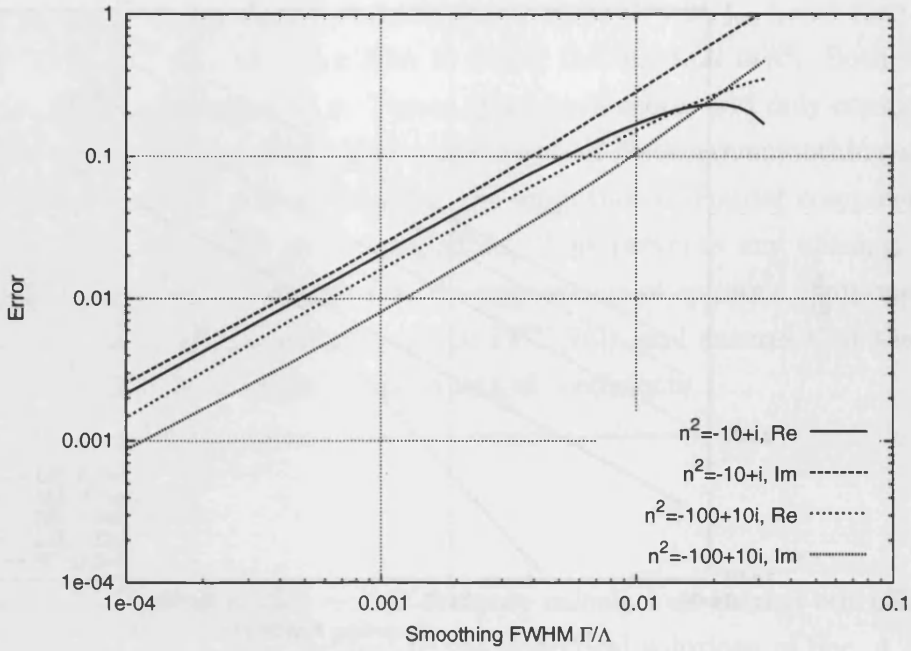


Figure 4.7: Variation of systematic error $|\Re(\beta - \beta_0)|\Lambda$ and $|\Im(\beta - \beta_0)|\Lambda$ with smoothing width Γ for the bound even TM mode of the 1D slab waveguide with $d = 0.1\Lambda$, for two metal-like dielectric constants.

accuracy of final result. For example, if $\beta\Lambda$ was required to within 0.01 for a glass of $n = 2.0$, the required smoothing width will be $\Gamma \lesssim 0.001\Lambda$.

The difference between TE and TM modes highlighted in Fig. 4.6 is a result of the different boundary conditions for the two polarisations. However, as previously noted, it is not possible in 2D to differentiate between TE and TM modes. The TM mode equation is the most similar to the vector wave equation in 2D and, as shown in Fig. 4.6, it is also the TM modes which have the more significant smoothing error. It is therefore important for estimates such as the one in the above paragraph to use the results for TM polarisation.

A similar plot for metal slabs ($n^2 = -10 + i$, representing gold at $\lambda \approx 0.6\,\mu$m and $n^2 = -100 + 10i$, gold at $\lambda \approx 1.6\,\mu$m [88]) is given in Fig. 4.7. Again the bound even TM mode is considered (see Fig. 4.5). Similar behaviour is seen here to that in Fig. 4.6, with error decreasing monotonically with smoothing width for all but the largest smoothings.

It is important to note from Fig. 4.7 that, although the errors in $\Re(\beta)$ and $\Im(\beta)$ behave similarly, the *relative* errors $|\Re\beta - \beta_0)|/\Re(\beta_0)$ and $|\Im(\beta - \beta_0)|/\Im(\beta_0)$ are

very different. To illustrate this, consider the analytical solution for $n^2 = -10+i$ which is, to 5 d.p., $\beta\Lambda = 10.56143+0.03795i$. At a smoothing width of $\Gamma = 10^{-4}\Lambda$ (i.e. the left-hand edge of Fig. 4.7), the relative errors in $\Re(\beta)$ and $\Im(\beta)$ are respectively 0.02% and 7%. Although the error in the real part is acceptable, that in the imaginary part is relatively large.

4.4.2 Plane-wave convergence

Once a level of smoothing has been chosen that results in a sufficiently small deviation from the true values of β, it is necessary to determine the size of basis set (i.e. the number of plane waves or, equivalently, size of FFT grid) needed to obtain a converged calculation. The 1D model provides an ideal means to analyse this behaviour because it can easily handle basis set sizes over many orders of magnitude. As will be demonstrated in this section, the error introduced by non-converged basis sets is generally less predictable than the smoothing error and it is important to ensure that all realistic calculations are carried out with converged basis sets.

The effect of smoothing on plane-wave convergence is first demonstrated by plotting values of β for a specific mode as the size of the FFT grid is increased. Because the cutoff $G_{\max}$ is extended to the edge of the FFT grid, the size N of the FFT grid is directly proportional to N_{PW}. The corresponding spatial resolution in real space is $\Delta x = \Lambda/N$.

Fig. 4.8 shows β values obtained for the bound even TM mode of the same example 1D model as that considered in Sec. 4.4.1 (dielectric slabs of thickness $d = 0.1\Lambda$ in air at a frequency of $k_0\Lambda = 10$), in this case with $n = 1.5$. Three levels of smoothing are shown together with the results obtained with no smoothing applied. The error introduced by smoothing is clearly shown by the converged values towards the right-hand side of the plot, which become closer to the analytical solution as the smoothing width is decreased. However, the calculation with an unsmoothed dielectric function appears *never* to converge to the correct value of β. Although it appears in Fig. 4.8 to converge to a value of $\beta\Lambda \approx 10.462$, there is no analytical solution at this value for either TM or TE modes[3]. When larger refractive index steps are used, the discrepancy is greater still: it is ≈ 5

[3]For completeness, the analytical values of $\beta\Lambda$ to 2 d.p. for this structure are: 11.34 (odd TE); 10.49 (even TM, shown); 8.87 (even TE); 8.29 (odd TM); 8.28 (even TM); 7.83 (odd TE).

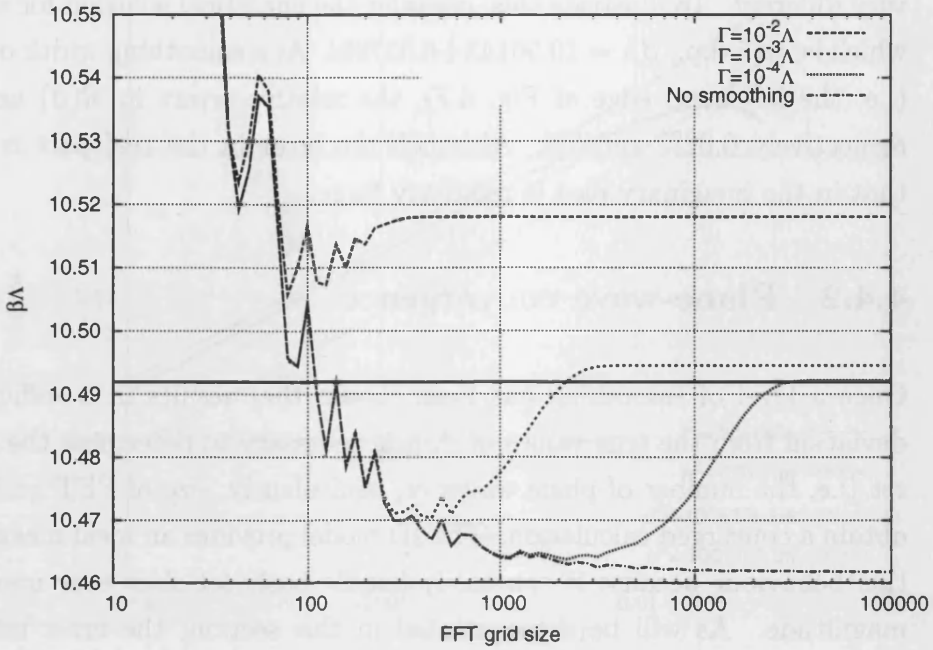


Figure 4.8: Plane-wave convergence of the bound even mode of the periodic dielectric slab waveguide with $d = 0.1\Lambda$, $k_0\Lambda = 10$ and $n = 1.5$. The solid horizontal line shows the analytical result for infinitely-sharp interfaces. Note the systematic error in the converged results introduced by smoothing and the spuriously 'converged' result for the unsmoothed structure.

for $n = 3.0$ (see Fig. 4.10). This error demonstrates that it is essential to use smoothing. If smoothing is not used (or insufficient smoothing is used), it is possible for calculations to appear to be converged with respect to basis set size, but with β values that are incorrect. A similar problem has been encountered in 2D [90].

It is noticeable that there is an approximate inverse proportionality between the FFT size needed to obtain converged results and the smoothing width Γ, which is consistent with the $\Gamma = 0$ results never converging at all. In order to demonstrate this we plot in Fig. 4.9 the same data but on the horizontal axis show $N_\mathrm{I} = \Gamma/\Delta x = \Gamma N/\Lambda$, which is a measure of the number of real-space sampling points (spaced by Δx) in the interface region (width Γ) of the unit cell. An equivalent plot for a glass with $n = 3.0$ is shown in Fig. 4.10.

The behaviour shown in Figs. 4.9 and 4.10 is that of convergence occurring at approximately N_I of order unity for all smoothings and both dielectric functions. Although not shown here, this behaviour is also evident across the range of glass-

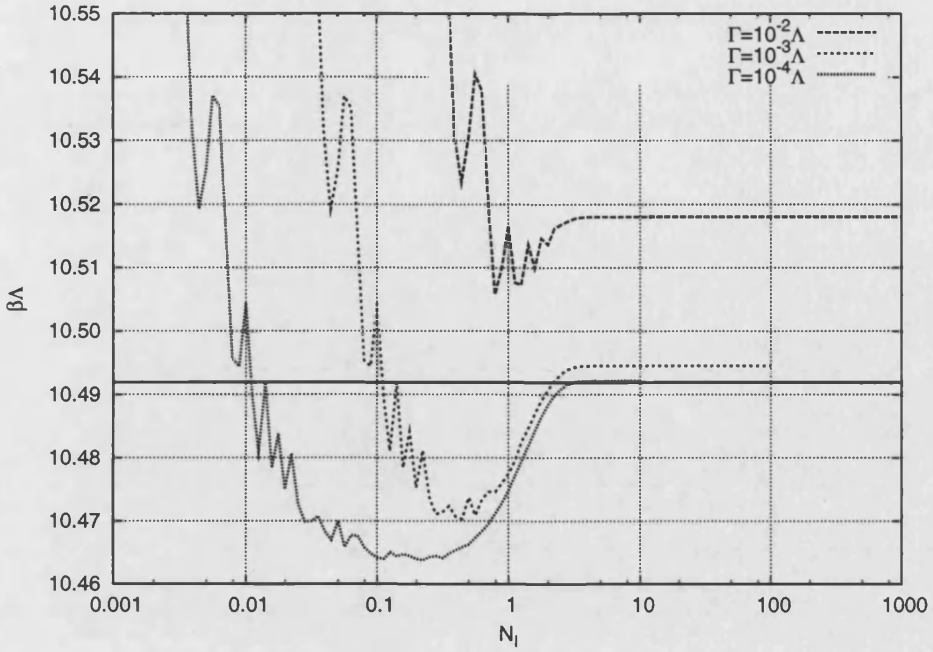


Figure 4.9: Plane-wave convergence of the waveguide described in Fig. 4.8 but with the horizontal axis showing $N_{\mathrm{I}} = \Gamma N/\Lambda$. The solid horizontal line shows the analytical result for infinitely-sharp interfaces. The unsmoothed structure is not plotted because $N_{\mathrm{I}} = 0$ for an infinitely-sharp interface.

like dielectric functions and for all TM modes, and is not dependent on the frequency or the width of dielectric slabs. TE modes generally converge sooner in the range $0.1 \lesssim N_{\mathrm{I}} \lesssim 1$, so we consider here only TM modes as the more difficult calculations to converge.

We can consider the effect of smoothing on plane-wave convergence qualitatively in terms of the sampling theorem. The sampling theorem, usually expressed with time t and frequency f as conjugate variables (but equally applicable to a spatial variable and spatial frequency), states that any continuous function $\phi(t)$ sampled discretely at intervals of Δt is *completely* determined by its samples provided that it has no components with frequencies $|f| \geq 1/2\Delta t$ [76]. If $\phi(t)$ contains higher frequencies, aliasing causes the sampled representation to be inaccurate. Because we use FFTs to create a sampled representation of functions in real space, we can use a similar idea to understand the sampling of a dielectric function. An interface of width Γ has equivalent spatial frequencies of order $1/\Gamma$, so we can expect a reasonably accurate description (i.e. convergence) to occur only when $1/\Delta x \gtrsim 1/\Gamma$, or $N_{\mathrm{I}} \gtrsim 1$.

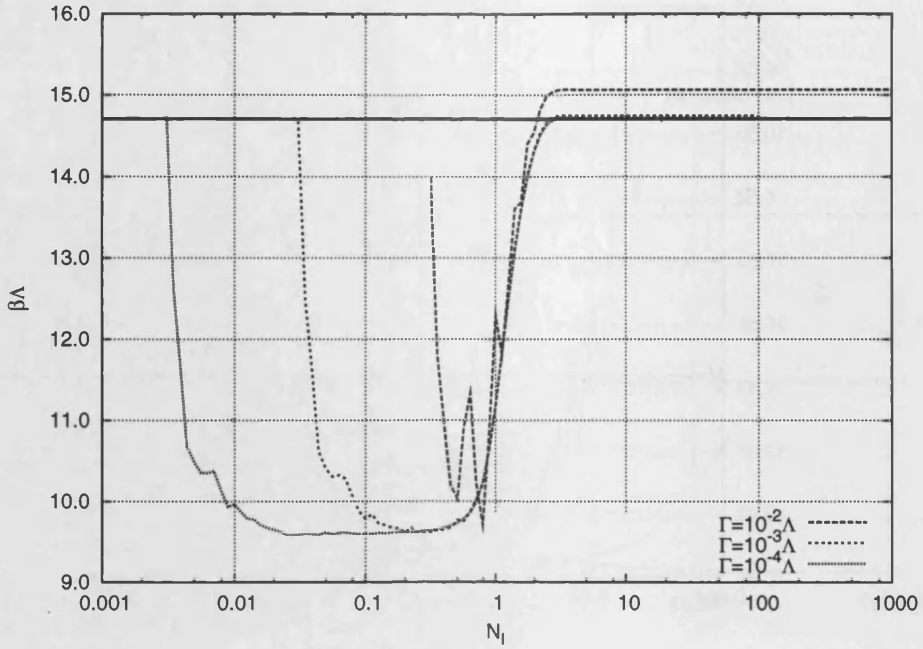


Figure 4.10: Plane-wave convergence of the bound even mode of the dielectric slab waveguide with $d = 0.1\Lambda$, $k_0\Lambda = 10$ and $n = 3.0$. The horizontal line shows the analytical result for infinitely-sharp interfaces. Note the change of vertical scale relative to Fig. 4.9; the error caused by lack of plane-wave convergence is greater for larger refractive index steps. The converged result for $\Gamma = 10^{-4}\Lambda$ is not visible beneath the solid black line on this scale.

The condition $N_\mathrm{I} \gtrsim 1$ specifies that there is on average at least one real-space sampling point in the smoothed interface region[4], and we refer to calculations carried out in this regime as being 'converged in principle'. Increasing the smoothing width or the density of grid points in real space (both of which increase N_I) enhances convergence, and if $N_\mathrm{I} \gtrsim 1$ we expect this convergence to be smooth[5]. Calculations that are not 'converged in principle' can produce erratic and unpredictable errors; this behaviour is demonstrated in Figs. 4.9 and 4.10.

It is also interesting to plot the convergence of $\Re(\beta\Lambda)$ and $\Im(\beta\Lambda)$ for metallic slabs to determine whether the same criterion for convergence can be applied to these structures. Fig. 4.11 shows the convergence of $\Re(\beta\Lambda)$, and Fig. 4.12 shows that of $\Im(\beta\Lambda)$. A similar criterion does seem to hold, but it is worth noting that $\Im(\beta\Lambda)$ in particular remains not only inaccurate but also unphysical (for $\Gamma \leq 10^{-2}\Lambda$,

[4]In 2D, we take this to mean at least one real-space sampling point in the direction normal to the dielectric interface at any point on the interface.

[5]The calculations shown in Sec. 4.4.1 were carried out on an FFT grid of dimension $N = 524,288$ and the narrowest smoothing widths shown were $\Gamma = 10^{-4}\Lambda$. This gives $N_\mathrm{I} \approx 52$ and confirms that these results were well converged.

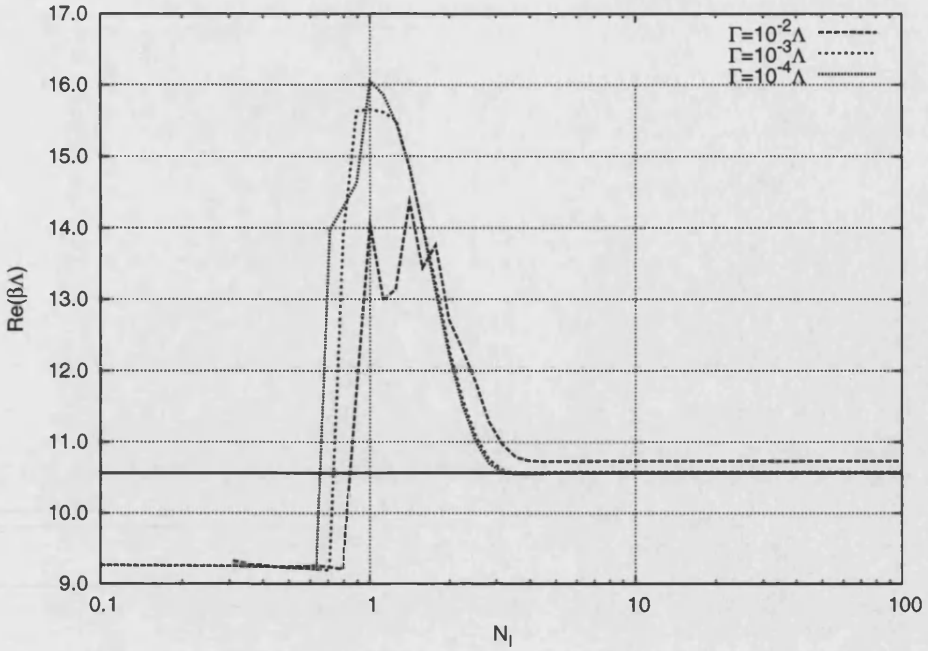


Figure 4.11: Plane-wave convergence of $\Re(\beta\Lambda)$ of the bound even mode of the dielectric slab waveguide with $d = 0.1\Lambda$, $k_0\Lambda = 10$ and $n^2 = -10 + i$.

$\Im(\beta\Lambda) < 0$, implying gain that the model does not contain) until $N_\mathrm{I} \approx 5$. This is slightly larger than the value of N_I needed for glasses and suggests that more care must be taken when considering metallic dielectrics.

To determine the feasibility of fixed-frequency plane-wave calculations involving metallic dielectrics, it is necessary to consider both the error produced by smoothing (analysed in Sec. 4.4.1) and the plane-wave convergence behaviour discussed here. Fig. 4.7 suggests that physically useful results can only be obtained for sharp interfaces with a width of $\Gamma \lesssim 10^{-4}\Lambda$. However, the 'convergence in principle' criterion then implies that FFT grids of size $N \approx 10,000$ are needed in order to obtain converged results. In 2D a practical limit set by the amount of memory available on a desktop PC is $N \approx 1,024$, suggesting that calculations on such large FFT grids are not feasible.

4.5 Conclusions

The 1D model presented in this chapter has demonstrated the systematic error introduced by Gaussian smoothing and shown why smoothing is necessary when using the fixed-frequency plane-wave method. The key results of the analysis

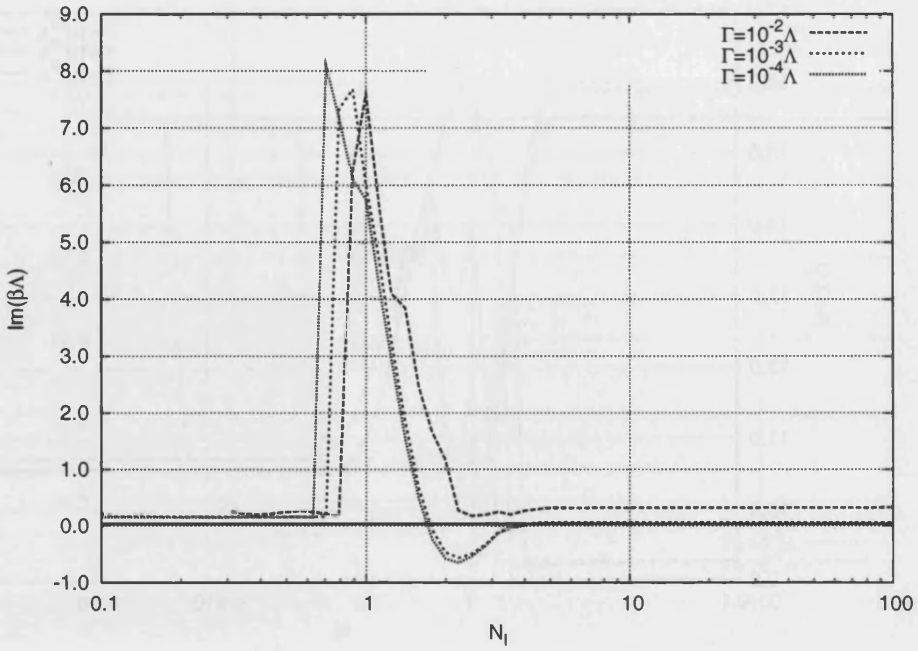


Figure 4.12: Plane-wave convergence of $\Im(\beta\Lambda)$ of the bound even mode of the dielectric slab waveguide with $d = 0.1\Lambda$, $k_0\Lambda = 10$ and $n^2 = -10 + i$.

of smoothing in Sec. 4.4.1 are summarised by Fig. 4.6, which demonstrates the systematic smoothing error and how it is a function of both the smoothing width and the dielectric function. Because the basic structure of the 2D eigenproblem is retained in 1D (at least for TM modes), these results are useful in providing order-of-magnitude estimates of smoothing errors at different smoothing widths which should also be applicable to 2D calculations.

The combination of errors introduced by smoothing, which is particularly significant for $\Im(\beta\Lambda)$, and plane-wave convergence (again apparently worse for imaginary parts), suggests that metallic dielectrics are particularly difficult structures to model accurately using the fixed-frequency plane-wave method. A very narrow smoothing width is needed in order to obtain useful results, and the corresponding FFT grids that are required are impractical in 2D. For this reason, metallic structures will not be considered further in this thesis. However, they remain an interesting possibility for further work, particularly using the techniques developed in Chapter 5.

The investigation of plane-wave convergence at different smoothing widths and the resulting 'convergence in principle' criterion developed in Sec. 4.4.2 is of crucial importance in selecting appropriate FFT grid sizes and ensuring that calcu-

lations are well converged. It is also highly relevant to the work described in the following chapter, in which a method to increase the number of sampling points at the dielectric interfaces is developed without increasing either the smoothing width or the FFT grid size.

Chapter 5

Adaptive curvilinear coordinates in the fixed-frequency plane-wave method

The analysis presented in the previous chapter suggests that plane-wave convergence can be achieved only when, on average, there is at least one real-space sampling point in any smoothed interface region of a dielectric function. In this chapter, a method is presented that allows the number of sampling points in interface regions to be increased without the need to increase either the FFT grid size or the smoothing width. This is achieved by reformulating the fixed-frequency plane-wave method in adaptive curvilinear coordinates, enabling position-dependent plane-wave cutoffs to be chosen and adapted to the specific structure under consideration.

The chapter begins with an introduction to the reasons for developing the plane-wave method in curvilinear coordinates and some of the general ideas that underpin the method. The reformulation of the method in curvilinear coordinates is then presented in detail, including how the method is implemented in practice. Also given is a general method for the generation of coordinates adapted to dielectric structures.

5.1 Introduction

It is sometimes convenient to express Maxwell's equations in generalised curvilinear coordinates (GCCs), and many examples exist of uses for GCCs in the field of photonics. These include uses in the finite-difference method [91] and FDTD [92], in grating theory [93], and to create perturbation-matched coordinates in coupled mode theory [94, 95]. However, the first application of GCCs specifically to the plane-wave method was proposed in the field of electronic structure in 1992 by Gygi [96, 97].

Gygi noted that solutions of the Schrödinger equation vary rapidly only in the vicinity of atomic cores, and they decay smoothly in the vacuum surrounding them. However, the FFT-based formulation of the plane-wave method requires uniform grid spacing throughout the unit cell, and so gaining an accurate description of the rapidly-varying wavefunctions near atoms implies also devoting a large computational effort to the empty parts of the unit cell. The necessity of using uniformly-spaced grids is therefore a cause of computational inefficiency. To reduce this inefficiency, Gygi introduced the reformulation of the plane-wave method in GCCs and demonstrated that it can be used to give a position-dependent plane-wave cutoff, which could be increased near atomic cores relative to its value in vacuum. Despite this nonuniformity, the use of FFTs can be retained and the favourable properties of the plane-wave method are therefore unaffected.

It is possible to apply similar reasoning to PCF. The structure of PCF is such that there are usually large regions of the unit cell with no variation of n^2, and these are separated by sharp interfaces. Like the atoms in the calculations considered by Gygi, these interfaces occupy only a relatively small fraction of the unit cell. It was shown in the previous chapter that the convergence behaviour of the plane-wave method is dominated by the sampling frequency at the interfaces. It would therefore be beneficial to increase the sampling frequency *only* at the interfaces in the same way that Gygi's method increased the sampling frequency near to atomic cores.

In order to understand how the GCC method gives a position-dependent plane-wave cutoff and sampling frequency, consider first the conventional method as described in Chapter 3. A cutoff condition $|\mathbf{G}| < G_{\mathrm{max}}$ is imposed to create a

finite set of basis functions, and these are connected by the FFT to an equivalent representation on a real-space grid with coordinates $\mathbf{x}$ having a uniform spacing $\Delta x \propto 1/G_{\max}$. Suppose instead that the real-space sampling of a function was uniform in another coordinate, $\boldsymbol{\xi}$, related to $\mathbf{x}$ by a smooth and invertible mapping. Because the '$\boldsymbol{\xi}$-space' grid is uniform, we are still able to use FFTs of quantities stored on it and therefore perform calculations rapidly. However, the grid of $\mathbf{x}$ points corresponding to the uniform $\boldsymbol{\xi}$-space grid is in general non-uniform. By choosing the coordinate transformation appropriately we can increase the density of grid points $1/\Delta x$ at chosen locations in $\mathbf{x}$ space whilst retaining uniformity in $\boldsymbol{\xi}$ space[1].

In the following section, the plane-wave method in GCCs is developed for an arbitrary coordinate transformation, and details of its implementation are then given in Sec. 5.3. A method to create an appropriate coordinate transformation for any given PCF structure is described in Sec. 5.4.

5.2 The fixed-frequency plane-wave method in generalised curvilinear coordinates

In this section the plane-wave method is developed in generalised curvilinear coordinates. The description given here is purely mathematical; details of how it is implemented in practice are presented in the following section.

Following the method of Gygi [96, 97], consider the arbitrary curvilinear coordinate $\boldsymbol{\xi}$, a vector of two Cartesian components ξ^1 and ξ^2, and define a $\boldsymbol{\xi}$-space lattice to have primitive lattice vectors $\mathbf{R}_1$ and $\mathbf{R}_2$. Note that these are the same lattice vectors as those in Fig. 2.1 on page 34; they define the repeating unit cell of the periodic structure under consideration. In order to make a mapping between the curvilinear coordinate and the Euclidean coordinate $\mathbf{x} = (x^1, x^2)$, we define a transformation by

$$\mathbf{x}(\boldsymbol{\xi}) = \boldsymbol{\xi} + \sum_{\mathbf{G}} \mathbf{x}_{\mathbf{G}} \exp(i\mathbf{G} \cdot \boldsymbol{\xi}), \tag{5.1}$$

where $\mathbf{G} = m_1\mathbf{G}_1 + m_2\mathbf{G}_2$ for $m_1, m_2 \in \mathbb{Z}$, and the primitive reciprocal lattice

[1] This is equivalent to having a 'local' effective plane-wave cutoff that is position-dependent: $G^{\text{eff}}_{\max}(\mathbf{x}) \propto 1/\Delta x(\mathbf{x})$.

vectors $\mathbf{G}_1$ and $\mathbf{G}_2$ are identical to those of the reciprocal lattice in Euclidean space. Eq. (5.1) ensures, by definition, that the transformation has the same periodicity as the dielectric function, which is necessary in order to use the plane-wave method. It is also a convenient choice because a plane-wave representation of the transformation allows rapid and simple computation of its derivatives, which are required elsewhere (see below). The coefficients $\mathbf{x_G}$ fully define the transformation and are given by

$$\mathbf{x_G} = \frac{1}{\Omega} \int_{\substack{\text{unit} \\ \text{cell}}} [\mathbf{x}(\boldsymbol{\xi}) - \boldsymbol{\xi}]\, e^{-i\mathbf{G}\cdot\boldsymbol{\xi}}\, \mathrm{d}^2\xi, \tag{5.2}$$

where Ω is the area of the unit cell. In general, two restrictions are placed on the transformation. Firstly, it must be invertible (i.e. a 1:1 mapping between $\mathbf{x}$ and $\boldsymbol{\xi}$). Secondly, it must have the same periodicity as the dielectric function under consideration (although the chosen form of Eq. (5.1) ensures this automatically).

The mathematics of coordinate transformations is commonly encountered in another area of physics making use of generalised curvilinear coordinates: that of general relativity [98]. It is common in general relativity to use the summation convention in which variables are written with either covariant (subscript) or contravariant (superscript) indices, and a sum is always implied over repeated covariant and contravariant indices (e.g. $g_{ab}g^{ac} \equiv \sum_a g_{ab}g^{ac}$). In this thesis, all sums have been shown explicitly.

It is useful for future use to define here the covariant metric tensor for the transformation (a quantity commonly used in general relativity) by

$$g_{pq} = \sum_{k=1}^{2} \frac{\partial x^k}{\partial \xi^p} \frac{\partial x^k}{\partial \xi^q}. \tag{5.3}$$

From g_{pq}, a position-dependent 2×2 matrix, it is also possible to define the related contravariant metric tensor g^{pq} by inversion at any point in space, i.e. $[g^{pq}] = [g_{pq}]^{-1}$. Another useful quantity used by Gygi is the 'geometric vector potential', which is defined by

$$A_p = \frac{1}{4}\frac{\partial}{\partial \xi^p} \ln g = \frac{1}{4g}\frac{\partial g}{\partial \xi^p}, \tag{5.4}$$

where $g = \det g_{pq}$ [97][2]. It is worth noting at this stage that g_{pq}, g^{pq} and A_p are all defined completely by the set of transformation coefficients $\mathbf{x_G}$. A description of how they are created and stored is given in Sec. 5.3.

It will be necessary later to change variables of integration between $\mathbf{x}$ and $\boldsymbol{\xi}$, and to do this it is necessary to know the Jacobian J appearing in the expression $\mathrm{d}^2x = J\,\mathrm{d}^2\xi$. It is well-known that J is given in 2D by

$$J = \det\left[\frac{\partial x^p}{\partial \xi^q}\right], \tag{5.5}$$

and it can be shown that, for arbitrary coordinates, $J = g^{1/2}$ [98]. For computational purposes either expression can be used.

To use the plane-wave method in curvilinear coordinates, it is necessary to define a set of plane-wave basis functions. Consider the set of basis functions given by

$$\chi_{\mathbf{k},\mathbf{G}}(\mathbf{x}) = g^{-1/4}(\mathbf{x})\, e^{i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}(\mathbf{x})}. \tag{5.6}$$

Orthogonality of the basis set is shown by taking the inner product:

$$\begin{aligned}
\int \chi^*_{\mathbf{k},\mathbf{G}}\, \chi_{\mathbf{k},\mathbf{G}'}\, \mathrm{d}^2x &= \int e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}} g^{-1/2}\, \mathrm{d}^2x && \text{(5.7a)}\\
&= \int e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}\, \mathrm{d}^2\xi && \text{(5.7b)}\\
&= \Omega\delta_{\mathbf{G},\mathbf{G}'}, && \text{(5.7c)}
\end{aligned}$$

where the property $\mathrm{d}^2x = g^{1/2}\mathrm{d}^2\xi$ has been used to transform the area element. It can also be shown that the set is complete [97]. The magnetic field and dielectric function can be represented as a sum over these basis functions, but note that the coefficients are different from those obtained using the conventional plane-wave method given by Eqs. (3.3) and (3.5):

$$\begin{aligned}
n^2(\mathbf{x}) &= g^{-1/4}(\mathbf{x}) \sum_{\mathbf{G}} n^2_{\mathbf{G}} \exp[i\mathbf{G}\cdot\boldsymbol{\xi}(\mathbf{x})] && \text{(5.8)}\\
h^i(\mathbf{x}) &= g^{-1/4}(\mathbf{x}) \sum_{\mathbf{G}} h^i_{\mathbf{k},\mathbf{G}} \exp[i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}(\mathbf{x})]. && \text{(5.9)}
\end{aligned}$$

[2]The geometric vector potential A_p is related to another quantity commonly encountered in general relativity, the 'affine connection' or Christoffel symbol of the second kind: $A_p = \frac{1}{2}\sum_q \Gamma^q_{qp}$ [97,98].

As in the conventional formulation of the fixed-frequency plane-wave method, the reciprocal-space matrix equation is found by substituting the plane-wave expansions of the fields into Eq. (2.18). A detailed derivation of the reciprocal-space matrix equation and its constituent matrix elements is given in Appendix B. In summary, the reciprocal-space matrix equation takes the form:

$$\begin{pmatrix} [\nabla_t^2 h^x]_{\mathbf{G}} + k_0^2 [n^2 h^x]_{\mathbf{G}} - [(\frac{\partial}{\partial y} \ln n^2)(\frac{\partial h^x}{\partial y})]_{\mathbf{G}} + [(\frac{\partial}{\partial y} \ln n^2)(\frac{\partial h^y}{\partial x})]_{\mathbf{G}} \\ [\nabla_t^2 h^y]_{\mathbf{G}} + k_0^2 [n^2 h^y]_{\mathbf{G}} + [(\frac{\partial}{\partial x} \ln n^2)(\frac{\partial h^x}{\partial y})]_{\mathbf{G}} - [(\frac{\partial}{\partial x} \ln n^2)(\frac{\partial h^y}{\partial x})]_{\mathbf{G}} \end{pmatrix} = \beta^2 \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}} \\ h^y_{\mathbf{k},\mathbf{G}} \end{pmatrix}, \quad (5.10)$$

where the elements appearing on the LHS may be constructed from the following set of operations:

$$[\nabla_t^2 h^i]_{\mathbf{G}} = -\sum_{p,q=1}^{2} \int \mathrm{d}^2\xi \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} \times (k_p + G_p - iA_p) \times g^{pq} \times (k_q + G'_q + iA_q) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}} \quad (5.11a)$$

$$\left[f(\boldsymbol{\xi}) \frac{\partial h^i}{\partial x^j} \right]_{\mathbf{G}} = i \sum_{p=1}^{2} \int \mathrm{d}^2\xi \, f(\boldsymbol{\xi}) \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} \frac{\partial \xi^p}{\partial x^j} (k_p + G'_p + iA_p) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}} \quad (5.11b)$$

$$[n^2 h^i]_{\mathbf{G}} = \int \mathrm{d}^2\xi \, n^2(\boldsymbol{\xi}) \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}, \quad (5.11c)$$

where $f(\boldsymbol{\xi}) = \frac{\partial \ln n^2}{\partial x^k}$ as discussed in Appendix B. Although Eq. (5.10) is presented as an equation in $\mathbf{G}$ only, it can be seen by reference to Eqs. (5.11a)-(5.11c) that it represents a matrix equation in $\mathbf{G}$,$\mathbf{G}'$ of the same form as Eq. (3.6) in the conventional formulation of the fixed-frequency plane-wave method. Expanding the full matrix showing the sum over $\mathbf{G}'$ is cumbersome and the form shown above has been used for brevity.

Note that if $\mathbf{x}_{\mathbf{G}} = \mathbf{0}$ (i.e. the $\boldsymbol{\xi}$ and $\mathbf{x}$ coordinate systems are identical), we obtain $g^{pq} = \delta^{pq}$, $A_p = 0$ and $\mathrm{d}^2\xi = \mathrm{d}^2x$: in this limit, Eqs. (5.10) and (5.11a)-(5.11c) reduce to those obtained for the conventional formulation of the fixed-frequency plane-wave method.

The following section discusses how the reciprocal-space matrix equation may be solved.

5.3 Implementation

The method of solution of the reciprocal-space matrix equation in GCCs is in essence similar to that of the conventional formulation of the fixed-frequency plane-wave method outlined in Chapter 3. As previously discussed, it is possible to write the reciprocal-space matrix equation in the form $\mathbf{Mv} = \beta^2\mathbf{v}$ and solve for β directly, but storage of $\mathbf{M}$ and direct diagonalisation is undesirable. It is instead preferable to use an iterative eigensolver and iterative linear solver in combination to determine the eigenvalues and eigenvectors of $\mathbf{M}$. The iterative linear solver requires only the result of the matrix-vector multiplication, which is provided by FFTs. The only differences between the conventional formulation of the method and that in GCCs are in how the matrix-vector operation is carried out using FFTs, and how the various quantities (e.g. g^{pq}) required to do this are generated.

In this section the generation of coordinate-dependent 'fixed' quantities (those that are evaluated once before the eigensolver is used) is first described, followed by the implementation of matrix-vector operations by FFTs.

5.3.1 Generation of coordinate-related fixed quantities

The general principle of the GCC method is that $\boldsymbol{\xi}$-space is sampled regularly, enabling the use of FFTs to perform calculations, but $\boldsymbol{\xi}$-space coordinates are mapped onto $\mathbf{x}$-space coordinates such that the sampling in $\mathbf{x}$-space is non-uniform, giving a position-dependent sampling frequency of the actual dielectric structure. Methods of choosing the mapping between $\boldsymbol{\xi}$ and $\mathbf{x}$ to achieve the desired sampling frequency are described below in Sec. 5.4. Once this coordinate transformation is obtained, it is possible to generate quantities relating to it and store these before performing a mode-solving calculation. The generation of these 'fixed' quantities relating to the mapping is described in this section.

It is assumed here that the set of coefficients $\mathbf{x}_\mathbf{G}$ with $|\mathbf{G}| < G^{\mathrm{x}}_{\mathrm{max}}$ (where $G^{\mathrm{x}}_{\mathrm{max}}$ is a cutoff imposed to make the set of coefficients finite) *define* the transformation and have previously been determined (for example, by the method described in Sec. 5.4.1). It is essential that $G^{\mathrm{x}}_{\mathrm{max}}$ is chosen to be sufficiently small when using a given FFT grid that there is no aliasing of $\mathbf{x}_\mathbf{G}$ or the metric tensor. Provided the coordinate transformation is properly described without aliasing, the GCC

formulation is exact: the choice of coordinates does not introduce any additional approximation into the plane-wave method.

From $\mathbf{x_G}$ it is possible to use FFTs to generate in real space the set of four first derivatives that comprise the metric tensor. Differentiating Eq. (5.1) with respect to ξ gives

$$\frac{\partial x^p}{\partial \xi^q} = \delta^p_q + \sum_{\mathbf{G}} iG_q x^p_{\mathbf{G}} e^{i\mathbf{G}\cdot\boldsymbol{\xi}}, \tag{5.12}$$

where the sum over $\mathbf{G}$ is evaluated by two FFTs (for $p = 1, 2$) and therefore the full set of derivatives ($p, q = 1, 2$) is evaluated using a total of four FFTs. Eq. (5.3) is then used to create the metric tensor g_{pq} on the $\boldsymbol{\xi}$-space grid, and inversion of this matrix gives g^{pq}, which is stored for future use. The determinant g given by $g = g_{11}g_{22} - g_{12}^2$ is also stored on a $\boldsymbol{\xi}$-space grid[3]. Taking the logarithm and using a single inverse FFT to determine the set of coefficients $[\ln g]_{\mathbf{G}}$, the geometric vector potential A_p can then be calculated using

$$A_p = \frac{1}{4} \sum_{\mathbf{G}} iG_p [\ln g]_{\mathbf{G}} e^{i\mathbf{G}\cdot\boldsymbol{\xi}}. \tag{5.13}$$

Again the sum over $\mathbf{G}$ is evaluated with two FFTs, and A_p is stored on a $\boldsymbol{\xi}$-space grid.

One further coordinate-dependent quantity is needed in order to perform matrix-vector operations by FFTs, the set of derivatives $\frac{\partial \xi^p}{\partial x^q}$. These derivatives may be obtained conveniently from quantities already stored, as follows:

$$\frac{\partial \xi^p}{\partial x^q} = \sum_{k=1}^{2} g^{kp} \frac{\partial x^q}{\partial \xi^k}. \tag{5.14}$$

5.3.2 Generation of structure-dependent fixed quantities

Once the coordinate transformation and related quantities have been defined, it is possible to generate and store quantities related to the dielectric function on $\boldsymbol{\xi}$-space grids. However, the use of GCCs introduces a complication in the way in which smoothing is applied, and this is discussed below.

[3]It is convenient at this point to check that $g \geq 0$ at all points on the grid. If this condition is not satisfied, the coordinate transformation is not invertible.

The dielectric function n^2 is first sampled on a fine FFT grid in $\boldsymbol{\xi}$-space; as in the conventional plane-wave method, we use a fine grid of size $\approx 8N \times 8N$ when the mode-solving calculation will be eventually carried out on an $N \times N$ grid. For each point in $\boldsymbol{\xi}$-space, the corresponding $\mathbf{x}$-space point is found using the mapping given by Eq. (5.1) and the value of the dielectric function there is multiplied by $g^{1/4}$ and stored on the uniform $\boldsymbol{\xi}$-space grid[4]. One inverse FFT is taken to reciprocal space to obtain a set of plane-wave coefficients $n^2_{\mathbf{G}}$ stored on a fine grid.

At this point in the conventional formulation of the plane-wave method (see Sec. 3.2.3), smoothing is applied by the multiplication of the reciprocal-space components of the dielectric function by a Gaussian envelope. The convolution theorem ensures that this is equivalent to convolving the dielectric function with a Gaussian in real space, thereby smoothing the interfaces. The FWHM of this function can easily be chosen to give the required amount of smoothing. It is not necessary to perform the computationally-expensive convolution in real space explicitly[5]. In the GCC method, however, there is no simple way to perform Gaussian smoothing working in reciprocal-space only. Nevertheless, we wish to avoid performing a convolution in real space. Instead, we sample a Gaussian function $G[\mathbf{x}(\boldsymbol{\xi})]$ in real space on the fine uniform $\boldsymbol{\xi}$-space grid, multiply by $g^{1/4}$ (as with the dielectric function), and take the inverse FFT to produce a set of coefficients in reciprocal-space. These coefficients are then multiplied by the unsmoothed coefficients $n^2_{\mathbf{G}}$ to obtain coefficients of the smoothed structure. This does not give Gaussian smoothing, but it does provide control over the smoothing width and allows an approximate equivalence to Gaussian smoothing to be obtained. This is discussed with reference to an example structure in Sec. 6.1.2.

Coefficients of $[\ln n^2]_{\mathbf{G}}$ are also required. These coefficients are determined by first carrying out an FFT of the smoothed $n^2_{\mathbf{G}}$ coefficients to real space, multiplying by $g^{-1/4}$ to obtain n^2, and then taking the logarithm. The resulting representation of $\ln n^2$ is then multiplied by $g^{1/4}$ and a single inverse FFT creates $[\ln n^2]_{\mathbf{G}}$. The coefficients satisfying the circular cutoff condition $|\mathbf{G}| < G_{\max}$ are selected for use

[4]The multiplication by $g^{1/4}$ arises from Eq. (5.8), from which it can be seen that the quantity described by the plane-wave expansion is $g^{1/4}(\mathbf{x})n^2(\mathbf{x})$. Therefore, in order to obtain plane-wave coefficients, it is necessary to calculate the inverse FFT of $g^{1/4}(\mathbf{x})n^2(\mathbf{x})$.

[5]A real-space convolution of two functions on $N \times N$ grids requires, in general, effort of $O(N^4)$. Convolution with a narrow function can be carried out more quickly by imposing a finite spatial cutoff, but this is still slow compared with performing the convolution in reciprocal space.

in the mode-solving calculation in the same way as in the conventional plane-wave method.

Although the reciprocal-space components $n^2_{\mathbf{G}}$ and $[\ln n^2]_{\mathbf{G}}$ are needed in order to generate a preconditioner (as described for the conventional method in Sec. 3.3.2), the required quantities appearing in Eqs. (5.11b) and (5.11c) are $n^2(\boldsymbol{\xi})$ and $f(\boldsymbol{\xi}) = \frac{\partial \ln n^2}{\partial x^k}$. These must be generated from reciprocal-space components on the coarse (size $N \times N$) FFT grid. As described previously, $n^2(\boldsymbol{\xi})$ can be obtained by a single FFT of $n^2_{\mathbf{G}}$ followed by pointwise multiplication in real space by $g^{-1/4}$.

To generate $\frac{\partial \ln n^2}{\partial x^k}$ the gradient operator of Eq. (5.11b) may be used, with $f(\boldsymbol{\xi})$ in Eq. (5.11b) replaced by unity and h^i replaced by $\ln n^2$. This gives:

$$\left[\frac{\partial \ln n^2}{\partial x^k}\right]_{\mathbf{G}} = i \sum_{p=1}^{2} \int \mathrm{d}^2\xi \sum_{\mathbf{G}'} [\ln n^2]_{\mathbf{G}'} \frac{\partial \xi^p}{\partial x^k} (G'_p + iA_p) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \tag{5.15}$$

Because the gradient quantity is required in real space, we take the Fourier transform and rearrange to obtain

$$\frac{\partial \ln n^2}{\partial x^k}(\boldsymbol{\xi}) = i \sum_{p=1}^{2} \frac{\partial \xi^p}{\partial x^k} \left(\sum_{\mathbf{G}'} G'_p [\ln n^2]_{\mathbf{G}'} e^{i\mathbf{G}'\cdot\boldsymbol{\xi}} + iA_p \sum_{\mathbf{G}'} [\ln n^2]_{\mathbf{G}'} e^{i\mathbf{G}'\cdot\boldsymbol{\xi}} \right). \tag{5.16}$$

The bracketed part of this equation is of a very similar form to Eq. (5.19) and the sums over $\mathbf{G}'$ can be carried out using FFTs in the same way as described below. The multiplication by $\frac{\partial \xi^p}{\partial x^k}$ is carried out pointwise in real space using the quantities evaluated as described in Sec. 5.3.1.

5.3.3 Matrix-vector operations by FFT

The application of the 'Maxwell' operator (corresponding to multiplication by $\mathbf{M}$) is carried out in two parts in the conventional formulation: the differential operators are applied in reciprocal space, and the 'potential' multiplication in real space, as described in Sec. 3.2.3. In the GCC method, all operations make use of both spaces. The FFT-based method of evaluating all matrix-vector operations is given in this section.

Laplacian operator

The ∇_t^2 operation on the component h^i as given by Eq. (5.11a) is

$$[\nabla_t^2 h^i]_{\mathbf{G}} = -\sum_{p,q=1}^{2} \int \mathrm{d}^2\xi \, (k_p + G_p - iA_p) g^{pq} e^{-i\mathbf{G}\cdot\boldsymbol{\xi}} T_q^i, \tag{5.17}$$

where we have introduced T_q^i:

$$T_q^i = \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} (k_q + G'_q + iA_q) e^{i\mathbf{G}'\cdot\boldsymbol{\xi}}. \tag{5.18}$$

Noting that A_q is stored on a real-space grid we can separate this equation into two sums over $\mathbf{G}'$:

$$T_q^i = \sum_{\mathbf{G}'} (k_q + G'_q) h^i_{\mathbf{k},\mathbf{G}'} e^{i\mathbf{G}'\cdot\boldsymbol{\xi}} + iA_q \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i\mathbf{G}'\cdot\boldsymbol{\xi}}, \tag{5.19}$$

and evaluate the sums as FFTs of $(k_q + G'_q) h^i_{\mathbf{k},\mathbf{G}'}$ (one FFT for each of $q = 1, 2$) and $h^i_{\mathbf{k},\mathbf{G}'}$ (one FFT). The addition (and multiplication by iA_q) is then carried out pointwise on real-space grids using the stored, fixed A_q. This results in two functions on real-space grids, T_1^i and T_2^i. Next the sum over q in Eq. (5.17) is performed, again pointwise on real-space grids, giving:

$$U^{ip} = \sum_{q=1}^{2} g^{pq} T_q^i, \tag{5.20}$$

and

$$[\nabla_t^2 h^i]_{\mathbf{G}} = -\sum_{p=1}^{2} \int \mathrm{d}^2\xi \, (k_p + G_p - iA_p) U^{ip} e^{-i\mathbf{G}\cdot\boldsymbol{\xi}}. \tag{5.21}$$

This equation can be rearranged, again noting that A_p is a quantity stored on a real-space grid, to give:

$$[\nabla_t^2 h^i]_{\mathbf{G}} = -\sum_{p=1}^{2} \left\{ (k_p + G_p) \int \mathrm{d}^2\xi \, e^{-i\mathbf{G}\cdot\boldsymbol{\xi}} U^{ip} \right\} + i \int \mathrm{d}^2\xi \, e^{-i\mathbf{G}\cdot\boldsymbol{\xi}} \sum_{p=1}^{2} A_p U^{ip}. \tag{5.22}$$

It is possible to perform the integrals in this expression by inverse FFTs. The first integral is the inverse Fourier transform of U^{ip} and can therefore be calculated by two inverse FFTs, one for each of $p = 1, 2$. The second integral may be evaluated by a single inverse FFT of the summed quantity $\sum_p A_p U^{ip}$ (after carrying out the

sum pointwise in real space using the stored A_p). Finally, carrying out the sum over p involving the leftmost integral and combining the two integrated quantities in reciprocal space yields $[\nabla_t^2 h^i]_{\mathbf{G}}$.

In summary, the number of FFTs required to carry out this operation is 6:

1. $h^i_{\mathbf{k},\mathbf{G}'} \rightarrow$ real space
2. $(k_1 + G'_1)h^i_{\mathbf{k},\mathbf{G}'} \rightarrow$ real space
3. $(k_2 + G'_2)h^i_{\mathbf{k},\mathbf{G}'} \rightarrow$ real space
4. $U^{i1} \rightarrow$ reciprocal space
5. $U^{i2} \rightarrow$ reciprocal space
6. $\sum_p A_p U^{ip} = A_1 U^{i1} + A_2 U^{i2} \rightarrow$ reciprocal space.

As h^x and h^y components must be evaluated separately, the total number of FFTs needed is 12. This should be compared to the 5 needed in the conventional formulation of the plane-wave method (see Sec. 3.2.3); it represents an increased cost of using the GCC method compared to the conventional method. The computational cost of the GCC method is discussed in more detail in Sec. 6.2.3.

Gradient operator

The gradient operation is given by Eq. (5.11b), and can be written as

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = i\sum_{p=1}^{2}\int \mathrm{d}^2\xi\, f(\boldsymbol{\xi})\frac{\partial \xi^p}{\partial x^j}e^{-i\mathbf{G}\cdot\boldsymbol{\xi}}T_p^i, \tag{5.23}$$

where T_p^i is the same quantity as that previously evaluated for use in calculating the action of the Laplacian operator, and defined in Eq. (5.19). Carrying out the sum over p and performing the multiplication pointwise in real space gives another quantity W_j^i:

$$W_j^i = f(\boldsymbol{\xi})\sum_{p=1}^{2}\frac{\partial \xi^p}{\partial x^j}T_p^i, \tag{5.24}$$

and the desired gradient operation is then

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = i\int \mathrm{d}^2\xi\, W^i_j e^{-i\mathbf{G}\cdot\boldsymbol{\xi}}, \tag{5.25}$$

which can be evaluated by inverse FFT. Reference to Eq. (5.10) indicates that there are 4 combinations of $f(\boldsymbol{\xi})$, i and j required. However, the inverse FFTs to reciprocal space needed in Eq. (5.25) can be carried out by adding the relevant integrand to the second term of Eq. (5.22) while evaluating the Laplacian. The gradient terms can therefore be calculated without any additional FFTs.

'Potential' operator

The 'potential' operation is given by Eq. (5.11c):

$$\left[n^2 h^i\right]_{\mathbf{G}} = \int \mathrm{d}^2\xi\, n^2(\boldsymbol{\xi}) e^{-i\mathbf{G}\cdot\boldsymbol{\xi}} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'}\, e^{i\mathbf{G}'\cdot\boldsymbol{\xi}}. \tag{5.26}$$

The sum over $\mathbf{G}'$ is carried out by a single FFT, but has already been evaluated as the second term of Eq. (5.19). Multiplication by $n^2(\boldsymbol{\xi})$ is then carried out pointwise and the integral is performed by an inverse FFT. Again, this can be combined into the second term of Eq. (5.22) so no additional FFTs are needed.

5.4 Adaptive coordinate generation

The description of the fixed-frequency plane-wave method in GCCs in the preceding sections has been given for arbitrary coordinate transformations. In this section a general way to create mappings based on a given dielectric function is presented.

No approximations are made in Sec. 5.2. It is therefore possible to choose any convenient mapping between $\mathbf{x}$ and $\boldsymbol{\xi}$ and the results obtained by the plane-wave method in GCCs will be exact (subject to plane-wave convergence).

The original method used by Gygi (and later developed by other authors) considered the coordinate transformation as a variational parameter in the plane-wave calculation [96,97,99–101]. This creates a total energy functional that depends on the coordinate mapping as well as the plane-wave coefficients, and by performing

a minimisation with respect to both sets of coefficients the convergence of the plane-wave method is enhanced. Similar methods are popular in the mathematics community to study systems where large solution variations occur over spatial scales that evolve rapidly in time, particularly 'blowup' solutions of nonlinear partial differential equations; these methods are generally known as 'moving mesh' methods [102, 103].

A different, simpler, approach has also been used, which involves choosing a fixed coordinate transformation before beginning the plane-wave variational minimisation [104–106]. In Ref. 104, for example, the following transformation was used:

$$\boldsymbol{\xi} = \mathbf{x} + \sum_{\alpha} (\mathbf{x} - \mathbf{X}_{\alpha}) f_{\alpha} \left(|\mathbf{x} - \mathbf{X}_{\alpha}| \right), \tag{5.27}$$

where $\mathbf{X}_{\alpha}$ is the position of atom α and the sum is over all atoms in the unit cell. The function $f_{\alpha}(r)$ is a positive and rapidly-decaying function chosen to give the desired enhancement in spatial resolution at atom α. The principal disadvantage of this method is that Eq. (5.27) provides $\boldsymbol{\xi}(\mathbf{x})$ (because the positions of atoms are known and fixed only in $\mathbf{x}$ coordinates), but ultimately we require instead $\mathbf{x}(\boldsymbol{\xi})$ so that the $\mathbf{x}$-coordinates of the sampling points on the uniform $\boldsymbol{\xi}$-space grid may be found. This requires inverting the coordinate transformation.

Fixing the coordinate transformation seems to be inherently more suitable for PCF, where there is no explicit time-dependence in the governing differential equation and it is reasonably obvious in advance which regions of a structure are likely to benefit from a greater spatial resolution. However, the method for electronic structure calculations described above increases the sampling density at discrete points in space (i.e. around atoms) rather than at interfaces. Instead, a method that can create transformations that increase the sampling density in the direction perpendicular to dielectric interfaces is needed. The 'fictitious energy' method described in the next section has been developed as a means to do this.

5.4.1 Fictitious energy method

One method to generate an adaptive coordinate transformation is to define and then minimise a 'fictitious energy' functional with respect to the transformation, including terms in the energy that favour a transformation of the form required. In order to increase the number of grid points at the dielectric interfaces, we

define an 'interface energy' E_{I} as

$$E_{\mathrm{I}}[\mathbf{x}(\boldsymbol{\xi})] = -\int \mathrm{d}^2\xi \, \left|\nabla_{\mathrm{t}} n^2(\mathbf{x}(\boldsymbol{\xi}))\right|. \tag{5.28}$$

Mapping the uniformly-spaced $\boldsymbol{\xi}$-space grid points onto regions of $\mathbf{x}$-space containing large $|\nabla n^2|$ (i.e. dielectric interfaces) reduces E_{I}. However, if this term were to be used alone to define the energy, the lowest-energy configuration would be with *all* grid points mapped onto the interfaces. In general such a transformation would not be invertible, and would be likely to lead to poor plane-wave convergence because the regions of the structure not at dielectric interfaces would be poorly represented. It is therefore useful to add an additional term to the fictitious energy that increases the energy whenever the grid points are moved closer together or further apart. One method to do this, as implemented by Gygi (though for control of grid distortion during a calculation, rather than in static grid generation), is to define an 'elastic' energy E_{E} by analogy with elasticity theory [97].

The coordinate transformation between $\boldsymbol{\xi}$ and $\mathbf{x}$ can be identified as a 'material' distortion if the grid points are considered as particles in a solid. The uniformly-spaced points in $\boldsymbol{\xi}$-space represent the undeformed solid, and the non-uniformly-spaced particles in $\mathbf{x}$ space represent a deformation. This deformation can be described as a strain, and in elasticity theory it is usual to define the 'Cauchy deformation tensor' for the strain as [107]:

$$C^{pq} = \sum_k \frac{\partial \xi^p}{\partial x^k} \frac{\partial \xi^q}{\partial x^k}, \tag{5.29}$$

which is identical to the definition of the inverse metric tensor g^{pq} (see Eq. (B.13)). Because any energy associated with this strain must be invariant under rotation and translation, it may depend only on the principal invariants of the deformation tensor [107, 108]. Of these, Gygi considered $\det g^{ij}$, which is associated with isotropic compression, and $\mathrm{tr}\, g^{ij}$, which is associated with shear [97]. In practice we find that the compression term alone is sufficient to control the deformation, and we define the 'elastic' energy to be:

$$E_{\mathrm{E}}[\mathbf{x}(\boldsymbol{\xi})] = \mu \int \mathrm{d}^2\xi \, \det g^{pq}, \tag{5.30}$$

where μ is a parameter chosen to determine the relative influence of the 'elastic'

energy and the 'interface' energy defined by Eq. (5.28). This association between $\det g^{ij}$ and compression is reasonable given that $\mathrm{d}^2x = g^{1/2}\mathrm{d}^2\xi$; the quantity $\det g^{ij} = 1/g$ is related to the local change in the area element when the grid is distorted. In Appendix C it is shown that the elastic energy as defined is stationary when the grid is undistorted (i.e. $\boldsymbol{\xi} \equiv \mathbf{x}$), and we find numerically that the stationary point is a minimum.

The introduction of an 'elastic' energy has two useful effects. Firstly, it gives control over the level of distortion and prevents grid points from becoming excessively close at the dielectric interfaces relative to elsewhere. Secondly, because $\det g^{pq} = 1/g$, it becomes increasingly unfavourable to reduce g as $g \to 0$ and this guarantees that the transformation remains invertible. This is discussed further in Sec. 5.4.2 below.

The total fictitious energy associated with the coordinate transformation is given by:

$$E[\mathbf{x}(\boldsymbol{\xi})] = E_{\mathrm{I}} + E_{\mathrm{E}} \tag{5.31a}$$

$$= \int \mathrm{d}^2\xi \, \left(- \left| \nabla_{\mathrm{t}} n^2(\mathbf{x}(\boldsymbol{\xi})) \right| + \mu \det g^{pq} \right). \tag{5.31b}$$

Note, however, that this choice for the fictitious energy is not unique. Any choice of E_{I} that gives an increase in energy when grid points are moved closer to dielectric interfaces would be suitable, and any function of any of the principal invariants of g^{ij} (provided that it is minimised by an undistorted grid and increases with distortion) could form a suitable 'elastic' energy. However, we have achieved good results using Eq. (5.31b), as will be discussed in Chapter 6.

Methods to create the coordinate transformation by minimising the fictitious energy are discussed in the next section.

5.4.2 Implementation of the fictitious energy method

The transformation between $\boldsymbol{\xi}$ and $\mathbf{x}$ is defined by Eq. (5.1) together with a finite circular cutoff for the coefficients such that $|\mathbf{G}| < G^{\mathrm{x}}_{\mathrm{max}}$. This cutoff is chosen to be as small as possible whilst giving grids with the required characteristics, and is discussed further in Sec. 6.1.1. If N_{x} $\mathbf{G}$-vectors satisfy the cutoff condition,

then locating the minimum in the fictitious energy becomes an $2N_x$-dimensional minimisation problem[6].

Many methods exist to solve minimisation problems, and in general the choice of method is constrained by the available information: if partial derivatives of the function to be minimised (i.e. its gradient) are available and can be calculated quickly, then usually a method that uses this information will be faster and more efficient than one that uses only evaluations of the function without derivatives [76][7]. In the case of the fictitious energy method presented here, the calculation of derivatives can be carried out easily using FFTs (see below) and we therefore choose a method that uses them.

The most efficient minimisation methods that use derivative information fall into two categories: conjugate gradient methods and variable metric methods. Conjugate gradient methods are common in plane-wave electronic structure calculations [72] and are also used by the fixed-wavevector plane-wave method in photonics (see Ref. 68 and Sec. 3.2.1). They are especially attractive when the number of dimensions is large because the required storage for an N-dimensional minimisation is $O(N)$. Variable metric methods instead require $O(N^2)$ storage but in some cases are more efficient [76]. We have successfully implemented the Broyden-Fletcher-Goldfarb-Shanno (BFGS) variable metric algorithm.

The BFGS algorithm takes a starting vector of $2N_x$ variables (the $x^p_{\mathbf{G}}$ coefficients) and makes 'downhill' steps in order to reach the nearest local minimum of E. In practice, the local minimum generally corresponds to the desired coordinate transformation provided a sufficiently large smoothing (as discussed below) is used. An obvious choice of starting vector is $\mathbf{x_G} = \mathbf{0}$ because this corresponds to an invertible transformation (with $g = 1$ everywhere). The form of the fictitious energy functional with its $1/g$ 'penalty' term (the elastic energy) ensures that small downhill steps away from an invertible transformation will create a new transformation that remains invertible.

The BFGS minimiser calculates new downhill steps by providing the vector of

[6]The factor of 2 here is a result of having $x^1_{\mathbf{G}}$ and $x^2_{\mathbf{G}}$ coefficients.

[7]Typically, an N-dimensional minimisation using a method utilising gradients must perform $O(N)$ one-dimensional minimisations, whereas a method that does not use gradients must perform $O(N^2)$ minimisations. This is especially important when N is large, as discussed in Sec. 6.2.1.

$x^p_{\mathbf{G}}$ components (the current estimate of the location of the minimum) and then taking as input the fictitious energy (a single scalar quantity) and its gradient, which is a vector of length $2N_{\mathrm{x}}$ whose components are given by

$$\frac{\partial E}{\partial x^p_{\mathbf{G}}} = \frac{\partial}{\partial x^p_{\mathbf{G}}} \int \mathrm{d}^2\xi \, \left(- \left| \nabla_{\mathrm{t}} n^2 \right| + \mu \det g^{pq} \right) . \tag{5.32}$$

In Appendix C a derivation is given of the following expression, which can be used in practice to obtain the components of the gradient:

$$\frac{\partial E}{\partial x^p_{\mathbf{G}}} = - \int \mathrm{d}^2\xi \left(\frac{\partial}{\partial x^p} \left| \nabla_{\mathrm{t}} n^2 \right| + \frac{2i\mu}{g} \sum_{r,q=1}^{2} g^{qr} \frac{\partial x^p}{\partial \xi^q} G_r \right) e^{i\mathbf{G}\cdot\boldsymbol{\xi}} . \tag{5.33}$$

The next sections describe the way in which E and its gradient, and the fixed quantities needed in order to determine them, are calculated.

Calculation of dielectric function and related quantities

Quantities relating to n^2 in Eqs. (5.31b) and (5.33) are fixed and can be calculated once and stored *on a uniform* $\mathbf{x}$*-space grid* (as in the conventional plane-wave method) before the minimisation begins[8]. First, n^2 is sampled on a fine real-space FFT grid and transformed to reciprocal space. This FFT grid is chosen to be larger than that used in the minimisation itself, to reduce errors introduced by interpolation (see below). The dielectric function is then smoothed with a smoothing FWHM of Γ^{x} by multiplication by a Gaussian envelope in the same way as described in Sec. 3.2.3 for the plane-wave method. However, the smoothing used here is not related to the smoothing used in the plane-wave eigenmode calculation. It is instead applied to the dielectric function for two reasons. Firstly, it ensures that there is a gradual transition of the coordinate transformation between regions of the structure with interfaces and the remainder of the unit cell. Secondly, smoothing the sharp features of $|\nabla_{\mathrm{t}} n^2|$ allows faster convergence of the minimisation and prevents the minimiser from locating spurious local minima in the sharp features. The smoothing widths Γ used for accurate mode-solving are typically much smaller than those used here for adaptive grid generation.

Having performed the smoothing in reciprocal space, it is then possible to calculate the reciprocal-space components of $\frac{\partial n^2}{\partial x^p}$ by multiplication by iG_p. Two FFTs

[8]Note that it is not possible to use a $\boldsymbol{\xi}$-space grid until the coordinate transformation between $\mathbf{x}$ and $\boldsymbol{\xi}$ is defined.

transform these quantities to real space, where $|\nabla_t n^2|$ is calculated as follows:

$$|\nabla_t n^2|(\mathbf{x}) = \left\{ \left(\sum_{\mathbf{G}} iG_1 n_{\mathbf{G}}^2 e^{i\mathbf{G}\cdot\mathbf{x}} \right)^2 + \left(\sum_{\mathbf{G}} iG_2 n_{\mathbf{G}}^2 e^{i\mathbf{G}\cdot\mathbf{x}} \right)^2 \right\}^{1/2}. \qquad (5.34)$$

The gradient $|\nabla_t n^2|$ is then stored on this fine $\mathbf{x}$-space grid for use in the minimisation as described below. As it is not a function of the coordinate transformation and depends only on the original dielectric structure, it can be calculated only once and then stored before the minimisation is carried out.

Calculation of fictitious energy

The fictitious energy is required by the BFGS algorithm and can be calculated from Eq. (5.31b) for a given set of transformation coefficients $\mathbf{x}_{\mathbf{G}}$. First, Eq. (5.1) is used to determine the real-space transformation:

$$x^p(\boldsymbol{\xi}) = \xi^p + \sum_{\mathbf{G}} x_{\mathbf{G}}^p \exp(i\mathbf{G}\cdot\boldsymbol{\xi}). \qquad (5.35)$$

The quantities $(x^p - \xi^p)$ are evaluated by taking one FFT for each of the vector components of the $\mathbf{x}_{\mathbf{G}}$ (i.e. two FFTs). The size of the FFT grid for these and subsequent FFTs used in grid generation must be large enough to contain all of the components satisfying $|\mathbf{G}| < G_{\max}^{\mathrm{x}}$, and also to give enough real-space points to sample $|\nabla_t n^2|$ accurately (i.e. of size $N \gtrsim 1/\Gamma^{\mathrm{x}}$). Because $G_{\max}^{\mathrm{x}} < G_{\max}$ and $\Gamma^{\mathrm{x}} > \Gamma$ (the quantities without the superscript x being those used in the eigenmode calculation), it is always sufficient to use an FFT grid of the same size as that used in eigenmode determination.

For every point on the $\boldsymbol{\xi}$-space grid, the corresponding $\mathbf{x}$-coordinate is determined by FFT as described. In general, this does not correspond to a grid point on the $\mathbf{x}$ grid, and therefore $|\nabla_t n^2|$ must be determined by interpolation. Any given point on the $\mathbf{x}$ grid is bounded by a triangle whose three corners are the nearest points at which $|\nabla_t n^2|$ is stored, and linear interpolation is used to provide the unknown value rapidly. Note that in the course of carrying out linear interpolation, the gradient $\frac{\partial|\nabla_t n^2|}{\partial x^p}$ is calculated and this can be stored for use in calculating the gradient functions as described in the next section.

Calculation of the metric tensor g_{pq} and its determinant has been described in

Sec. 5.3.1, and g is stored on a $\boldsymbol{\xi}$-space grid. To evaluate the 'elastic' component of the fictitious energy (the second term of Eq. (5.31b)), $1/g$ is summed pointwise over the entire grid and multiplied by μ.

Calculation of gradient

From the set of coefficients $x^p_{\mathbf{G}}$, the mapping $\mathbf{x}(\boldsymbol{\xi})$ can be obtained as described above. Using this, for every point on the $\boldsymbol{\xi}$-space grid, the corresponding $\mathbf{x}$ point is known and the value of $\frac{\partial}{\partial x^p}|\nabla_t n^2|$ is also known because it was calculated previously as part of the linear interpolation for $|\nabla_t n^2|$. The entire set of derivatives of E_I is then given using Eq. (5.33) by a single inverse FFT to reciprocal space for each of $p = 1, 2$:

$$\frac{\partial E_\mathrm{I}}{\partial x^p_{\mathbf{G}}} = -\int \mathrm{d}^2\xi \left(\frac{\partial}{\partial x^p}|\nabla_t n^2|\right) e^{i\mathbf{G}\cdot\boldsymbol{\xi}}. \tag{5.36}$$

The gradient of the elastic energy can also be calculated by FFT. Sec. 5.3.1 describes how the derivatives $\frac{\partial x^p}{\partial \xi^q}$ and the determinant g are obtained. On the real-space $\boldsymbol{\xi}$ grid the following quantity is calculated:

$$Z^{pr}(\boldsymbol{\xi}) = \frac{1}{g}\sum_q g^{qr}\frac{\partial x^p}{\partial \xi^q}, \tag{5.37}$$

and the gradient of the elastic energy is then given by

$$\frac{\partial E_\mathrm{E}}{\partial x^p_{\mathbf{G}}} = -2\mu i \sum_r G_r \int \mathrm{d}^2\xi\, Z^{pr} e^{i\mathbf{G}\cdot\boldsymbol{\xi}}, \tag{5.38}$$

where the integrals can be computed by four FFTs (for $p, r = 1, 2$) and the multiplication by G_r is carried out in reciprocal space. As with the interface energy, the FFTs provide all components of $\frac{\partial E_\mathrm{E}}{\partial x^p_{\mathbf{G}}}$ simultaneously.

All of the information required by the BFGS minimisation algorithm is contained within Sec. 5.4.2. In order to generate adaptive grids using the fictitious energy method, it remains only to choose $G^{\mathbf{x}}_{\max}$ to set the number of $\mathbf{x}_{\mathbf{G}}$ coefficients that are allowed to vary. The choice of $G^{\mathbf{x}}_{\max}$ is discussed with reference to example structures in Chapter 6.

5.4.3 Characterisation of adaptive grids

It was shown in Chapter 4 that plane-wave convergence can be expected only when the number of sampling points at dielectric interfaces is $N_\mathrm{I} \gtrsim 1$. Extending this 'convergence in principle' argument to 2D, we can expect plane-wave convergence only when the number of sampling points in the direction perpendicular to dielectric interfaces is $N_\mathrm{I} \gtrsim 1$. It is therefore useful to know how N_I is changed by a particular coordinate transformation.

The method used by Gygi to define a local (position-dependent) effective plane-wave cutoff was to use the Jacobian for the transformation [97]. In terms of the number of real-space sampling points, we can use the transformation expression $\mathrm{d}^2x = g^{1/2}\mathrm{d}^2\xi$ to determine that the local density of sampling points per unit area on the $\mathbf{x}$ grid is increased by a factor of $g^{-1/2}(\mathbf{x})$ relative to a uniform grid with the same number of sampling points.

The use of g to determine a local effective cutoff was suitable for the electronic structure work of Gygi, but in the case of adaptive grids for PCF it is necessary instead to find the increase in the number of sampling points in a given direction (i.e. perpendicular to dielectric interfaces). This can be done by treating the coordinate transformation as a spatial deformation of particles in a solid, as described in Sec. 5.4.1, and using established results from the field of continuum mechanics.

The ratio of lengths of a distorted and undistorted line element along a unit vector $\hat{\mathbf{n}}(\mathbf{x}) = (n^1, n^2)$ is a quantity known as the 'stretch ratio' $\lambda_{\hat{\mathbf{n}}}$, and is determined from the Cauchy deformation tensor (defined by Eq. (5.29) and identical to g^{ij}) by [107]:

$$\frac{1}{\lambda_{\hat{\mathbf{n}}}^2} = \sum_{p,q=1}^{2} C^{pq} n^p n^q. \tag{5.39}$$

The quantity of interest when considering plane-wave convergence is the spacing of grid points in a specific direction $\hat{\mathbf{n}}$, which is related to the stretch ratio in that direction by $\Delta x_{\hat{\mathbf{n}}} = \lambda_{\hat{\mathbf{n}}} \Delta x^0_{\hat{\mathbf{n}}}$, where $\Delta x^0_{\hat{\mathbf{n}}}$ is the undistorted spacing of grid points in the same direction. In particular, we require the spacing of grid points in the direction perpendicular to the dielectric interfaces. The perpendicular unit

vector is given by

$$\hat{n}_{\perp}^{p} = \frac{1}{|\nabla_t n^2|} \frac{\partial n^2}{\partial x^p}, \tag{5.40}$$

and therefore the spacing of grid points in that direction is given by

$$\Delta x_{\perp} = \Delta x_{\perp}^{0} \left(\frac{1}{|\nabla_t n^2|^2} \sum_{p,q=1}^{2} g^{pq} \frac{\partial n^2}{\partial x^p} \frac{\partial n^2}{\partial x^q} \right). \tag{5.41}$$

Because $\frac{\partial n^2}{\partial x^q}$ and $|\nabla_t n^2|$ are both calculated as part of the minimisation process (see Eq. (5.34)), this quantity can be easily obtained for the final calculated distortion and used to characterise the grid. In Chapter 6, grids are shown and characterised by the quantity $N_{\mathrm{I}} = N_{\mathrm{I}}^{0}/\Delta x_{\perp}$, which relates N_{I} to its value N_{I}^{0} when sampled on an undistorted grid.

In general, N_{I} varies both parallel and perpendicular to any interface and it is necessary to take an average of N_{I} including only sampling points in the unit cell that are considered to be at an interface. The simplest way to do this is to define a point as being at an interface if $|\nabla_t n^2| > \alpha$ for some tolerance α. The maximum value of $|\nabla_t n^2|$ is $\approx 2[n^2(\mathrm{max}) - n^2(\mathrm{min})]/\Gamma^{\mathrm{x}}$ for a sharp interface between $n^2(\mathrm{max})$ and $n^2(\mathrm{min})$. We arbitrarily take α to be 5% of this maximum. However, it should be noted that the calculated N_{I} does depend on this tolerance. Although they provide a useful way to compare different grids, the values presented in Chapter 6 can only be considered approximate[9].

5.5 Conclusions

The results of Chapter 4 suggest that the convergence of the fixed-frequency plane-wave method is dependent on the number of real-space sampling points in the regions of a structure corresponding to the sharp interfaces in the dielectric function. It therefore follows that any method that can increase the number of sampling points at interfaces should improve plane-wave convergence. However, in the conventional formulation of the plane-wave method, this can be achieved only by either increasing the level of smoothing applied to the dielectric function (which perturbs β values), or increasing the size of FFT grids (which requires

[9] A more accurate value of N_{I}, though still approximate, would be found by using Γ rather than Γ^{x} in the definition of α. However, this introduces an additional complication because $\Delta x_{\perp}$ is then dependent on the eigensolver smoothing width rather than being an independent property of the adaptive coordinate grid.

more memory and time to perform the calculation). In this chapter, a method has been developed to give *position-dependent* control over the density of sampling points, and therefore allow the number of grid points to be increased at dielectric interfaces without increasing either the smoothing width or the FFT grid size.

The reformulation of the fixed-frequency plane-wave method presented in this chapter shares some similarities with the conventional method outlined in Chapter 3. Like the conventional method, it places no restrictions on the geometry of the repeating unit cell; it can handle material dispersion and complex dielectric constants; it uses an iterative eigensolver and iterative linear solver in combination to extract interior eigenvalues efficiently; and—most importantly—it also uses FFTs to perform matrix-vector operations and is therefore capable of performing calculations rapidly.

Also presented in this chapter is a method for the generation of adaptive grids for use with the plane-wave method in curvilinear coordinates. This novel method of grid generation based on minimising a fictitious energy functional creates a system of curvilinear coordinates specifically for use with a given dielectric function, and therefore allows the selective increase in the sampling density of grid points at dielectric interfaces that should enhance the convergence of the plane-wave method. A method of characterising adaptive grids and determining the increase in the number of grid points at dielectric interfaces has also been described.

The following chapter demonstrates the use of the fictitious energy method for grid generation both for single unit cells of PCF cladding and also for supercells. The enhancement of plane-wave convergence brought about by using the plane-wave method in GCCs is then shown.

Chapter 6

Applications of the plane-wave method in curvilinear coordinates

The previous chapter contained a description of the reformulation of the fixed-frequency plane-wave method in generalised curvilinear coordinates, and presented a method for the generation of adaptive grids based on the idea of a fictitious energy functional. In this chapter, results are first presented demonstrating the generation of grids adapted for PCF cladding structures, i.e. single unit cells. The enhanced convergence behaviour of the plane-wave method in GCCs when using these grids is then shown and compared to that of the conventional formulation of the fixed-frequency plane-wave method. The generation of adaptive grids for supercell calculations is then demonstrated using a realistic hollow-core PCF structure as an example. The performances of the GCC and conventional plane-wave methods are also compared for this structure. The problem of increased computational effort arising when using the GCC method to solve for the eigenmodes of supercells is then discussed.

6.1 Application to PCF cladding structure

In order to test the plane-wave method in GCCs and demonstrate its associated improvement relative to the conventional formulation of the plane-wave method, it is convenient to study a PCF cladding structure for which exact numerical results are readily available. We therefore choose to use an array of circular holes

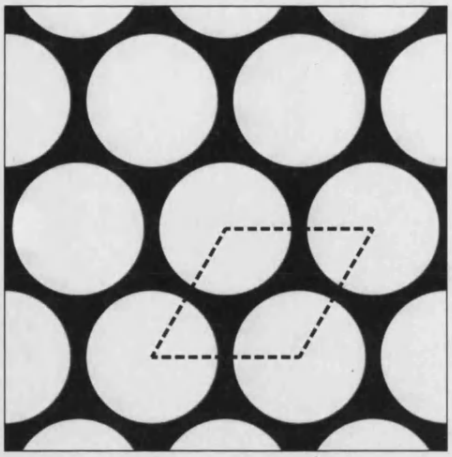


Figure 6.1: PCF cladding structure of a triangular array of circular air holes (shown in white) in glass (black). The circles are of radius 0.45Λ where Λ is the pitch of the lattice. The unit cell is marked by dashed lines.

in glass, which forms a semi-realistic structure for a PCF cladding but has modes that can be obtained easily using an exact method.

The test system used in this section is a triangular array of circular air holes with refractive index $n = 1$ embedded in a glass of $n = 1.5$. The air holes have a radius of 0.45Λ, where Λ is the lattice pitch; this corresponds to an air-filling fraction of 73%, and creates a perfectly periodic PCF cladding structure. A typical value of $k_0\Lambda = 10$ is chosen, together with (for simplicity) Bloch wavevector $\mathbf{k} = \mathbf{0}$. The PCF cladding structure is shown in Fig. 6.1.

The following sections describe the steps taken to use the plane-wave method in GCCs to determine the eigenmodes of the test structure, beginning with the generation of a suitable adaptive grid.

6.1.1 Adaptive grid generation

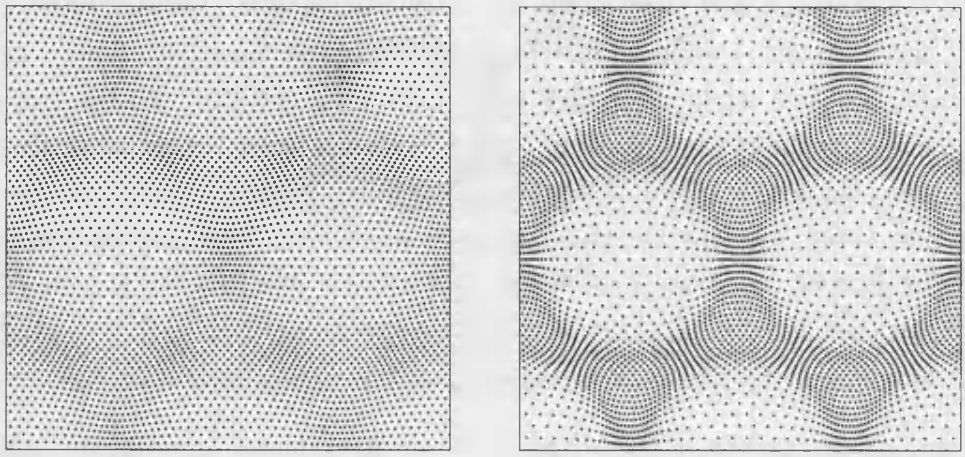
In order to generate an adaptive grid using the method of Sec. 5.4.1, it is first necessary to choose a level of Gaussian smoothing for the dielectric structure. Note that this is *only* for the grid generation; it is independent of that used in the eigenmode determination, as discussed in Sec. 5.4.2. The resulting grid tends to be relatively independent of the choice of smoothing width Γ^x, but increasing it tends to increase the rate of convergence of the BFGS minimisation.

It is also necessary to choose an appropriate cutoff $G^{\mathbf{x}}_{\max}$ for the reciprocal-space components of the coordinate transformation $\mathbf{x}_{\mathbf{G}}$. Increasing this cutoff generally lowers the minimum of the fictitious energy because it increases the number of degrees of freedom in the coordinate transformation, allowing the grid to adapt more freely. However, it also increases the amount of effort required to perform the minimisation. To reduce this effort as much as possible, we choose to enforce the full P6mm symmetry of the dielectric function. The components of $\mathbf{x}_{\mathbf{G}}$ are thus constrained such that of the total of 19 $x^1_{\mathbf{G}}$ components and 19 $x^2_{\mathbf{G}}$ components of the first three reciprocal-space 'stars', there are only 3 independent variables. Sec. D.2 of Appendix D gives details of this procedure. For this structure, 3 independent variables are sufficient to obtain a good adaptive grid; in general, calculation of N_{I} (by the method described in Sec. 5.4.3) together with visual inspection of the sampling grid using plots such as those of Fig. 6.2 is sufficient to determine whether enough variables have been used.

The smoothing width used to create $|\nabla_{\mathrm{t}} n^2|$ was $\Gamma^{\mathbf{x}} = 0.03\Lambda$, and the fine grid of $|\nabla_{\mathrm{t}} n^2|$ (in undistorted $\mathbf{x}$-space) was of dimension 512×512. The FFT grid size used for the grid generation presented in this section was 128×128. The generated grids are relatively independent of all of these quantities provided the conditions on the FFT grid size in Sec. 5.4.2 are met. Convergence of the BFGS minimiser tends to be enhanced by using large fine grids for $|\nabla_{\mathrm{t}} n^2|$, but this is ultimately limited by available storage and the time required to perform the FFTs on large grids.

Sec. 5.4.1 described the use of an 'elastic' energy to control the level of distortion of a coordinate transformation. The grids plotted in Fig. 6.2 show the effect of the compression energy parameter μ on the grids produced by the fictitious energy method. Each grid typically required ≈ 10 BFGS iterations (i.e. ≈ 10 one-dimensional line minimisations) to locate the fictitious energy minimum; the total time taken was approximately 1 second on a desktop PC. Also given is the enhancement (relative to an undistorted grid, of the type used in the conventional formulation of the plane-wave method) of the number of grid points in the direction perpendicular to the dielectric interfaces, calculated using the method described in Sec. 5.4.3.

Fig. 6.2 shows clearly the effect on the resulting grid of changing μ. The larger the value of μ, the larger the cost in elastic energy of distorting the grid: as



(a) $\mu = 0.4$, $N_\mathrm{I} \approx 1.3 N_\mathrm{I}^0$

(b) $\mu = 0.3$, $N_\mathrm{I} \approx 2.3 N_\mathrm{I}^0$

(c) $\mu = 0.2$, $N_\mathrm{I} \approx 3.7 N_\mathrm{I}^0$

(d) $\mu = 0.1$, $N_\mathrm{I} \approx 4.7 N_\mathrm{I}^0$

Figure 6.2: Grids of real-space sampling points generated for an array of circular air holes in glass (see Fig. 6.1) smoothed with FWHM $\Gamma^\mathrm{x} = 0.03\Lambda$. The grids shown were calculated on an FFT grid of size 128×128, with every fourth point plotted to give a grid of 32×32. The approximate number of points in the dielectric interface regions perpendicular to the interfaces is given relative to that of the undistorted grid, N_I^0.

$\mu \to \infty$, the amount of grid distortion must tend to zero. Smaller values of μ allow the grid to distort more, and as μ is decreased the increase in the number of grid points at the dielectric interfaces is visible.

Another effect visible in Fig. 6.2 is an increase in the number of grid points in the regions of the structure containing glass. The fictitious energy functional does not explicitly favour increasing the number of points in the glass regions, because the only term involving the dielectric function in the fictitious energy is the interface term, a function of $|\nabla_t n^2|$ and not n^2 explicitly. Instead, the observed increase of grid points in glass is a consequence of the trade-off between interface energy and elastic energy. If the grid were to have a similar density of grid points in glass and air, the observed results suggest that an increase in the density of points at interfaces must cause a large increase in elastic energy.

6.1.2 Dielectric smoothing in the GCC method

Once a suitable coordinate grid for a calculation has been created, it is necessary to generate the structure-dependent fixed quantities that are needed by the mode-solving calculation. These are n^2, $\ln n^2$ and $\frac{\ln n^2}{\partial x^k}$; the method of generating them was discussed in Sec. 5.3.2.

As in the conventional plane-wave method, the generation of structure-dependent fixed quantities in the GCC method involves application of smoothing. Details of how this is done were given in Sec. 5.3.2 but it is useful here to show the difference between exact Gaussian smoothing and the smoothing used in the GCC method with reference to the test structure. An interface smoothed with each method is shown for comparison in Fig. 6.3; the best match between the two smoothings has been determined by eye.

In general, it is desirable to perform calculations with a level of smoothing chosen such that it introduces negligible error into the results relative to an unsmoothed interface. Provided the width of smoothing can be controlled, it should always be possible to choose a width sufficiently narrow to obtain useful results. In this limit, when the smoothing error is small, the specific functional form of dielectric interfaces is unimportant. However, the inability to perform exact Gaussian smoothing makes it more difficult to make an exact comparison between the results of the conventional plane-wave method and those of the method in GCCs.

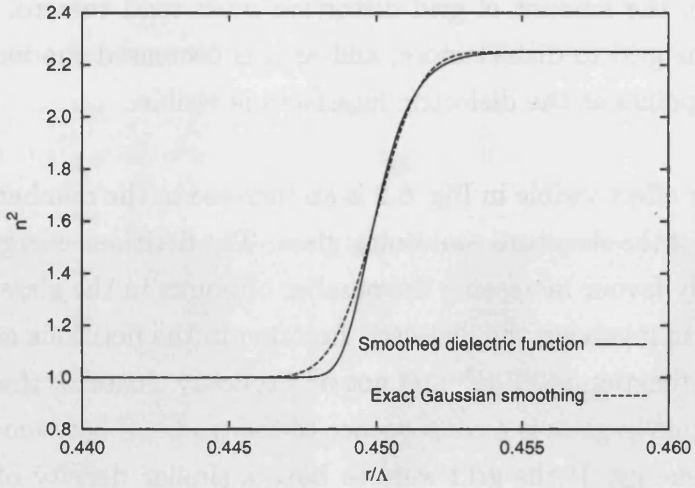


Figure 6.3: Smoothed dielectric function of the test system described in Sec. 6.1 plotted in a direction normal to the dielectric interface. The radius $r = |\mathbf{x}|$ is measured from the centre of the hole along the line from $\mathbf{x} = \mathbf{0}$ to $\mathbf{x} = \mathbf{R}_1 + \mathbf{R}_2$.

The calculations presented in the remainder of this chapter give a comparison between the conventional plane-wave method and the GCC method, and in these cases the smoothing has been matched as well as possible (as in Fig. 6.3). All smoothing widths for the GCC method are given as 'approximate Gaussian equivalent' widths.

6.1.3 Plane-wave calculations

The aim of the reformulation of the plane-wave method in GCCs is to provide an increase in the number of real-space sampling points at dielectric interfaces for a given FFT grid size. This should manifest itself as enhanced convergence behaviour, i.e., it should be possible to obtain converged results using smaller basis set sizes than with the conventional plane-wave method. In order to demonstrate this, we calculate the eight distinct highest-β modes of the test system previously described using the $\mu = 0.1$ coordinate grid (as shown in Fig. 6.2d), and observe the convergence of these modes with respect to the size of the FFT grid when a range of different smoothing widths are applied. In Fig. 6.4 we plot the calculated values of $\beta\Lambda$ as a function of the FFT grid size, using three different Gaussian smoothing widths and two smoothings (whose approximate equivalent Gaussian widths are given) for GCC calculations.

Fig. 6.4 shows that in general the β values of the GCC method converge signifi-

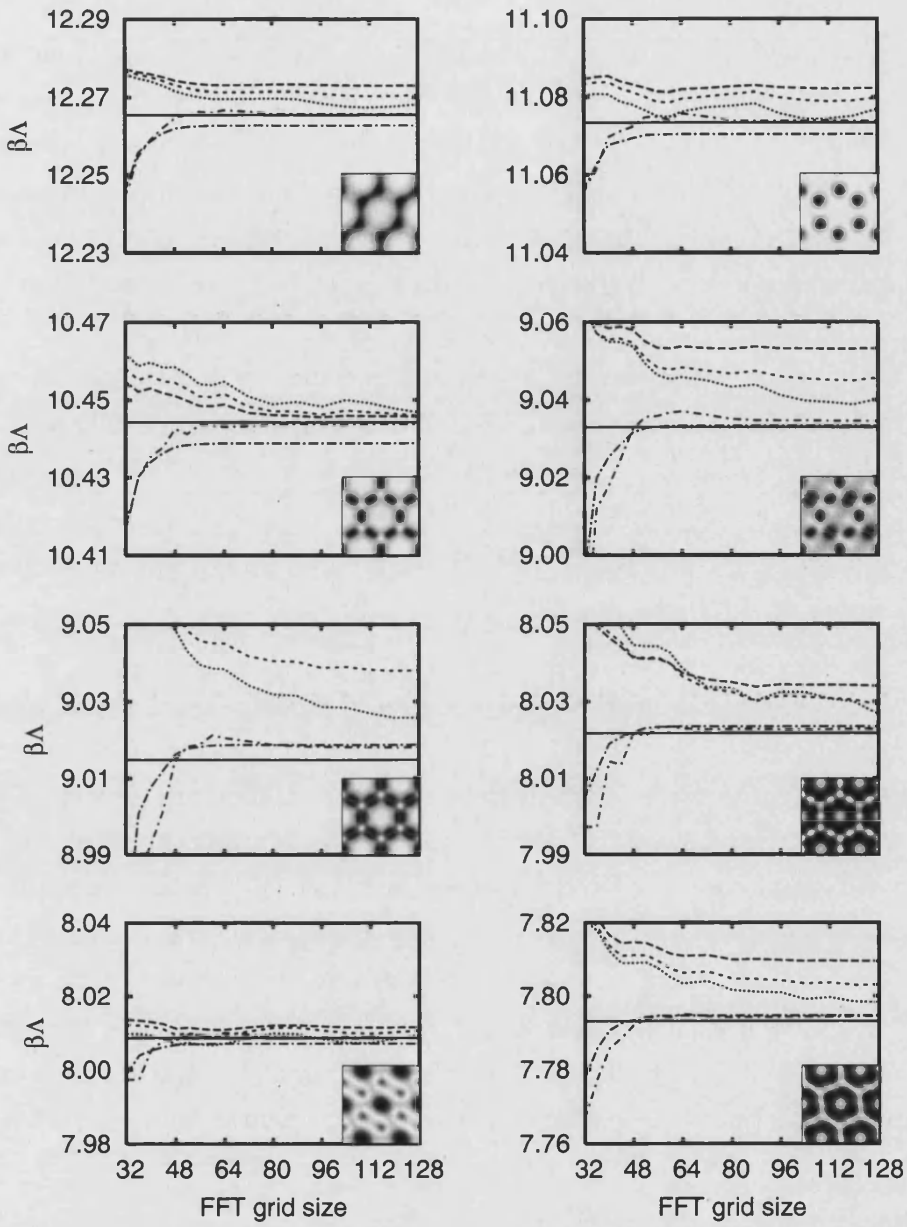


Figure 6.4: Convergence of the 8 highest-β modes of the test system of Sec. 6.1 as FFT grid size is increased. In each plot the solid horizontal line is the result from an exact (unsmoothed) method correct to 4 d.p. [109,110], the dashed lines show convergence of the conventional method, and the mixed dot/dash lines show the GCC results. The Gaussian smoothing widths for the conventional results are (in order of increasing dash length) $10^{-3}\Lambda$, $6 \times 10^{-3}\Lambda$ and $10^{-2}\Lambda$. The approximate equivalent widths for the GCC curves are (again in order of increasing dash length) $10^{-3}\Lambda$ and $2.5 \times 10^{-3}\Lambda$. Inset figures show Poynting vector magnitude (large in black regions). The modes at $\beta\Lambda = 12.2654$, 11.0734, 9.0329, 8.0086, and 7.7931 are degenerate pairs; only one of each pair is shown.

cantly more quickly with respect to the basis set size than those produced by the conventional method. It is not clear why the GCC β values should converge from below the correct value while those from the conventional plane-wave method converge from above, but this is not a general property: in 1D, for example, the β values for glass in the conventional method tend instead to converge from below (see, e.g., Fig. 4.8 on page 70), and metals tend to converge from above (see, e.g., Fig. 4.11 on page 73). Figs. 6.11-6.14 later in this chapter show GCC β values converging from either above or below depending on the level of grid distortion.

Comparison of the two curves representing the smallest smoothing widths shown in Fig. 6.4, both of which represent $\Gamma = 10^{-3}\Lambda$, makes the enhanced convergence behaviour of the GCC method particularly clear. For all of the modes shown, the GCC results are converged at an FFT size of ≈ 48, whereas for some modes the results of the conventional plane-wave method do not appear to have converged even at an FFT size of 128.

Because the GCC method converges more rapidly, it is possible to use narrower smoothing widths with the method while retaining plane-wave convergence. Consequently, the set of β values from the GCC method are consistently closer than those of the conventional method to the exact (unsmoothed) values, which have been calculated by a KKR-based method [109,110]. It is worthwhile to note that the systematic error introduced by smoothing, while remaining of the same order of magnitude, tends to vary considerably between modes. The most smoothed conventional-method curve for $\beta\Lambda = 9.0148$ is not visible on the scale shown: it reaches 9.0529 at FFT size 128. This uncertainty in the accuracy of results can be resolved only by using smaller smoothing widths throughout, demonstrating a major advantage of the GCC method.

The improvement in convergence behaviour due to the use of GCCs can be understood in terms of the 'convergence in principle' (CIP) arguments developed in Chapter 4, although it is important to note that these are intended to be qualitative and do not necessarily hold quantitatively in 2D. As the FFT grid size is increased, the density of real-space grid points increases; only when our CIP criterion of having at least one real-space sampling point in the interface region (i.e. $N_{\mathrm{I}} \gtrsim 1$) is met can we expect to see convergence of β values. Because the grid used for the GCC calculations had $N_{\mathrm{I}} \approx 4.7 N_{\mathrm{I}}^{0}$, convergence can be expected for these calculations at FFT grid sizes that are smaller than a similarly-sized

conventional grid by a factor of ≈ 5. This is consistent with the observation that the GCC results in Fig. 6.4 are converged but those of the conventional plane-wave method are not, although there is no exact quantitative agreement between the apparently-converged FFT grid sizes and 'convergence in principle'. Given the approximate nature of the CIP criterion (particularly when applied to 2D calculations), and the uncertainty inherent in defining N_{I} (as discussed in Sec. 5.4.3), this is not unexpected. However, the general principle remains valid and the improvement brought about by GCCs is clear from Fig. 6.4.

6.2 Application to supercell calculations

The methods developed in Chapter 5 (both the reformulation of the plane-wave method in GCCs and the 'fictitious energy' method of adaptive grid generation) are equally applicable to calculations involving supercells as to those for single unit cells. Exact results are less readily available, but it is still possible to make a direct comparison between the conventional formulation of the plane-wave method and the reformulation in GCCs.

The test system used to study the GCC method is a cladding consisting of a triangular array of rounded hexagonal air holes in a background of refractive index $n = 1.5$, with a 7-cell core chosen to be typical of air-silica hollow-core PCF. The details of the cladding structure are shown in Fig. 6.5, and those of the core are shown in Fig. 6.6. We use an 8×8 supercell as shown in Fig. 6.7. Although the calculations presented here are a demonstration of the GCC method, the structure used is realistic and relevant to real PCF structures. Fig. 1.1b on page 13 and Fig. 7.3a on page 130 show real air-silica hollow-core PCF structures for comparison.

The next sections describe the generation of a suitable adaptive grid for the supercell calculations and then compare the results obtained with those from the conventional plane-wave method.

6.2.1 Adaptive grid generation

The generation of adaptive grids for supercells does not differ in principle from that for single unit cells; the procedure described for single cells in Sec. 6.1.1 is

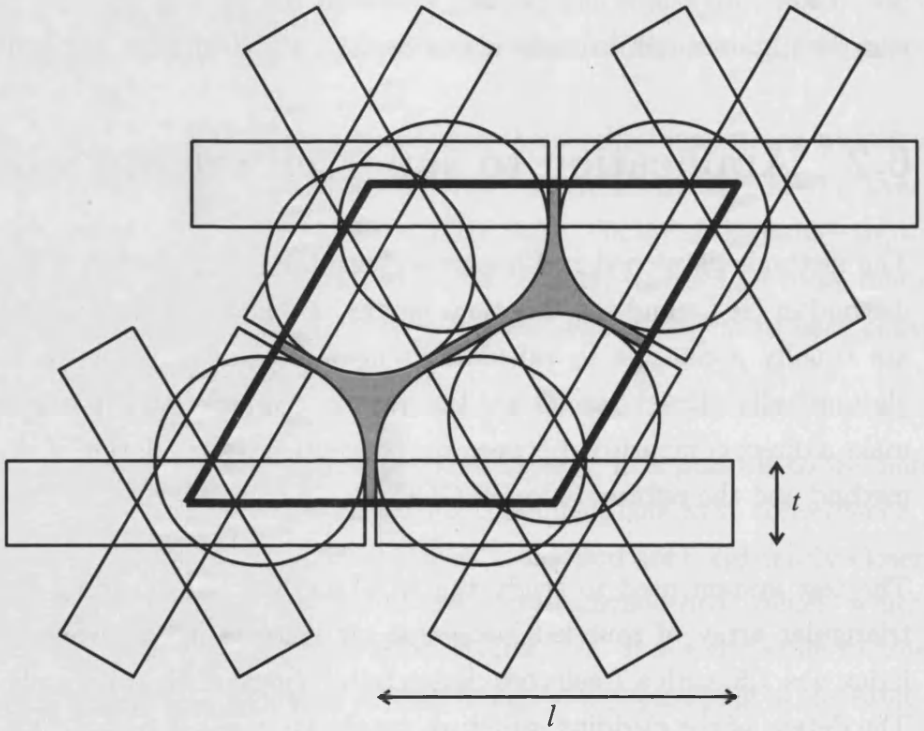


Figure 6.5: Cladding design for test system of air-silica PCF using circles and rectangles, similar to that of Ref. 111. Silica regions are shown in grey, and the unit cell (pitch Λ) is marked in a thicker line. The structure is constructed by starting with a silica unit cell and then placing the air rectangles, then adding the circles whose radii and centres are chosen to make the gradient continuous where they touch the rectangles. The adjustable parameters are the lengths l and t; these determine respectively the thickness of struts in the cladding (given by $(\Lambda - l)$) and the 'roundedness' of the hexagons. Decreasing t makes the structure more rounded. For the test system, $l = 0.97\Lambda$ (i.e. strut thickness 0.03Λ) and $t = 0.2309\Lambda$; the resulting structure has an air-filling fraction of 91%.

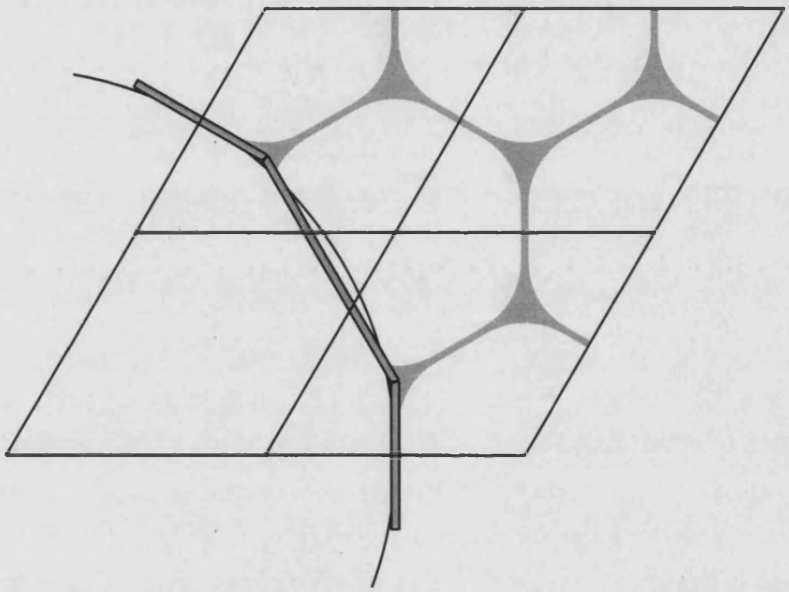


Figure 6.6: Core design for a test system of air-silica PCF using circles and hexagons. Only $\frac{1}{6}$ of the design is shown; the remainder can be obtained by applying rotational symmetry. An air circle of radius $r = 3\Lambda/2$ is first placed at the centre of the supercell comprising cladding unit cells (as marked), and the core wall is then constructed by superimposing three silica rectangles of the desired width. The core wall width shown (and used for the test structure) is 0.03Λ.

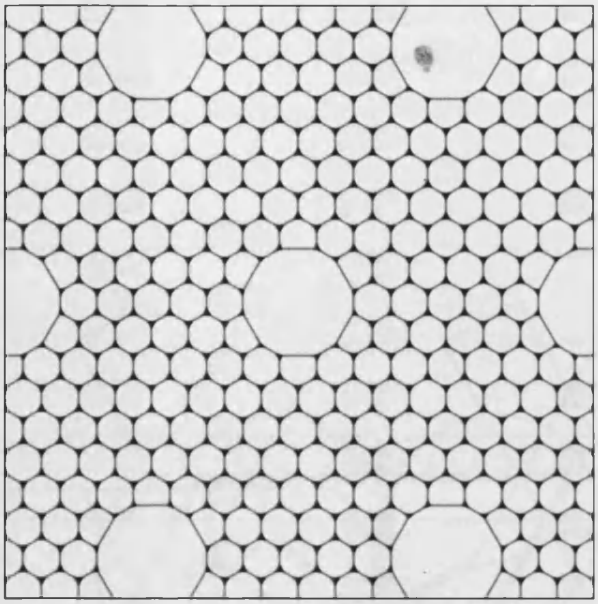
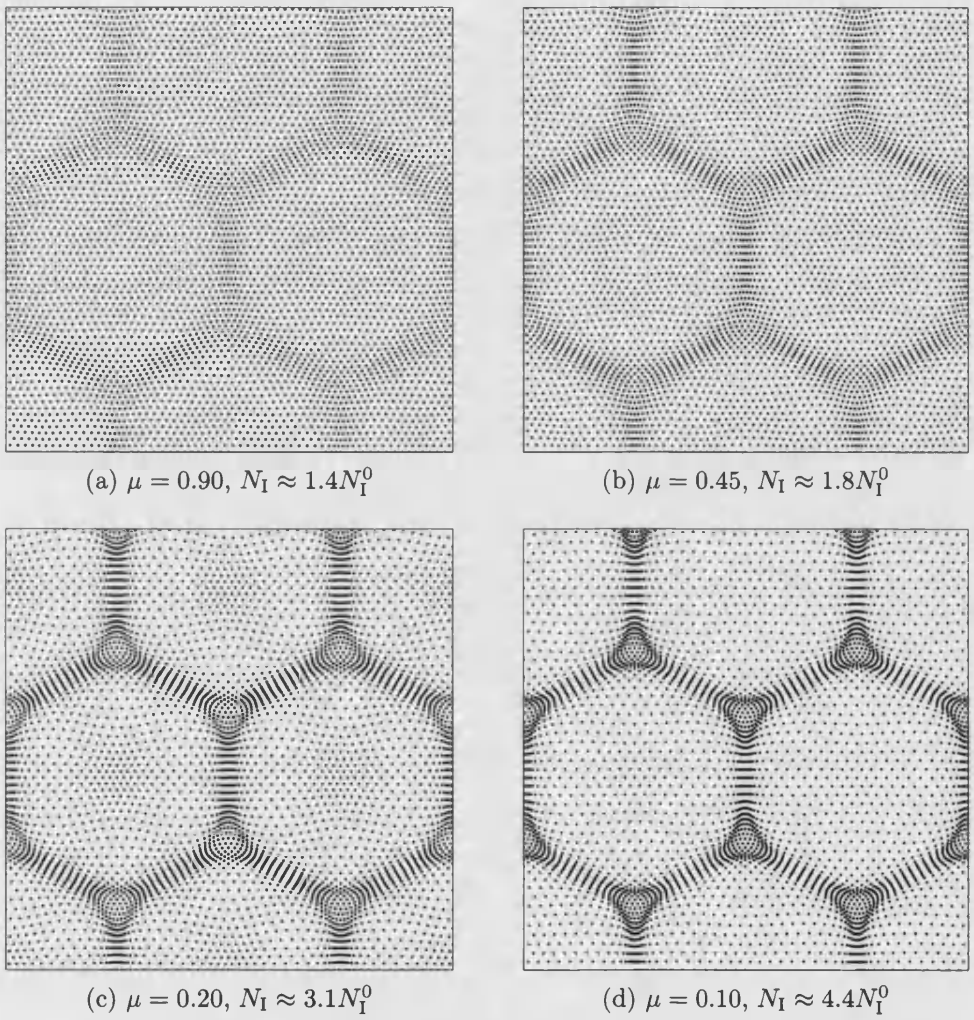


Figure 6.7: Test system for supercell calculations. The cladding consists of rounded hexagonal air holes ($n = 1$) in a glass background ($n = 1.5$) giving an air-filling fraction of 91%, and a 7-cell air core. Details of the cladding and core are given in Figs. 6.5 and 6.6.

unchanged. In order to generate a grid, it is necessary first to choose a value of the compression parameter μ that results in a grid with the appropriate enhancement of sampling point density at the dielectric interfaces. To demonstrate the effect on the rounded hexagon structure of changing μ, we show in Fig. 6.8 a set of PCF cladding grids calculated for a range of values of μ. These grids are not used in practice for supercell calculations, but they are shown here for illustrative purposes and because it is convenient to determine appropriate values of μ using rapid single cell calculations before performing slower supercell grid generations. Each grid was created by applying P6mm symmetry as described in Sec. D.2 and minimising 6 independent variables[1].

Supercell grids can be generated in the same way as single cell grids, and using the same values of μ for supercells and single cells tends to produce grids with a similar level of distortion. In order to describe the greater complexity of supercell structures, however, a greater number of $x^p_{\mathbf{G}}$ coefficients is needed. As a result, the number of variables in the minimisation generating the grid is larger. Typi-

[1]Only 3 variables were used for the structure with circular holes in Sec. 6.1.1, but this produced unsatisfactory results for the rounded hexagon structure, perhaps owing to its increased structural complexity.



(a) $\mu = 0.90$, $N_\mathrm{I} \approx 1.4 N_\mathrm{I}^0$

(b) $\mu = 0.45$, $N_\mathrm{I} \approx 1.8 N_\mathrm{I}^0$

(c) $\mu = 0.20$, $N_\mathrm{I} \approx 3.1 N_\mathrm{I}^0$

(d) $\mu = 0.10$, $N_\mathrm{I} \approx 4.4 N_\mathrm{I}^0$

Figure 6.8: Grids of real-space sampling points generated for an array of rounded hexagonal air holes in glass (the cladding of Fig. 6.7; also shown magnified in Fig. 2.1 on page 34) smoothed with FWHM $\Gamma^\mathrm{x} = 0.03\Lambda$ (as for the grids in Sec. 6.1.1). The grids shown were calculated on an FFT grid of size 128×128, with every fourth point plotted to give a grid of 32×32. The approximate number of points in the dielectric interface regions perpendicular to the interfaces is given relative to that of the undistorted grid, N_I^0.

cally 100-300 variables are required; using too few produces grids that are poorly adapted to the structure[2]. It is essential to use an efficient minimisation method, i.e. a method making use of calculated gradients (as discussed in Sec. 5.4.2), when the number of variables is of this order. This is demonstrated by the rate of convergence of the BFGS algorithm, which for 200 variables requires only ≈ 350 iterations.

To illustrate an adaptive grid created for a supercell structure, Fig. 6.9 shows the grid created with compression parameter $\mu = 0.2$ for the supercell of Fig. 6.7. 200 variables were used in the minimisation and, as with the single unit cells, P6mm symmetry was applied in order to reduce the computational effort. The FFT grid used for the minimisation was of size 512×512, and that for $|\nabla_t n^2|$ was of size 1024×1024. The total time taken on a desktop PC was ≈ 15 min.

The following section describes the improvements in plane-wave convergence brought about by using the GCC method.

6.2.2 Plane-wave calculations

Before demonstrating the use of the GCC method, it is helpful to know the position of bandgaps in the PCF cladding and also the mode trajectories (plots of the variation of β with k_0) of guided modes within the bandgap. It is also necessary to determine an appropriate representative frequency at which to perform calculations. The conventional plane-wave method was used to locate the position of the first bandgap of the cladding structure and then to determine mode trajectories, and the results are shown in Fig. 6.10 for a frequency range within the first bandgap. To distinguish between different modes, the modes are labelled by symmetry type. Modal symmetry can be determined by examination of the magnetic field $\mathbf{h}_t$ in reciprocal space, the full method for which is described in detail in Sec. D.1 of Appendix D.

The mode trajectories shown in Fig. 6.10 demonstrate modes guided in the air core of the fibre (visible with near-vertical slope), and also 'surface' modes associated with the thin silica wall around the core (with a shallower slope). Although the specific issues surrounding surface modes are not relevant to this chapter, the

[2]This can be checked visually, and by calculation of N_I to determine the increase of grid point density at the dielectric interfaces.

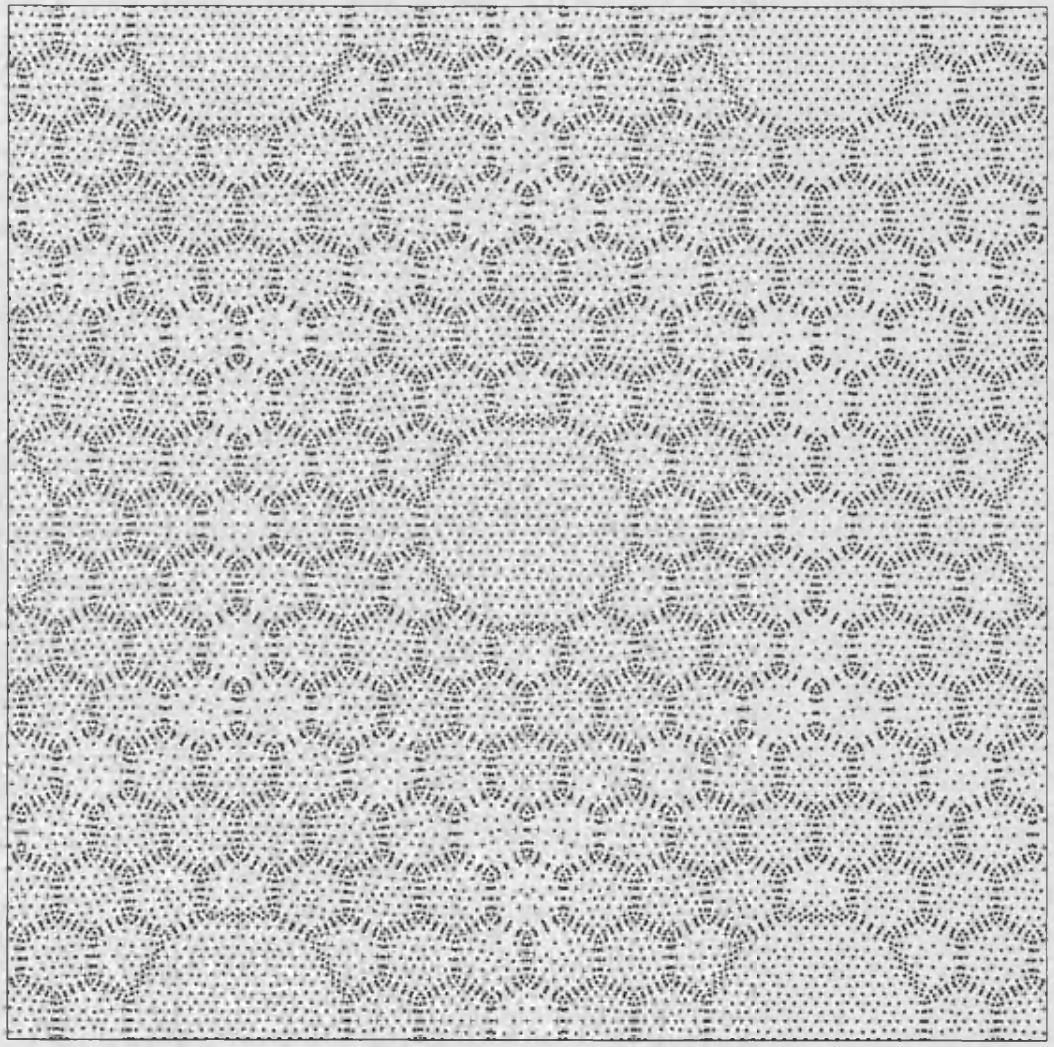


Figure 6.9: Grid of real-space sampling points for the supercell structure shown in Fig. 6.7, calculated with $\mu = 0.2$. The FFT grid size used was 512×512 and every eighth point is plotted to give a grid of 64×64. The smoothing used was $\Gamma^{\mathrm{x}} = 0.03\Lambda$. The increase in the number of grid points perpendicular to dielectric interfaces is $N_{\mathrm{I}} \approx 2.5 N_{\mathrm{I}}^{0}$.

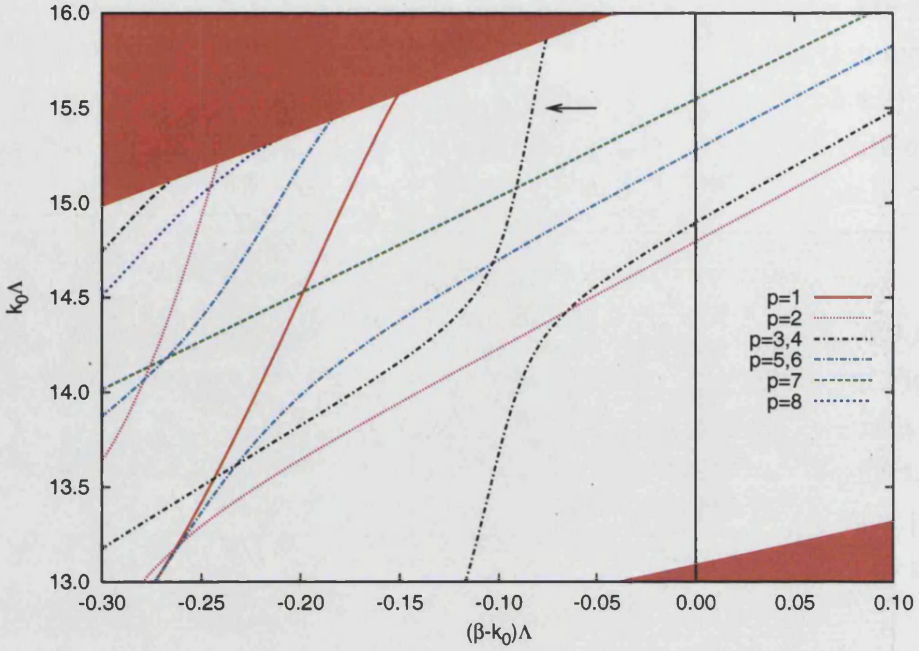


Figure 6.10: Modes of the PCF structure of Fig. 6.7, determined using the conventional plane-wave method with an FFT grid of dimension 512×512 and smoothing $\Gamma = 8 \times 10^{-3}\Lambda$. The horizontal axis shows the propagation constant plotted as $(\beta - k_0)\Lambda$, and the vertical axis shows normalised frequency $k_0\Lambda$. The red shaded regions in the upper-left and lower-right corners show the band edges of the infinitely-periodic cladding, and the air-line is marked with a vertical black line. The fundamental air-guided mode is marked with an arrow. Symmetry types are shown using the notation of Ref. 112 (see also Appendix D).

problems associated with them, and means to suppress them, are discussed in greater detail in Chapter 7.

To make a comparison between the convergence behaviour of the GCC method and that of the conventional plane-wave method, we calculate the four largest distinct β values below the air-line at $k_0\Lambda = 14.5$ (corresponding to three doubly-degenerate modes and one non-degenerate mode). This is a representative frequency at which the fundamental mode as well as surface modes are present.

The width of Gaussian smoothing used for calculations using the conventional plane-wave method was $\Gamma = 8 \times 10^{-3}\Lambda$. Consequently, $N_\mathrm{I} \approx 0.5$ at an FFT grid size of $N = 512$ and the results are not expected to be fully converged unless grid sizes above 512×512 are used. However, by using the GCC method, it is possible to enhance convergence. This is demonstrated in Fig. 6.11, in which the convergence of the conventional plane-wave method and that of the GCC method

are compared for the GCC grid with slight distortion ($\mu = 0.90$, corresponding to $N_\mathrm{I} \approx 1.4 N_\mathrm{I}^0$ and therefore $N_\mathrm{I} \approx 0.7$ at $N = 512$). The smoothing of the dielectric function in the GCC method has been chosen to be a close approximation to Gaussian smoothing of the same width as that used in the conventional method, $\Gamma = 8 \times 10^{-3}\Lambda$ (see discussion above in Sec. 6.1.2).

Fig. 6.11 shows clearly a slight improvement in the convergence behaviour of the GCC method compared to the conventional plane-wave method although, as expected, neither method is fully converged at $N = 512$. The series of calculations shown in Figs. 6.11-6.14 shows the behaviour of the GCC method for a range of values of the compression parameter μ. In each figure, the results from the GCC method are compared against the same results from the conventional method. A gradual improvement in the rate of convergence of the GCC method is clearly visible as μ is decreased, i.e. as the density of sampling points is increased. The slight difference in results at large FFT grid sizes is partly a result of the lack of convergence of the conventional method, and partly because of the inexact equivalence of smoothing in the GCC method to Gaussian smoothing.

The most distorted GCC grid is that with $\mu = 0.10$ with convergence behaviour shown in Fig. 6.14. No results are presented for $N > 288$ at this level of distortion because the GMRES linear solver is unable to converge. This problem is discussed in the following section.

6.2.3 Computational cost

The reformulation of the plane-wave method in GCCs shows improved convergence behaviour relative to the conventional plane-wave method, but there are several costs associated with this improvement. One cost, discussed in Sec. 5.3.3, is the increase in the number of FFTs required to perform matrix-vector operations: 5 are needed for each matrix-vector multiplication required by GMRES in the conventional method, whereas 12 are needed in the GCC method. However, this is not in general the limiting factor in determining the total time taken by the GCC method. Instead, the number of iterations required by GMRES in order to reach convergence is the most important difference between the two methods affecting the total calculation time.

It is well-known that sets of linear equations (represented as a matrix) are much

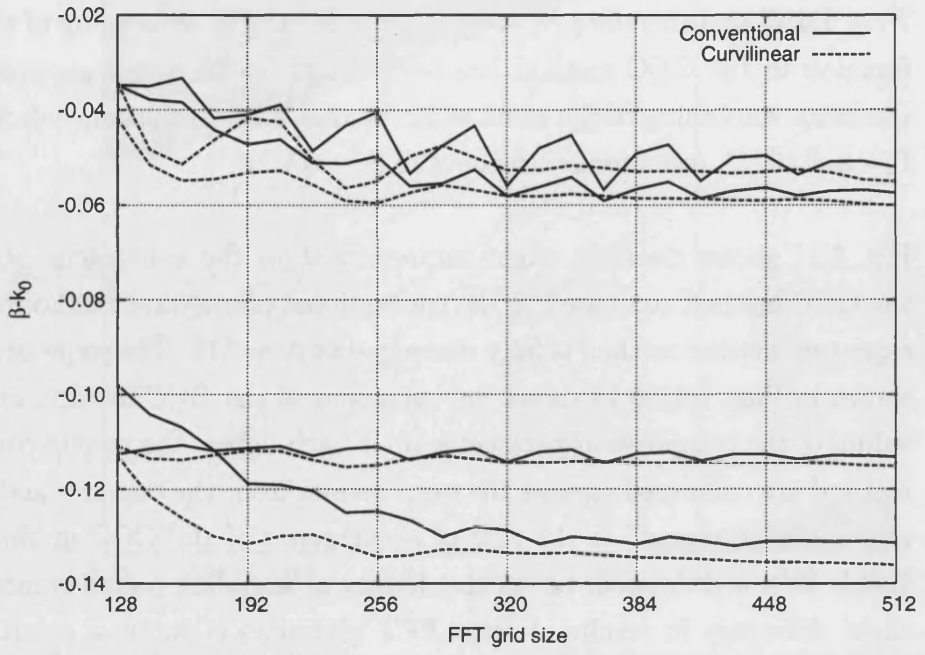


Figure 6.11: Convergence of the conventional and GCC plane-wave methods as the FFT grid size is increased, for a GCC grid with $\mu = 0.90$. The frequency is $k_0\Lambda = 14.5$. In both cases the smoothing of the dielectric function is $\Gamma = 8 \times 10^{-3}\Lambda$.

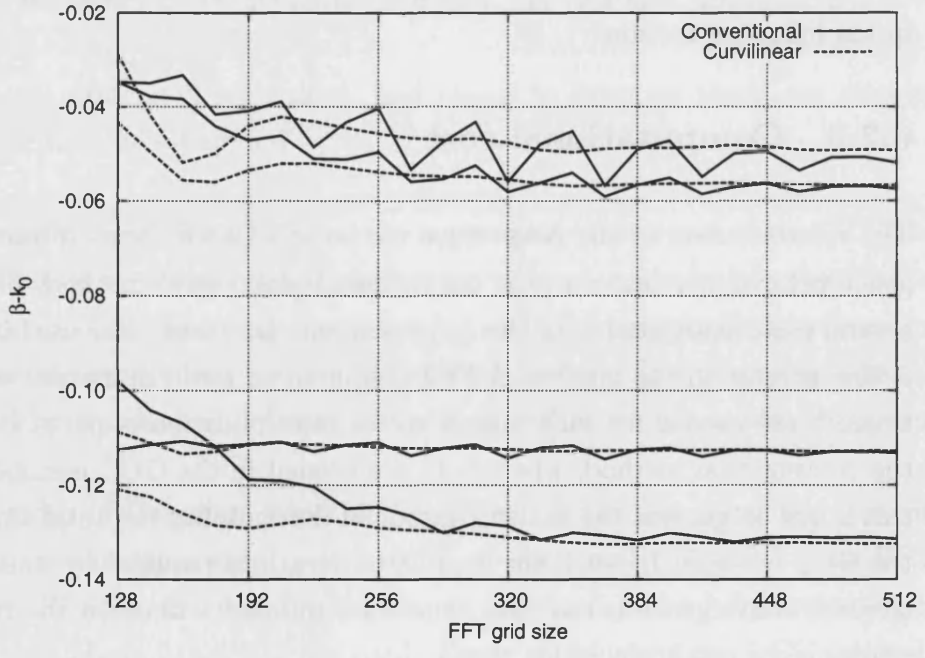


Figure 6.12: Convergence of the conventional and GCC plane-wave methods as in Fig. 6.11 but with $\mu = 0.45$.

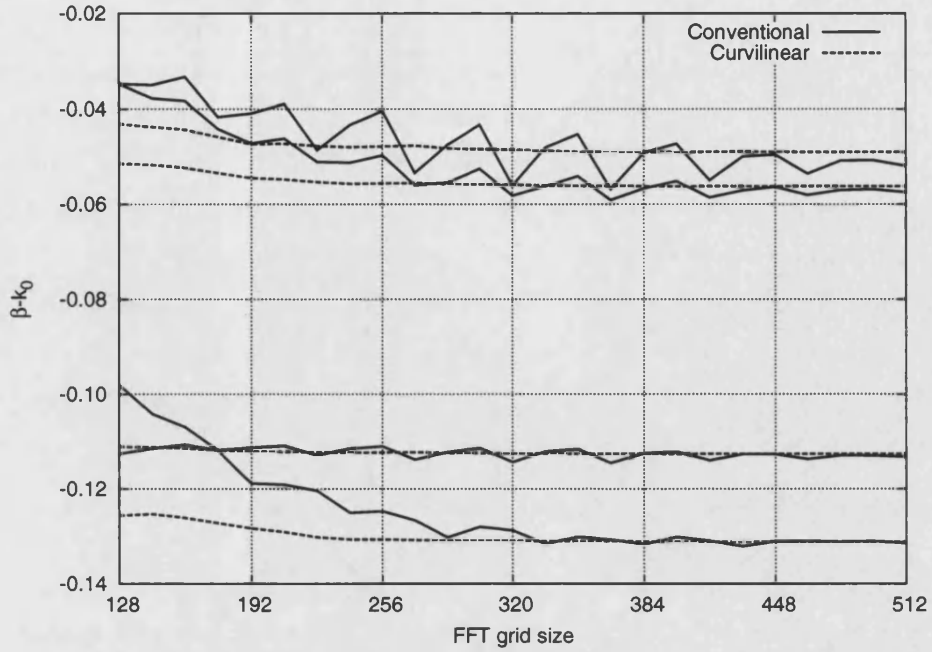


Figure 6.13: Convergence of the conventional and GCC plane-wave methods as in Fig. 6.11 but with $\mu = 0.20$.

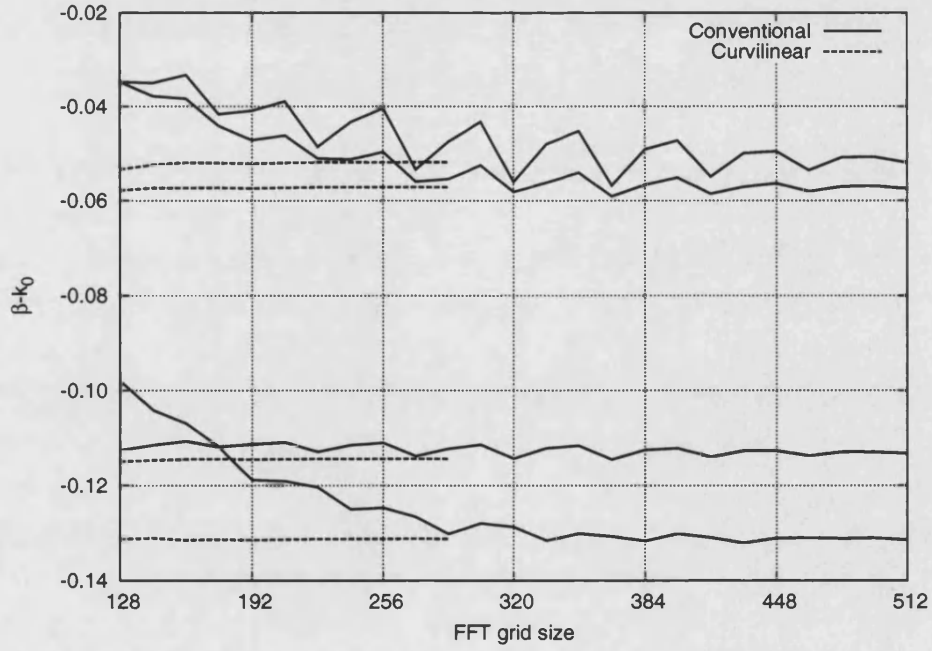


Figure 6.14: Convergence of the conventional and GCC plane-wave methods as in Fig. 6.11 but with $\mu = 0.10$. Note that, in contrast to Figs. 6.11-6.13, the β values of the GCC method generally converge from below rather than above.

more easily solved if they are *sparse*, i.e. contain only small numbers of non-zero elements [76, 80, 81]. The fixed-frequency plane-wave method (unlike some methods, such as the finite-difference method [59]) does not create a sparse matrix problem, but the structure of the matrix $(\mathbf{M} - \sigma\mathbf{I})$ defining the linear equation solved by GMRES (as described in Sec. 3.2.2) is such that the elements decay in magnitude away from the leading diagonal[3]. However, it is inevitable that if the compactness of the basis set is increased by the use of GCCs, then the matrix elements decay less rapidly away from the leading diagonal and create a more dense matrix. As a simple example of this, consider the matrix element for $\nabla_t^2 h^i$ which, in the conventional plane-wave method, is (from Eq. (3.7a)):

$$[\nabla_t^2 h^i]_{\mathbf{G}} = -\sum_{\mathbf{G}'} |\mathbf{k}+\mathbf{G}'|^2 \, \delta_{\mathbf{G},\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} \tag{6.1a}$$

$$= -|\mathbf{k}+\mathbf{G}|^2 \, h^i_{\mathbf{k},\mathbf{G}}, \tag{6.1b}$$

whereas in the GCC method it is

$$[\nabla_t^2 h^i]_{\mathbf{G}} = -\sum_{p,q=1}^{2} \int \mathrm{d}^2\xi \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} (k_p + G_p - iA_p) g^{pq} (k_q + G'_q + iA_q) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \tag{6.2}$$

The matrix representing Eq. (6.1a) is diagonal (and therefore sparse), because the RHS is zero if $\mathbf{G} \neq \mathbf{G}'$. However, the matrix representing Eq. (6.2) is dense: it has non-zero RHS in general for all $\mathbf{G}$,$\mathbf{G}'$.

The consequence of the increased density of matrix $\mathbf{M}$ is that the number of iterations needed by GMRES to invert is also increased. Table 6.1 shows the increase in the number of GMRES iterations as the compression parameter μ is decreased, together with approximate total computation times. The number of iterations and the total time taken for $\mu = 0.10$ is greatly increased by the need to restart GMRES after 100 iterations; the requirement for restarting is discussed in greater detail in Sec. 7.3.1.

It can be seen from Table 6.1, by comparing results from the conventional method and a small distortion of $\mu = 0.90$, that the increase in time by a factor of $12/5 = 2.4$ caused by performing additional FFTs is not noticeable. This is because a large preconditioner is used, and the time taken to apply the preconditioner by direct matrix-vector multiplication (which is not increased by using

[3]This can be seen by considering Eq. (3.6) on page 41 and noting that Gaussian smearing causes $n^2_{\mathbf{G}-\mathbf{G}'}$ and $[\ln n^2]_{\mathbf{G}-\mathbf{G}'}$ to decay as $(\mathbf{G}-\mathbf{G}')$ increases.

μ	$N_\mathrm{I}/N_\mathrm{I}^0$	GMRES iterations	Time/s
Conventional	1	20	400
0.90	1.4	25	550
0.45	1.7	35	700
0.20	2.5	90	2,100
0.10	3.8	450	12,000

Table 6.1: Approximate number of GMRES iterations per matrix-vector operation and total time taken to determine 8 modes of the PCF structure of Fig. 6.7 with an FFT grid of dimension 256×256 and preconditioner of size $N_\mathrm{P} = 6500$. The time taken to create the preconditioner is ≈ 700 s in each case and is not included in the above times. The increase in number of sampling points N_I is also given.

the GCC method) dominates over the time taken to perform FFTs. With larger FFT grid sizes the factor of 12/5 (see Sec. 5.3.3) could be more noticeable, but in practice the preconditioner effectiveness worsens and the total number of GMRES iterations becomes the dominant factor in determining the total computational time.

It is particularly evident from Table 6.1 that increasing the grid distortion rapidly causes an increase in the required number of GMRES iterations. As a result, it is not practical to use very large distortions (small μ) without making some attempt to reduce the number of iterations. Methods for controlling the total number of GMRES iterations in the context of the conventional plane-wave method are discussed further in Sec. 7.3.1, but for the GCC method described in this chapter we limit the discussion to distortions of $\mu = 0.20$, which can be handled without recourse to changes in preconditioner design or GMRES restarts.

Fig. 6.13 shows that, for $\Gamma = 8 \times 10^{-3}\Lambda$, GCC calculations at $\mu = 0.20$ are converged at an FFT grid size of $N \approx 256$, whereas those for the conventional plane-wave method require at least $N \approx 512$. To determine whether the GCC method provides an improvement in the computational time for similar calculations, it is helpful to make a direct comparison of these two cases. The conventional plane-wave method requires $\approx 1,000$ s, and the GCC method requires $\approx 2,100$ s (excluding preconditioner creation time of ≈ 700 s in both cases). This illustrates that, in general, the GCC method does not allow faster computation of equivalent results than the conventional plane-wave method. However, we can consider $N = 1,024$ to be a practical limit on the FFT grid size used on a desktop PC, and the requirement of plane-wave convergence consequently limits smoothing widths

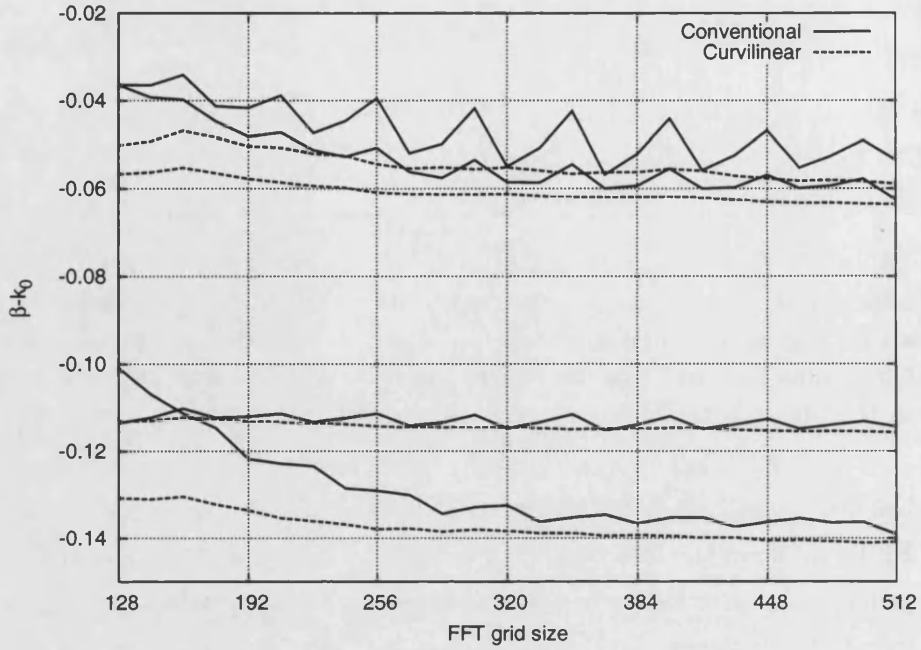


Figure 6.15: Convergence of the conventional and GCC ($\mu = 0.20$) plane-wave methods. The frequency is $k_0\Lambda = 14.5$. In both cases the smoothing of the dielectric function is $\Gamma = 8 \times 10^{-4}\Lambda$; note the lack of convergence of the conventional method.

to $\Gamma \approx 5 \times 10^{-3}\Lambda$. Although it is not faster than the conventional plane-wave method, the GCC method provides a way to carry out converged calculations with smaller smoothing widths than can be practically used in the conventional method. This is demonstrated in Fig. 6.15, in which the results shown for the conventional plane-wave method for a very narrow smoothing width are clearly unconverged, whereas those of the GCC method are better converged.

6.3 Conclusions

The results presented in this chapter demonstrate that the reformulation of the fixed-frequency plane-wave method in GCCs, as described in Chapter 5, provides a practical way to improve plane-wave convergence. In order to use the method, it is first necessary to generate coordinate transformations that are adapted to the structure under consideration; the efficient method of fictitious energy minimisation developed in Sec. 5.4.1 has been successfully demonstrated both for single unit cell (cladding) and for supercell (defect) geometries.

Sec. 6.1 demonstrated the application of the GCC method to single unit cells

describing the cladding of a PCF. The results of these calculations are shown in Fig. 6.4; it is clear from this figure that the method can give a reduction in the required FFT grid size from over 128×128 to $\approx 48 \times 48$ while retaining plane-wave convergence. Because the GCC grid used provided an increase in N_{I} by a factor of ≈ 5, this improvement is in approximate agreement with the 'convergence in principle' argument developed in Chapter 4.

The GCC method can also be used together with the supercell approximation to determine the guided modes of a defect introduced in a PCF cladding, and this use of the method has also been demonstrated above. Again, the 'convergence in principle' argument gives an approximate quantification of the expected improvement in the convergence behaviour, and this is demonstrated in Figs. 6.11-6.14.

The most significant limitation to the use of GCCs lies in the fact that the matrix problem solved by GMRES is more difficult to solve when using GCCs than with the conventional plane-wave method—a difference that we attribute in Sec. 6.2.3 to the reduced sparseness of the matrix. This places a practical upper limit on the amount of distortion that it is possible to use, at least without further development of preconditioning methods better adapted to the matrix problems created by the method. Together with the intrinsic increase in computational cost of the GCC method (by a factor of 2.4 resulting from the need to carry out more FFTs), the convergence problems associated with using large distortions prevent the GCC method from being faster than the conventional plane-wave method for equivalent calculations.

Although the GCC method is not faster than the conventional plane-wave method for carrying out equivalent calculations, it does possess the one important advantage of requiring smaller basis sets. Because smaller basis sets can be used while retaining plane-wave convergence, the GCC method allows well converged calculations with small dielectric smoothings to be carried out, whereas the basis sets needed to obtain convergence with the conventional plane-wave method would be prohibitively large.

Chapter 7

Large index contrast bandgap-guiding PCF

This chapter describes the design of a hollow-core PCF structure that is suitable for the guidance of light with wavelengths in the mid- to far-infrared. In addition to having practical use, such fibres are of interest computationally because suitable glasses tend to have a higher refractive index than silica. As has been demonstrated in Chapter 4, this creates a more challenging problem for the fixed-frequency plane-wave method: both the smoothing error and the plane-wave convergence error are greater than when modelling air-silica fibres. There is also an additional difficulty caused by a decrease in preconditioner effectiveness, and this has made necessary a more detailed study of preconditioning and how it can be improved for use with high-contrast dielectrics.

The design of all hollow-core PCF is constrained by a need to ensure that surface mode crossings (described below) do not cause significant degradation of the guidance properties of the fibre. This problem is discussed in detail, and the fixed-frequency plane-wave method is used to investigate a set of potential designs for core walls in order to minimise the effect of surface modes.

An introduction to the reasons for studying fibres with large index contrast is given below, followed by an analysis of the difficulties associated with modelling such fibres with the fixed-frequency plane-wave method. Potential fibre designs

are then discussed and analysed, and results demonstrating their properties are shown.

7.1 Introduction

There is considerable interest in developing hollow-core PCF for use in the mid- to far-infrared (IR) wavelength region of $\approx$ 2-10 μm. This has applications such as power delivery (e.g. of CO_2 laser beams), LIDAR ('light detection and ranging'), mid-IR spectroscopy, and fibre-based sensors and devices [45]. However, although much previous work has focussed on hollow-core PCF formed from silica glass, it is not suitable for the guidance of light in the mid- to far-IR wavelength range for reasons to be discussed below.

Guidance in air-silica fibres is usually dependent on the bandgap that arises between the 4^{th} and 5^{th} bands (counting from the band with the largest β)[1]. This bandgap is appreciable at high air-filling fractions of $\gtrsim$ 80% [41, 45], and is denoted a 'type-I' bandgap in the notation of Ref. 45. Fig. 7.1a illustrates the density of states (DOS) for a typical air-silica structure showing the type-I bandgap. However, the use of the type-I bandgap is limited to the wavelengths over which silica glasses are transparent. Outside this window of transparency, losses become large even when the fraction of light in glass is small. The upper wavelength edge of this window for silica is in the mid-infrared region at $\approx 2\,\mu$m [10], although there has been demonstration of PCF guidance at wavelengths up to $\approx 3\,\mu$m [113]. For wavelengths above this, silica is not a suitable material for PCF.

There exist glasses that are transparent at wavelengths above $2\,\mu$m, such as tellurites [114, 115] and chalcogenides [116, 117], but these glasses have a higher refractive index than silica (nominally 1.8-2.3 and 2.4-2.7 respectively [45]). The increased refractive index of the glass causes the type-I bandgap to close up and move to the high-β side of the air-line [45]. Instead, a different bandgap must be used. A suitable bandgap is the so-called 'type-II' bandgap, which occurs between the 8^{th} and 9^{th} bands of the PCF cladding bandstructure when the refractive index of the glass is higher than that of silica ($n \gtrsim 2.0$) and the air-filling fraction is relatively low ($\approx$ 60%); this gap has been shown to be robust

[1]The calculations of the modes of the air-silica fibre described in Sec. 6.2 (see, e.g., the mode trajectories of Fig. 6.10 on page 116) used this bandgap.

to perturbations in hole radius and refractive index [45, 118]. Fig. 7.1b gives the DOS for a typical structure comprising high-index glass and demonstrates the type-II bandgap. Tellurite and chalcogenide glasses are both transparent over the required wavelength range, and are therefore promising candidates for the manufacture of fibres based on the type-II bandgap. In the remainder of this chapter the modelling work uses the representative refractive index of $n = 2.4$.

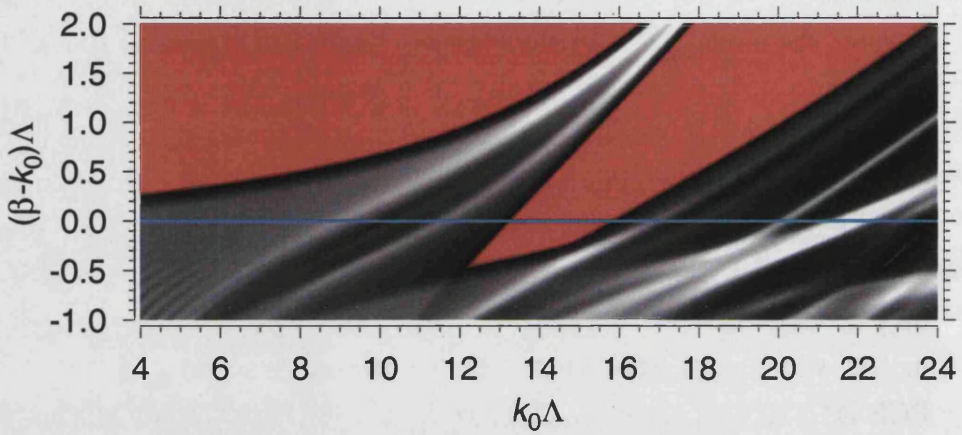
Although it has been established that type-II bandgaps are present in PCF claddings comprising high-index glasses, this does not necessarily imply that it is possible to design useful fibres making use of these gaps. In order to guide light, any practical PCF must have a core, and in hollow-core PCF this core must be surrounded by a 'core wall' of glass. The core wall may introduce 'surface modes' which are known to be a major cause of transmission loss in hollow-core PCF [41, 119, 120]. It is necessary to design the core of high-contrast fibres carefully to suppress these modes. The issue of core wall design is discussed in more detail below in Sec. 7.2.1.

7.2 Fibre design

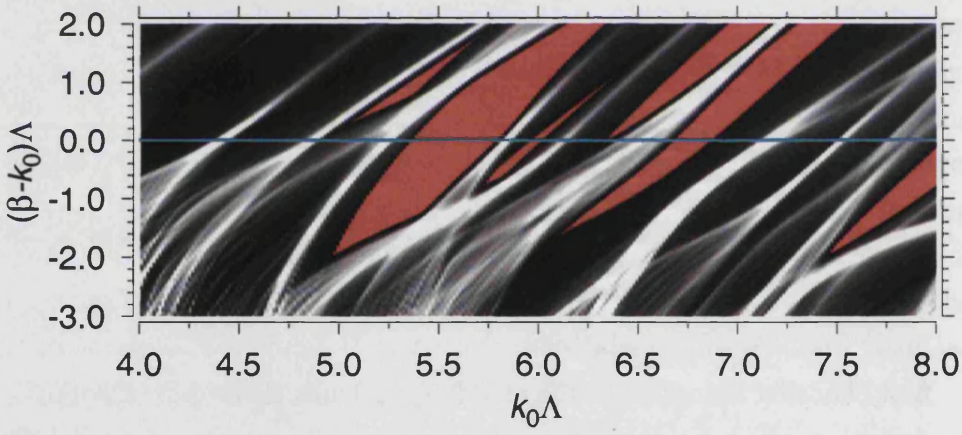
The design of PCF in general involves a trade-off between developing structures that have desirable properties (such as wide bandgaps and 'clean' core-guided modes without the presence of surface modes), and ensuring that these structures are physically realisable. In this section the design of a realistic PCF structure made from high-index glass using the type-II bandgap is described. This involves first choosing a design for the cladding that has the required bandgap, and then designing an appropriate shape for the core. The need for careful consideration of the core design is discussed in detail.

For a triangular lattice of circular air holes in a glass of refractive index $n = 2.4$, the type-II bandgap at the air-line is at its greatest frequency width when the radius of air holes arranged in a triangular lattice of pitch Λ is $r \approx 0.4\Lambda$.[2] A reasonable starting point for the design of a hollow-core fibre using this bandgap is therefore a cladding of air holes with $r = 0.4\Lambda$. The corresponding air-filling fraction of 58% is much lower than that needed for the type-I bandgap in silica,

[2] A map of the photonic density of states of this structure for the frequency range of interest is given in Fig. 7.1b, and for a range of hole radii in Ref. 45.



(a) Type-I bandgap in an array of rounded hexagonal air holes in a glass of $n = 1.5$ (structure as in Fig. 6.5 on page 110). The air-filling fraction is 91%.



(b) Type-II bandgap ($5.3 \lesssim k_0\Lambda \lesssim 5.8$ at the air-line) in an array of circular air holes of radius $r = 0.4\Lambda$ in a glass of $n = 2.4$ similar to Fig. 6.1 on page 102. The air-filling fraction is 58%.

Figure 7.1: Density of states for two PCF cladding structures showing type-I and type-II bandgaps. In both cases the horizontal line is the air-line. Red regions show bandgaps (zero DOS), and the grey-scale colouring is such that black represents low DOS and white represents high DOS. Details of the definition and method of calculation of the density of states are given in Appendix E.

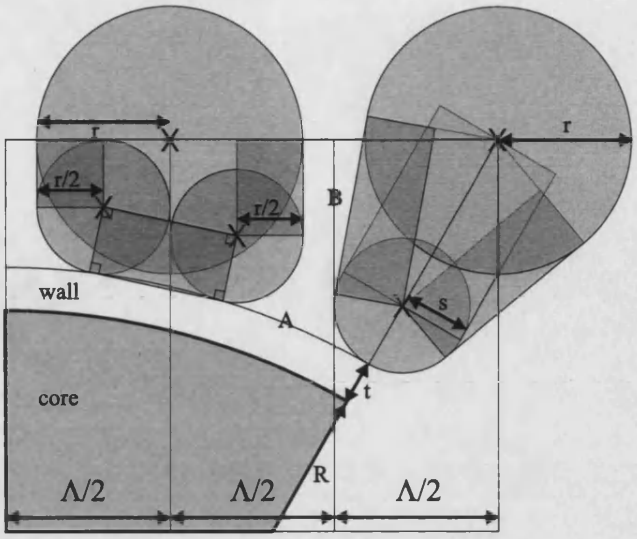


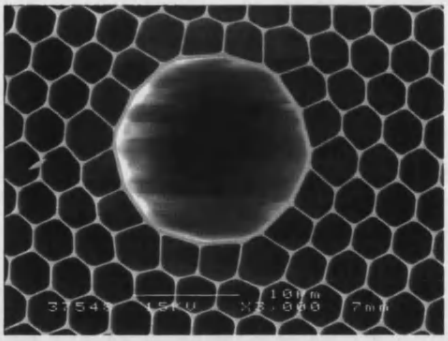
Figure 7.2: Core design of high-index PCF using geometrical shapes, showing air regions in grey (different shades of which are shown for clarity) and glass in white. For simplicity, only the unique $\frac{1}{12}$ of the design is given; the remainder may be obtained by applying C_{6v} symmetry operations. The two large circles lie on the cladding lattice, and the radius s is chosen such that the corresponding circle touches arc A and line B. The adjustable parameters are the cladding hole radius r and the core wall thickness t, which is controlled by the radius R of the central hole; we consider here a fixed cladding hole radius $r = 0.4\Lambda$ and vary R. Design and figure courtesy of Dr John Pottage.

and therefore circular air holes may be used for the type-II cladding instead of the rounded hexagons usually used in air-silica claddings.

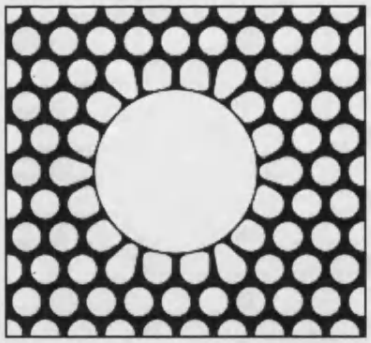
For ease of modelling, a 19-cell PCF core was designed using only geometrical shapes (rectangles and circles), as shown in Fig. 7.2. The details of the design were chosen using currently existing silica fibres as a guide; the similarity between the designs is shown in Fig. 7.3. In particular, the relatively constant core wall thickness of the silica fibre is reproduced, together with the fact that the cladding structure returns to being undistorted beyond one ring of holes around the core. The design provides an adjustable thickness of core wall (marked as t in Fig. 7.2), the importance of which is discussed in the following section.

7.2.1 Core wall design

Both experimental evidence [41, 119, 121] and theoretical studies [120, 122] have shown that the anti-crossings between surface modes (modes associated with the core surround) and core modes of the same symmetry play a major role in



(a) An SEM image of a typical silica hollow-core fibre (fabricated by Blaze Photonics).



(b) The model design for high-index glass created using geometrical shapes. The core wall thickness shown is $t = 0.05\Lambda$.

Figure 7.3: Hollow-core air-silica PCF and model design for high-index PCF.

causing the transmission loss of hollow-core PCF. It is known that in air-silica PCF, thicker core walls tend to support more surface modes [123]. It is reasonable to expect that using glasses of a higher refractive index than that of silica may also allow the core wall to support a greater number of surface modes. As a result, it is necessary to take care to design a core wall such that it suppresses surface modes as much as possible.

The problem of suppressing surface modes has been considered in air-silica fibres for 'undistorted' cores, which are cores formed by superimposing an air circle on an otherwise periodic PCF cladding. This may be done either with or without a thin wall of glass around the air circle forming a core wall, but without any distortion of the ring of air holes around the core (in contrast to, for example, Fig. 7.3, where the ring of air holes around the core has a distorted shape). If there is no core wall, there exist core radii such that no surface modes are supported [124, 125]. If a core wall does exist, then the core radius should be carefully selected and the wall thickness should be as thin as possible to suppress surface modes [123]. However, in fabricating real fibres, it is not yet possible to create an 'undistorted' core. Distortion inevitably occurs in practice as a result of the fibre drawing process, and all real fibres have a core wall.

In addition to the studies of 'undistorted' cores in air-silica PCF, there has also been some study of distorted (real) cores. Suggestions for suppressing the surface modes in distorted cores include adding 'fingers' of glass protruding into the core to reduce the distortion [120, 124], using a thin core wall [122], or choosing the wall thickness to be 'anti-resonant' [126]. The model design shown in Fig. 7.2

provides the ability to test the effectiveness of choosing an appropriate core wall thickness to suppress surface modes, and the results presented below in Sec. 7.4 concern a range of core wall thicknesses.

Having chosen a particular design, it is then possible to use the fixed-frequency plane-wave method to determine its properties. The way in which this was done is described in the following section.

7.3 Modelling large index contrast fibres

The fixed-frequency plane-wave method as described in Chapter 3 can be used to model PCF structures with a large index contrast, but the index contrast causes difficulties for the method. The principal problem, outlined below, is slow convergence of the linear solver. A discussion is also given of why, as a result of these convergence problems, the GCC method cannot be used to provide an improvement. A method to overcome the slow convergence using an improved preconditioner in the conventional fixed-frequency plane-wave approach is then described.

As discussed in Sec. 3.3.2, a linear solver is needed in the fixed-frequency plane-wave method in order to solve equations of the form

$$\mathbf{P}(\mathbf{M} - \sigma\mathbf{I})\mathbf{w} = \mathbf{P}\mathbf{u}, \tag{7.1}$$

where $\mathbf{u}$ is known and $\mathbf{w}$ is required. The operation $(\mathbf{M} - \sigma\mathbf{I})$ can be performed by FFTs, and $\mathbf{P}$ is a preconditioner chosen to be an approximate inverse of $(\mathbf{M} - \sigma\mathbf{I})$. We use the GMRES linear solver, which determines $\mathbf{w}$ iteratively and requires as input only the action of either $(\mathbf{M} - \sigma\mathbf{I})\mathbf{y}$ or $\mathbf{P}\mathbf{y}$ on an arbitrary vector $\mathbf{y}$.

As described in Sec. 3.3.2, there exist situations in which the 'Block/Jacobi' preconditioner—a combination of an exactly-inverted square sub-matrix of dimension $N_\mathrm{P} \times N_\mathrm{P}$ together with the Jacobi preconditioner—can be ineffective. The modelling of PCF structures with a large index contrast is one such situation. In order to demonstrate the problem, consider Fig. 7.4 in which typical convergence behaviour of GMRES for the high-index structure shown in Fig. 7.3b is plotted. If the size of the exactly-inverted sub-matrix is insufficiently large,

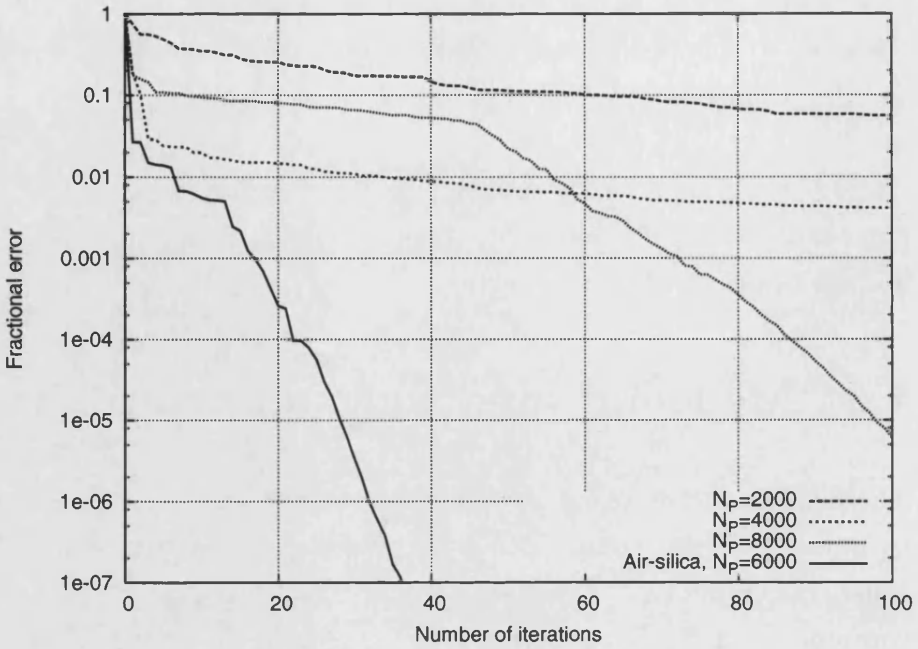


Figure 7.4: Typical convergence of the GMRES linear solver for a glass of $n = 2.4$ and structure as shown in Fig. 7.3b with a 9×9 supercell and FFT grid size 512×512. The fractional error is given by $\left|\mathbf{P}(\mathbf{u} - (\mathbf{M} - \sigma\mathbf{I})\mathbf{w}^{(m)})\right| / |\mathbf{P}\mathbf{u}|$, where $\mathbf{w}^{(m)}$ is the solution after m iterations. For comparison, the solid line shows the convergence of GMRES for a typical air-silica structure at the same FFT grid size. Note the stagnation of GMRES for the high-index structures with $N_\mathrm{P} < 8,000$. The approximate times required to create the preconditioners by exact inversion on a desktop PC are: $N_\mathrm{P} = 2,000$, 10 s; $N_\mathrm{P} = 4,000$, 90 s; $N_\mathrm{P} = 8,000$, 12 min.

the algorithm can require a large number of iterations in order to converge. For $N_\mathrm{P} = 2,000$ and $N_\mathrm{P} = 4,000$ the behaviour demonstrated is that of 'partial stagnation', in which the improvement brought about by each successive iteration becomes progressively less.

It can be shown that GMRES can locate the solution of any $N \times N$ matrix problem in at most N iterations and therefore, in principle, GMRES always converges [82]. In practice, for the reasons discussed below, it is not possible to provide the required storage or computational time if the algorithm does not converge considerably more favourably than this[3]. For this reason only the first 100 iterations are shown in Fig. 7.4.

To understand why slow convergence is a particular problem, it is necessary to

[3]Although there exist pathological cases for which N iterations are needed ('complete' stagnation), it is generally the case that GMRES needs $\ll N$ iterations [127].

consider the storage requirement and computational overhead (the computation required to determine each successive iterative step) associated with using GMRES. Both the storage and computational overhead are consequences of the way in which GMRES uses information from all previous iterations in order to determine the optimal next step towards the solution. When solving a system with N variables (in this case $N = 2N_{\mathrm{PW}}$), the storage requirement after m iterations is $O(mN)$ and the overhead associated with determining the next vector is $O(m^2N)$ [82]. These requirements place a practical upper limit on the number of iterations; it is not usually desirable to continue the algorithm for more than ≈ 100 iterations. Ways to circumvent this problem are discussed in the following section.

Sec. 6.2.3 demonstrated that the decrease in matrix sparseness brought about by using the GCC method made the solution of the linear matrix equation by GMRES more difficult, and the required number of GMRES iterations increased with grid distortion. When using the GCC method to study high-index structures such as those considered in this chapter, the convergence problems encountered in the conventional plane-wave method would be worsened further by using GCCs. As a result, it is not possible to use the GCC method to study these high-index structures without considerable improvements in preconditioning. In the remainder of this chapter, only improvements in preconditioning in the conventional plane-wave method will be considered.

7.3.1 Restarting and preconditioning

There are several existing methods by which the number of GMRES iterations can be kept within the tolerable upper limit. One way is to restart GMRES after a fixed number of iterations. Restarting involves using as an initial guess for the next set of iterations the best estimate of $\mathbf{w}$ obtained so far, but discarding the Krylov subspace built up by the previous applications of $(\mathbf{M} - \sigma\mathbf{I})$. However, the loss of information dramatically reduces the rate of convergence immediately after restarting (see Refs. 80,81 and Fig. 7.5 below), and in general when studying PCF it is important to avoid the need to restart the algorithm wherever possible.

Another way to reduce the number of required GMRES iterations is to improve the preconditioning of the problem such that it converges more quickly. This is equivalent to making $\mathbf{P}$ a better approximation to the inverse of $(\mathbf{M} - \sigma\mathbf{I})$.

Clearly the simplest way to do this with the 'Block/Jacobi' preconditioner is to increase N_P as demonstrated in Fig. 7.4, but this suffers from being expensive, requiring storage $O(N_\mathrm{P}^2)$, time to perform the inversion of $O(N_\mathrm{P}^3)$, and time to perform exact matrix-vector multiplication of $O(N_\mathrm{P}^2)$. Another possibility is to use 'flexible' GMRES (FGMRES), which allows the use of a preconditioner that varies between iterations. It has been suggested [81] that using a few iterations of GMRES as a preconditioner in the FGMRES algorithm could enhance convergence, but we have not found any significant improvement using this method.

The perfect preconditioner for Eq. (7.1) would be an exact inverse given by $\mathbf{P} = (\mathbf{M} - \sigma\mathbf{I})^{-1}$, but clearly determining this inverse to create the perfect preconditioner is as expensive as solving the original problem. Consider instead the inverse given by $(\mathbf{M}' - \sigma\mathbf{I})^{-1}$, where $\mathbf{M}'$ is the matrix corresponding to a system that is similar (but not identical) to that described by $\mathbf{M}$. This inverse can provide an approximation to $(\mathbf{M} - \sigma\mathbf{I})^{-1}$ that is useful as a preconditioner in Eq. (7.1) if $\mathbf{M}$ and $\mathbf{M}'$ are sufficiently similar, and provided $(\mathbf{M}' - \sigma\mathbf{I})$ is somehow easier to invert than $(\mathbf{M} - \sigma\mathbf{I})$. If the action of this preconditioner on an arbitrary vector $\mathbf{y}$ is denoted $\mathbf{z}$, then determining $\mathbf{z}$ is equivalent to solving a new system given by

$$(\mathbf{M}' - \sigma\mathbf{I})\mathbf{z} = \mathbf{y}. \tag{7.2}$$

Like Eq. (7.1), this equation can also be solved by GMRES. However, convergence is enhanced significantly by also preconditioning *this* system. The simple 'Block/Jacobi' preconditioner, denoted $\mathbf{Q}$ here, can be used:

$$\mathbf{Q}(\mathbf{M}' - \sigma\mathbf{I})\mathbf{z} = \mathbf{Q}\mathbf{y}, \tag{7.3}$$

where $\mathbf{Q}$ is created using the matrix $\mathbf{M}'$ rather than $\mathbf{M}$. We denote the original problem of Eq. (7.1) the 'outer' GMRES loop, and this method of iteratively preconditioning the original problem the 'inner' GMRES loop. Note that both iterative loops are independent: there are two separate GMRES algorithms being used to solve different matrix equations. In summary, the determination of eigenvalues of $\mathbf{M}$ is carried out using these nested iterative methods as follows, where indentation is used to denote the different levels of iterative loops:

1. ARPACK provides arbitrary vector $\mathbf{u}$ to GMRES and requires $\mathbf{w}$ given by $\mathbf{w} = (\mathbf{M} - \sigma\mathbf{I})^{-1}\mathbf{u}$
2. Outer GMRES provides an arbitrary vector $\mathbf{y}$, and requires one of:
3. Multiplication operation $(\mathbf{M} - \sigma\mathbf{I})\mathbf{y}$ by FFT
4. Preconditioning operation $\mathbf{z} = (\mathbf{M}' - \sigma\mathbf{I})^{-1}\mathbf{y}$ by inner GMRES
5. Inner GMRES provides arbitrary vector $\mathbf{t}$, and requires one of:
6. Multiplication operation $(\mathbf{M}' - \sigma\mathbf{I})\mathbf{t}$ by FFT
7. Preconditioning operation $\mathbf{Qt}$ by 'Block/Jacobi' method
8. Go to 5 until converged, then return $\mathbf{z}$ to outer GMRES
9. Go to 2 until converged, then return $\mathbf{w}$ to ARPACK
10. Go to 1 until converged, then return eigenvalues/eigenvectors of $(\mathbf{M} - \sigma\mathbf{I})$.

Before this method can be used, it is necessary to find an appropriate structure on which to base the matrix $\mathbf{M}'$. The results from the 1D model of Chapter 4 suggest that plane-wave convergence is enhanced by increasing the width of Gaussian smoothing applied to dielectric structures. Previous work suggests that, especially for high-index structures, increasing the smoothing width can also enhance the rate of convergence of GMRES [74]. This observation shows that a structure with more smoothing provides an ideal candidate for $\mathbf{M}'$: it is an easily-controllable approximation to $\mathbf{M}$ (since $\mathbf{M}' \to \mathbf{M}$ as its smoothing is reduced to that of $\mathbf{M}$), and empirically we find that GMRES for smoother structures converges more quickly. A suitable smoothing width for the structure from which $\mathbf{M}'$ is created can be found by experimentation. If it is too narrow, the inversion of $\mathbf{M}'$ will suffer from the same convergence problems as $\mathbf{M}$; if it is too wide, then $\mathbf{M}'$ and $\mathbf{M}$ will be dissimilar and the preconditioner will be ineffective.

The following section presents an analysis of the properties of the model structure described in Sec. 7.2. These results have been obtained using the fixed-frequency plane-wave method together with the iterative 'inner GMRES' preconditioner.

7.4 Results

Fixed-frequency plane-wave calculations of the photonic bandstructure have been carried out for the PCF design shown in Fig. 7.2 for a range of core wall thicknesses $t/\Lambda = 0.03, 0.05, 0.07$. In this section the way in which the calculations were carried out is first described, and the numerical behaviour of the method (i.e. the improvement due to preconditioning by the 'inner GMRES' method) is quantified. The results of calculations are then given and discussed.

7.4.1 Numerical behaviour

The plane-wave method requires periodicity of the dielectric function, so in order to model a structure such as that shown in Fig. 7.3b it is necessary to use a supercell approximation. A 9×9 supercell was found to be the minimum size giving physically useful results, i.e. without showing significant coupling between adjacent cores. Gaussian smoothing of the dielectric function with FWHM of $\Gamma = 0.03\Lambda$ was used[4], and the size of the FFT grid was 512×512. For the smoothed preconditioner described in Sec. 7.3.1, a larger smoothing width of $\Gamma = 0.1\Lambda$ (determined by trial-and-error, observing convergence behaviour of both inner and outer GMRES loops) was used. This width is clearly much larger than the width desirable for accurate eigenmode determination, but it affects only the preconditioner and does not change the values of β resulting from the actual mode-solving calculation.

Typically we find that 3-6 matrix-vector operations are required per eigenvalue, each of which requires 20-50 outer GMRES operations. Each outer GMRES iteration requires an application of the iterative preconditioner, which for the highly smoothed structure converges within 15-30 inner GMRES operations. Typical behaviour of the outer GMRES algorithm with and without the iterative preconditioner is shown in Fig. 7.5. The iterative preconditioner requires less storage for GMRES because fewer iterations are needed, and the GMRES overhead (which is $O(m^2)$ after m iterations) is insignificant for the same reason. In comparison, the 'Block/Jacobi' preconditioner must be restarted and the GMRES overhead is much larger.

[4]Referring to Fig. 4.6 on page 67, the 1D model predicts an error in β of ≈ 0.5 for $n = 2.5$ at this level of smoothing. In practice the smoothing error in these 2D calculations is less. The accuracy of the results presented in this chapter was confirmed by using other widths of smoothing and observing the size of the resulting perturbation in β.

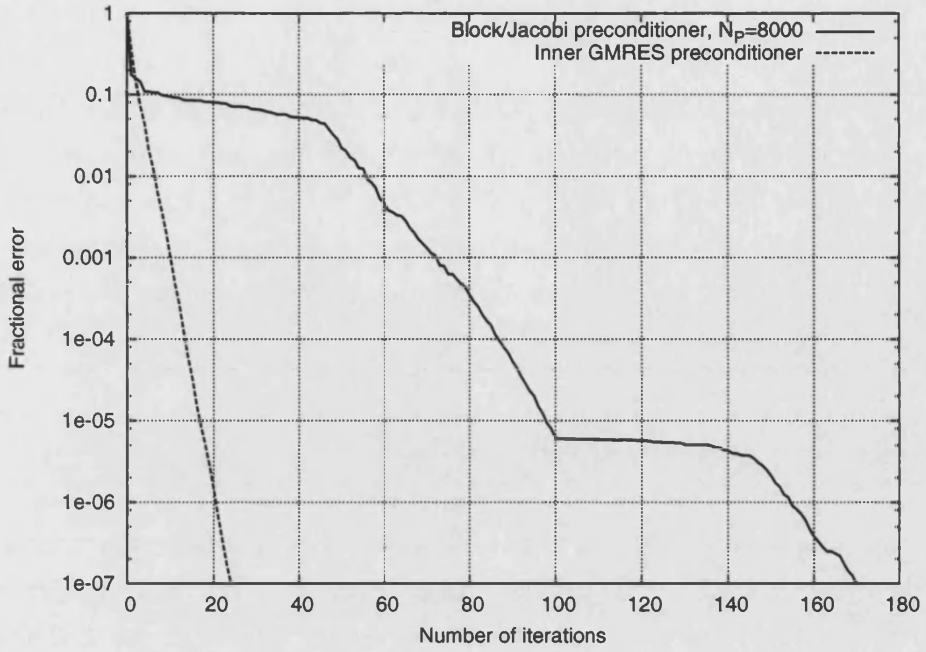


Figure 7.5: Typical convergence of the GMRES linear solver for a glass of $n = 2.4$ and structure as shown in Fig. 7.3b with a 9×9 supercell and FFT grid size 512×512. The fractional error is as defined in Fig. 7.4. The solid line shows the convergence for a large 'Block/Jacobi' preconditioner ($N_P = 8,000$) with a restart at 100 iterations; note the slow progress immediately following the restart. The dashed line shows the convergence of the outer GMRES loop when using the iterative (inner GMRES) preconditioner.

Applicability to the plane-wave method in curvilinear coordinates

It is worth considering whether the improved method of preconditioning demonstrated in this section would also provide an improvement when applied to the GCC method. One conclusion of Chapter 6 was that the convergence of GMRES is slower when using large distortions, and this places an upper limit on the amount of distortion that can be used. While the results presented in this section show that the inverse of $\mathbf{M}'$ constructed using a smoothed dielectric function provides an effective preconditioner, this behaviour relies on the fact that smoother structures converge more quickly than sharp structures. In the case of poor convergence in the GCC method, it is instead less distorted coordinate transformations which show more rapid convergence.

There is a fundamental difference between using a smoothed dielectric structure and using a less distorted coordinate transformation to construct $\mathbf{M}'$. Smoothing a dielectric structure changes the sets of plane-wave coefficients $n^2_{\mathbf{G}}$ and $[\ln n^2]_{\mathbf{G}}$,

but it does not change the definition of $h^i_{\mathbf{k},\mathbf{G}}$. In the case of the GCC method, a less distorted coordinate transformation not only has different $n^2_{\mathbf{G}}$ and $[\ln n^2]_{\mathbf{G}}$, but the matrix containing these coefficients (or equivalently the set of FFT-based operations) must act upon a vector of magnetic field components that are defined differently from those describing the desired (more distorted) solution. As a result, it is unlikely that the exact inverse of a matrix formed using a less distorted coordinate transformation would provide an effective preconditoner for large distortions.

7.4.2 Computational results

Fig. 7.6a shows a selected region of the type-II bandgap for the structure with core wall thickness $t = 0.03\Lambda$. The horizontal axis shows $(\beta - k_0)\Lambda$, extending from the air-line (the right-hand edge of the plot) to $\beta\Lambda = k_0\Lambda - 0.4$. The vertical axis shows normalised frequency, $k_0\Lambda$. Air-guided modes can be seen in this figure as near-vertical lines, whereas surface modes (being instead localised to glass rather than air) are visible as lines with a shallower slope. The fundamental air-guided mode can be seen as the rightmost near-vertical line and is marked with an arrow; it is followed as β decreases by a group of three modes of higher order (a non-degenerate mode, a doubly-degenerate pair, and another non-degenerate mode).

A large group of surface modes is visible in the lower-right corner of Fig. 7.6a. If the frequency range over which the fundamental mode is free of anti-crossings is to be extended, it is necessary to shift these modes to higher β. This may be achieved by the addition of glass, i.e. by increasing the core wall thickness slightly. In Fig. 7.6b we show the same region of the bandstructure but for a wall thickness of $t = 0.05\Lambda$. At this thickness the fundamental mode is 'clean' over almost the entire range shown, and its anti-crossings with surface modes are weaker (that is, the fundamental mode and surface modes of the same symmetry type approach closely at anti-crossings). An expanded bandstructure for this optimal thickness is shown in Fig. 7.7.

Increasing the core wall thickness further, as shown by Fig. 7.6c for $t = 0.07\Lambda$, continues to sweep the surface modes to higher β away from the fundamental mode (they can still be seen in the lower-right corner of the plot), but also leads to the appearance of new surface modes introduced at the upper edge of the

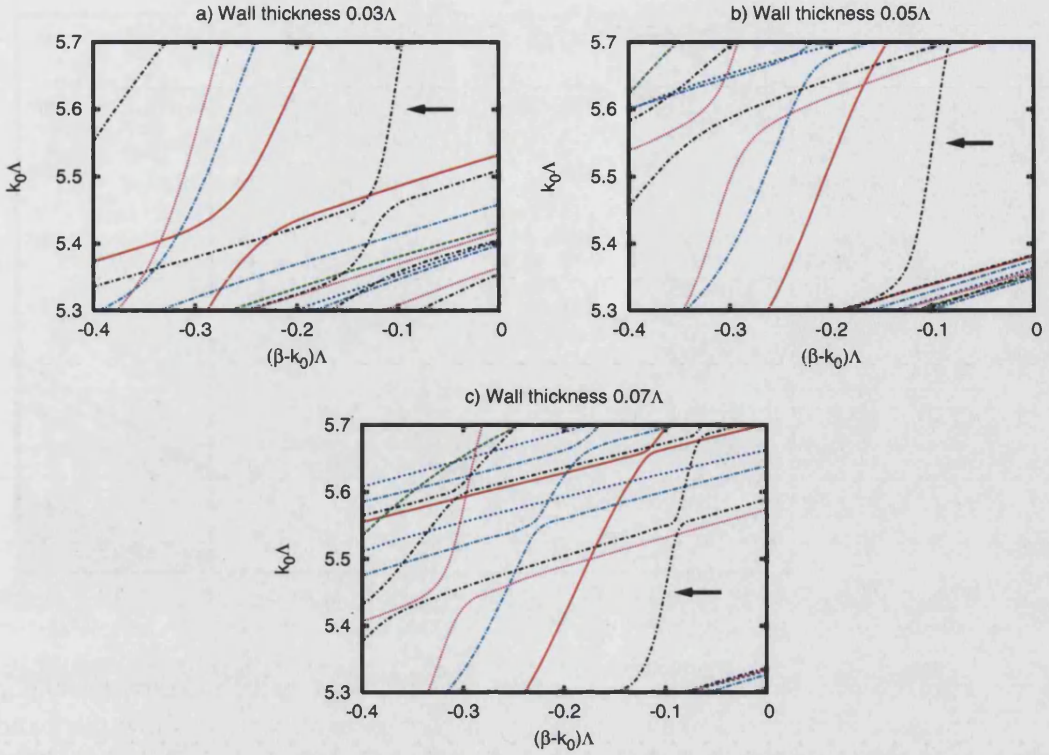


Figure 7.6: Mode trajectories for a range of three core wall thicknesses of the PCF design shown in Figs. 7.2 and 7.3b. The mode symmetry types are calculated using the method of Sec. D.1 and are labelled using the notation of Ref. 112; the key to symmetry types is given in Fig. 7.7. The fundamental air-guided mode is shown by an arrow in each figure.

bandgap. These new modes reduce the range over which the fundamental mode is 'clean'. A range of thicker walls has been considered but these significantly worsen the problem of surface mode crossings. This is consistent with expectations and with known results for very thick walls of silica glass [123].

Properties of the 'optimal' core wall design

In this section the 'optimal' structure with core wall thickness $t = 0.05\Lambda$ is studied further. We first show plots of the axial Poynting vector for the surface modes, and demonstrate that there is unlikely to be any simple underlying origin of the large 'clean' region, and then show that core wall anti-resonance does not explain its origin. The fraction of power in air for the structure is also calculated.

The nature of the surface modes supported by the $t = 0.05\Lambda$ structure is shown in Fig. 7.8. It is noticeable that the surface modes at the high-frequency edge of

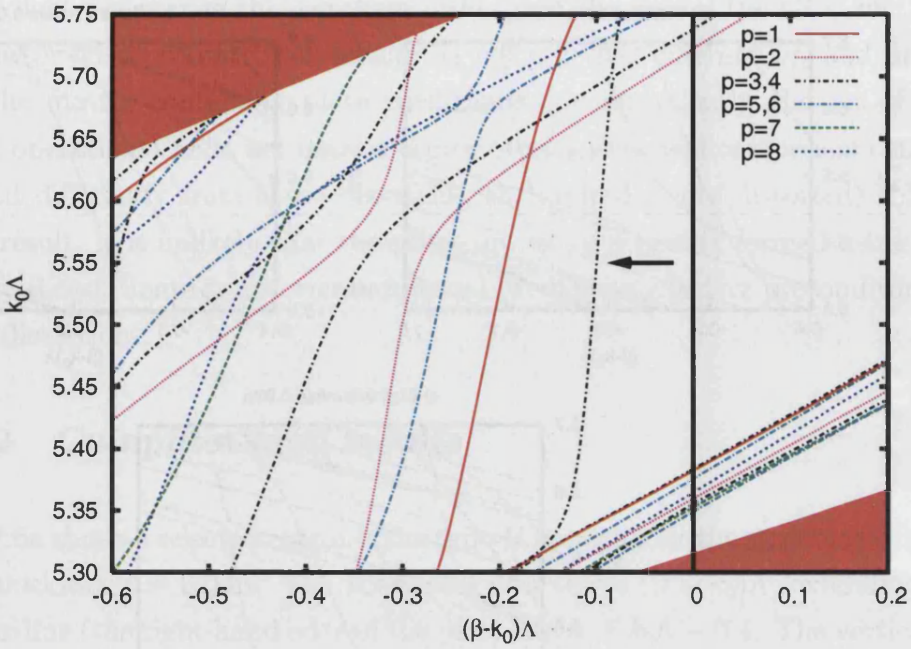


Figure 7.7: Modes of the PCF structure with core wall thickness $t = 0.05\Lambda$. The red shaded regions in the upper-left and lower-right corners show the band edges of the infinitely-periodic cladding, and the air-line is marked with a vertical black line. The fundamental air-guided mode is marked with an arrow. Symmetry types are calculated using the method of Sec. D.1 and shown using the notation of Ref. 112.

the 'clean' region tend to have intensity localised in the glass struts supporting the core wall and in the core wall itself, whereas those towards the low-frequency edge of the 'clean' region appear to be peaked further from the core wall in the glass surrounding the ring of distorted air holes. This is not a general feature of all of the surface modes and it therefore does not provide a simple way of understanding the reason for the existence of the large 'clean' region. However, the observation that surface modes are associated with the glass around the ring of distorted air holes, as well as with the cladding struts and the core wall itself, suggests that the underlying 'clean' region does not have a simple physical origin. Instead, it is likely to be a result of combinations of resonances of the different structural features, and sensitive to geometry (both the thickness of the core wall and the distortion around it).

One potential cause of the suppression of surface modes in PCF structures is core wall anti-resonance, and it is worthwhile to consider whether this is the cause of the 'clean' region. Using the principle behind the anti-resonant reflecting optical waveguide (ARROW) [128], anti-resonant (AR) core wall thicknesses are those

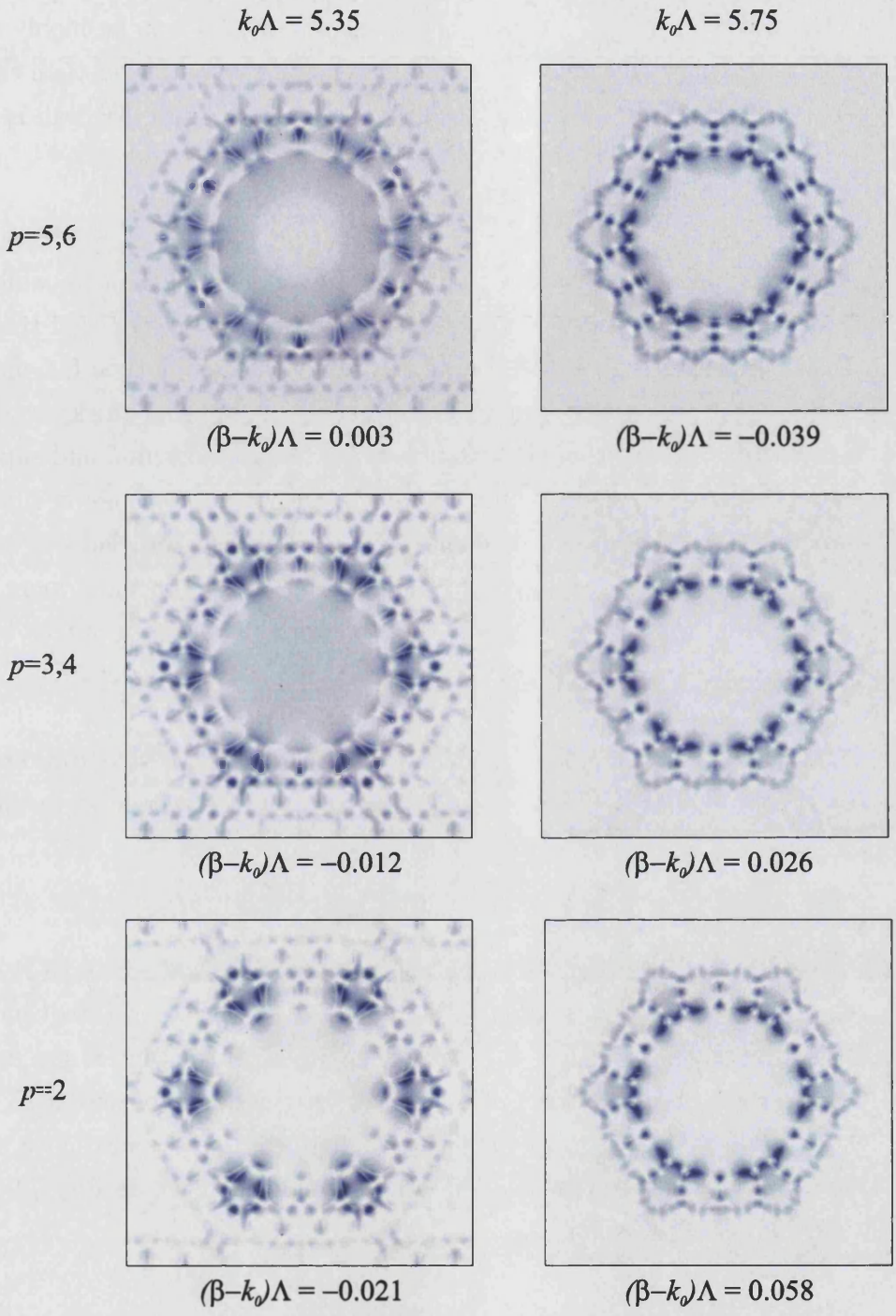


Figure 7.8: Plots of the axial Poynting vector (normalised to unity over the supercell, and shown on a linear scale) for selected surface modes of the structure with $t = 0.05\Lambda$ near to the air-line, at frequencies above and below the 'clean' region of Fig. 7.7. Each row of the figure shows two modes of the same symmetry type. The intensities of the two modes of each doubly-degenerate pair have been added to show their structure more clearly.

such that the wall forms an AR Fabry-Perot 'reflector' for the radial component of the wavevector [126]. This approach has been shown to be highly effective in silica fibres, for which the AR core wall is significantly thicker than the width of struts in the cladding [126]. The thickness of the thinnest AR wall is given by

$$t = \frac{\pi}{2(n^2 k_0^2 - \beta^2)^{1/2}}, \tag{7.4}$$

where n is the refractive index of the core wall and t is the core wall thickness. If we use $n = 2.4$, we find anti-resonance at the air-line for wall thicknesses in the range $0.12\Lambda < t < 0.14\Lambda$ for values of $k_0\Lambda$ within the type-II bandgap. This range of AR core walls is much thicker than the optimal thickness determined numerically; in fact, the AR thicknesses are far from optimal and support many surface modes. This can be understood by noting that, unlike in silica, the AR core wall is much thinner than the strut thickness in the cladding (which is of order 0.2Λ). In this case, glass in the distorted region becomes more important than that in the core ring itself and the AR core wall approach can not be expected to be reliable.

The six lowest-order guided modes of the 'optimal' PCF structure are shown in Fig. 7.9 at $k_0\Lambda = 5.5$. Also given is the fraction of power in air, which is given by

$$P_{\text{air}} = \frac{\int_{\text{air}} \mathbf{S} \cdot \hat{\mathbf{z}} \, \mathrm{d}A}{\int \mathbf{S} \cdot \hat{\mathbf{z}} \, \mathrm{d}A}, \tag{7.5}$$

where $\mathbf{S}$ is the Poynting vector and the integrals are over the area A of the unit cell. For the fundamental mode, P_{air} remains over 95% over a frequency range of approximately $5.4 \leq k_0\Lambda \leq 5.7$, comprising $\approx 5\%$ of the central gap frequency at the air-line. The maximum values of P_{air} for this mode of over 98% occur over the range $5.5 \leq k_0\Lambda \leq 5.6$. Note that higher-order modes also have a relatively high fraction of power in air, with the six modes shown in Fig. 7.9 all having $P_{\text{air}} > 95\%$ at $k_0\Lambda = 5.5$.

7.5 Conclusions

The work presented in this chapter concerns the modelling of a PCF structure that is suitable for the guidance of light in the mid- to far-infrared wavelength region. In order to use the fixed-frequency plane-wave method to model these structures with large refractive index steps, it was first necessary to design an

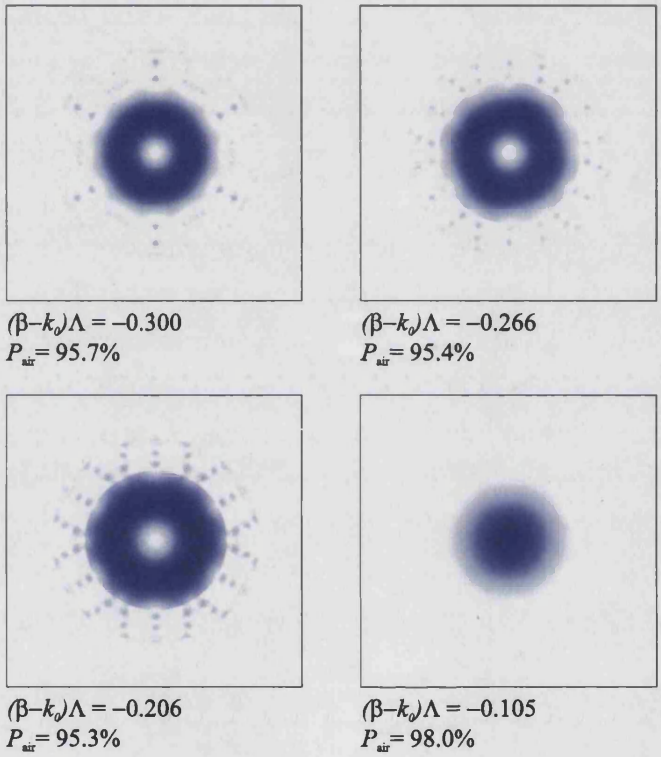


Figure 7.9: Plots of the axial Poynting vector (normalised to unity over the supercell, and shown on a linear scale) for the lowest-order air-guided modes at frequency $k_0\Lambda = 5.5$. The intensities of the two modes of each doubly-degenerate pair (at $\beta\Lambda = -0.266, -0.105$) have been added, and the colour scale used is the same as that in Fig. 7.8.

improved method of linear equation preconditioning. This method, described in Sec. 7.3, is compared to the usual 'Block/Jacobi' preconditioning method and shown to provide a considerable improvement, making the study of large index contrast PCF feasible using the plane-wave method.

When designing any hollow-core PCF, it is important to take into account the effect of surface modes that are supported by the wall of glass surrounding the core. Although it might be expected that structures comprising high-index glass are particularly susceptible to surface modes, the results of this chapter show that the effect of these modes can be reduced by an appropriate choice of core wall thickness. The 'optimal' design presented shows a fundamental mode free of surface mode crossings over a large fraction of the width of the bandgap, and we have demonstrated that the fraction of power in air in this design can be as high as 98%. No simple underlying physical reason for this behaviour (such as anti-resonance) is obvious, but the results show that detailed numerical modelling can successfully identify designs with favourable properties.

Because the 'optimal' design shown in this chapter is guided by the fabrication constraints of real PCFs, it should be possible to fabricate fibres with this structure; recent work has already demonstrated success in fabricating high-index PCF [115,117]. Similar designs also using the type-II bandgap may also be practicable for glasses with refractive indices over the range of approximately 2.0-2.8.

Chapter 8

Bend loss in low contrast bandgap-guiding PCF

The work presented in Chapters 4-7 shows improvements to the fixed-frequency plane-wave method aimed at increasing its applicability to PCF structures by attempting to overcome some of its most significant practical problems. As shown in Chapters 4 and 7, perturbation of β values caused by smoothing and convergence problems of the linear solver are both increased when the refractive index contrast in structures is increased. In this chapter, the fixed-frequency plane-wave method is instead applied to PCF structures containing *low* contrast index steps (comprising only different types of glass, rather than both air and glass), for which results are less affected by smoothing and the linear solver convergence is generally much more rapid.

Low-contrast fibres provide an interesting insight into light propagation in PCF because they can be formed from simple structures whose photonic bandstructure can be analysed and understood more easily than that of air-glass fibres. They are also more susceptible to bend loss (i.e. loss of light from the fibre core when the fibre is bent). The computational and experimental results presented in this chapter demonstrate the relationship between photonic bandstructure and bend loss in a low contrast fibre.

This chapter begins with an introduction to low contrast fibres and the reasons for studying them, and then describes how the fixed-frequency plane-wave

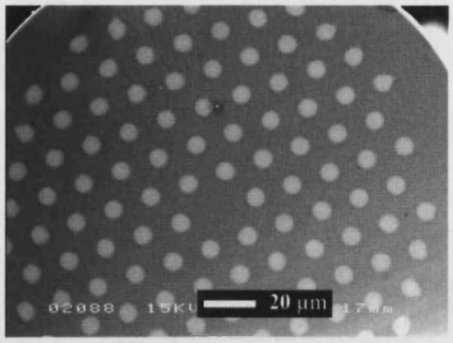


Figure 8.1: SEM image of an all-solid low contrast PCF. The background is undoped silica with $n \approx 1.458$ and the light regions, with $n \approx 1.458 + 2\%$, are germanium-doped silica rods. The fibre core consists of one missing rod. Image courtesy of Dr Feng Luan.

method has been used to determine the density of states (DOS) and position of the fundamental mode for a particular fibre that has been fabricated experimentally. Bend loss is then discussed in theoretical terms, and comparisons are made between calculated estimates of bend loss behaviour and experimental results obtained using fabricated fibres. It is demonstrated that the observed bend loss characteristics are a result of particular features of the bandstructure, and the origin of these features is also discussed. Finally, the methods developed to estimate bend loss susceptibility are also applied to an air-silica structure, and it is demonstrated that they correctly show that air-silica fibres are less susceptible to bend loss than low contrast fibres.

8.1 Introduction

Guidance in low-contrast PCF was initially demonstrated by filling the holes of an index-guiding PCF (an air-silica structure similar to that shown in Fig. 1.1a on page 13, though with larger holes) with a high-index liquid of $n \approx 1.8$ [48,129]. More recently, there has been progress in fabricating low-contrast PCF structures using glass alone. Examples include fibres made from the silicate glasses LLF1 and SF6, with refractive indices at visible wavelengths of $n \approx 1.54$ and $n \approx 1.79$ (index step 16%) [49]; and fibres made from doped silica glass with a typical index contrast of a few percent [50] and as low as 1% [130, 131]. The fibre described later in this chapter is a typical all-solid low contrast fibre. Its structure, showing the arrangement of high- and low-index regions, is visible in Fig. 8.1.

Low contrast fibres have several potential applications. The high-index liquid used in Ref. 48 had a refractive index that varied with temperature T as $\mathrm{d}n/\mathrm{d}T \approx -10^{-3}\ \mathrm{K}^{-1}$, and this type of liquid-filled fibre could prove useful to make temperature sensors or devices with easily-tunable optical properties. Recently a three-level neodymium laser has been demonstrated with a laser transition at 890-950 nm, by using an all-solid bandgap fibre to suppress undesirable competition from a four-level transition [132]. This may be useful in applications such the study of optical absorption by water vapour and for frequency doubling [132]. Low contrast fibres are likely to find further similar applications as spectral filters [133]. It has also been noted that all-solid fibres are attractive because they can be easier to fabricate and splice owing to their lack of holes [50].

In addition to their uses in potential practical applications, low contrast fibres are a valuable tool to investigate the phenomenon of bend loss. Bend loss has been investigated extensively in conventional optical fibres [10,11,134,135] and to some extent in index-guiding PCF [136]. It has received less attention in bandgap-guiding PCF, largely because it is usually negligible in hollow-core PCF [137]. However, it has been shown that it is more significant in low contrast fibres [131]. This makes low contrast fibres an ideal tool for the study of bend loss in bandgap-guiding PCF.

Because low contrast fibres can comprise simple structures—often a periodic lattice of circular raised-index regions ('rods'), with one missing rod to create a core, as shown in Fig. 8.1—their guidance properties are more easily understood than those of hollow-core PCF (see, for example, Ref. 138 and Sec. 8.5). Nevertheless, guidance in low contrast fibres is still a result of a photonic bandgap and, therefore, it is hoped that the insights into guidance gained from the study of low contrast fibres may prove useful in understanding hollow-core PCF. In Sec. 8.5 below, a simple analysis is presented that explains the origin of specific features in the bandstructure of a low contrast fibre.

8.2 Modelling an all-solid low contrast fibre

In this section an all-solid low contrast fibre that has been fabricated experimentally is described. The results of modelling of this fibre using the fixed-frequency plane-wave method are then given, together with a method to extract from the results the rate of decay of the fundamental mode into the fibre cladding.

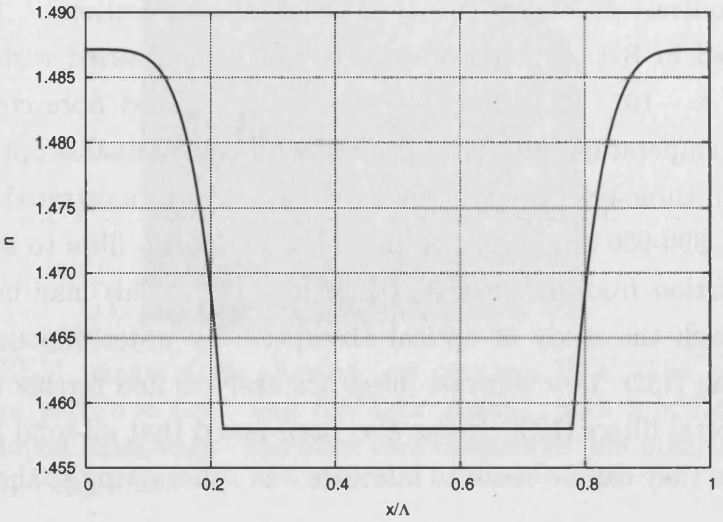


Figure 8.2: Refractive index profile $n(x)$ given by Eq. (8.1) across one edge of the cladding unit cell along $\mathbf{R}_1$ from $(0,0)$ to $(\Lambda,0)$.

8.2.1 Fibre structure

The fibre shown in Fig. 8.1 comprises high-index doped rods in a background of refractive index $n_0 = 1.458$. Each rod has a nominal index profile (based on manufacturers' data) given at radius ϱ from its centre by:

$$n(\varrho) = n_0 \left(1 + \Delta n_0 \left[1 - \left(\frac{\varrho}{\varrho_\mathrm{d}} \right)^{4.7} \right] \right), \tag{8.1}$$

where $\varrho_\mathrm{d} = 0.22\Lambda$ is the radius of the doped rod and $\Delta n_0 = 2\%$. The pitch, Λ, was measured to be $15.2\,\mu$m. Fig. 8.2 shows the refractive index profile across one edge of the cladding unit cell containing a doped rod in each corner. The calculations presented in the remainder of this section, and the experimental results in Sec. 8.4, concern this fibre structure.

8.2.2 Plane-wave calculations

To determine the density of states (DOS) for the cladding structure, the fixed-frequency plane-wave method was used with an FFT grid size of 32×32 and smoothing $\Gamma = 0.01\Lambda$, and 127 $\mathbf{k}$-points in the irreducible wedge of the first Brillouin zone[1]. Note that because the index profile shown in Fig. 8.2 is graded, the structure is 'smoothed' by definition and only the gradient of $n(\varrho)$ is dis-

[1]See Appendix E for the definition and method of calculation of the DOS.

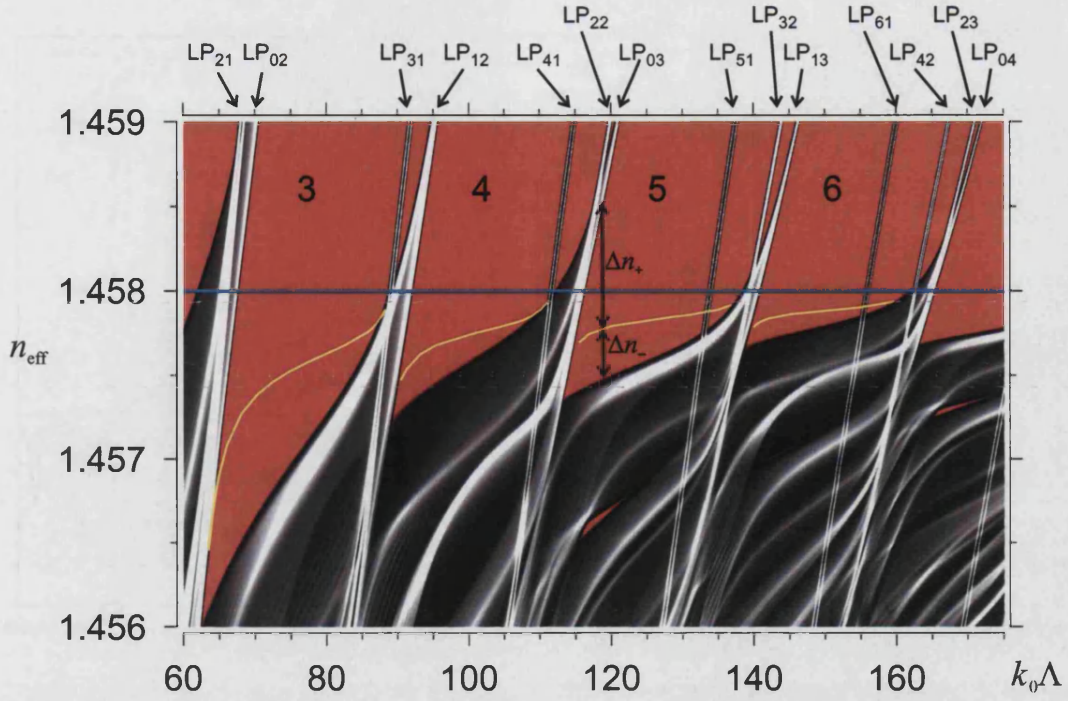


Figure 8.3: Calculated DOS for a triangular lattice of graded-index rods (see Eq. (8.1) and Fig. 8.2) in a low-index background of $n_0 = 1.458$. The axes are effective index $n_{\text{eff}} = \beta/k_0$ and normalised frequency $k_0\Lambda$, and the horizontal line is drawn at $n_{\text{eff}} = n_0$. The colour scheme is as in Fig. 7.1 on page 128. The yellow line is the locus of the fundamental core-guided mode as determined by a supercell calculation. Also given are mode designations LP_{lm} (see Sec. 8.5) and effective index mismatches Δn_- and Δn_+, which are described in the text of Sec. 8.3. The bandgaps are numbered in order of increasing $k_0\Lambda$. For ease of comparison with experiment (see Sec. 8.4), only bandgaps 3-6 are shown.

continuous at $\varrho = \varrho_{\text{d}}$; as a result, the structure (and calculated values of β) is largely unaffected by Gaussian smoothing. The calculated DOS for the structure is shown in Fig. 8.3.

In addition to the DOS, it is useful to know the position of the fundamental guided mode in the fibre core and, for later use, the field profile in the core. In order to ascertain this, an 8×8 supercell was used with FFT grid size 256×256 and smoothing $\Gamma = 0.01\Lambda$, and a defect was introduced by the omission of a single doped rod. The locus of the fundamental mode is shown superimposed on the density of states of the cladding in Fig. 8.3.

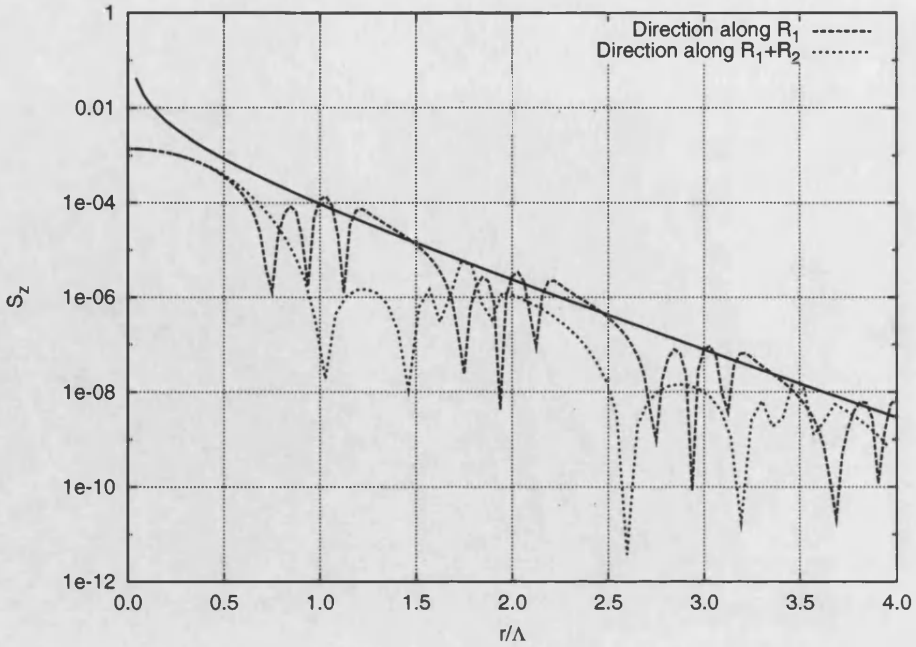


Figure 8.4: Plot of the axial Poynting vector S_z of the fundamental mode (normalised to unity over the unit cell) at $k_0\Lambda = 70$ along the directions of $\mathbf{R}_1$ and $\mathbf{R}_1 + \mathbf{R}_2$ (see Fig. 2.1 on page 34 for definitions of the supercell unit vectors $\mathbf{R}_1$ and $\mathbf{R}_2$). The horizontal axis shows $r = \sqrt{x^2 + y^2}$ for $r \leq 4\Lambda$. The solid line is the approximate fit $S_z = 1.9 \times 10^{-3} \exp(-3.0r/\Lambda)/(r/\Lambda)$ (see text).

8.2.3 Fundamental mode decay rate

For later use in quantifying bend loss, it is also possible to use the supercell approximation to calculate the decay rate of the fundamental mode into the cladding[2]. To obtain a visual representation of the field, the eigenvector of $\mathbf{h}_t$ corresponding to the fundamental mode is used to calculate the axial component of the Poynting vector S_z (as outlined in Sec. 3.1), and S_z is then plotted as a function of $r = \sqrt{x^2 + y^2}$ along given directions in the unit cell. An example of such a plot is given in Fig. 8.4.

In a conventional optical fibre of core refractive index n_{co} and radius a, and cladding refractive index n_{cl}, the Poynting vector of the fundamental mode in the cladding ($r \geq a$) has r-dependence of the form [11]:

$$S_z(r) \propto \left\{ c_1 K_0^2(Wr/a) + c_2 K_2^2(Wr/a) + c_3 K_0(Wr/a) K_2(Wr/a) \right\}, \qquad (8.2)$$

[2]Note that this decay rate is fundamentally linked to the confinement loss discussed in Sec. 2.3.2. A faster rate of decay into the PCF cladding gives a smaller field amplitude at the outside edge of the cladding, and consequently a lower confinement loss.

where $W = a\sqrt{\beta^2 - k_0^2 n_{\mathrm{cl}}^2}$, K_ν are modified Bessel functions of the second kind, and c_1, c_2 and c_3 depend on refractive index and frequency but not on r. If r is sufficiently large, the Bessel functions can be replaced by their asymptotic forms for large argument, given by [11]:

$$K_\nu(\zeta) = \sqrt{\frac{\pi}{2\zeta}} e^{-\zeta} \left(1 + \frac{4\nu^2 - 1}{8\zeta} + \ldots\right), \tag{8.3}$$

where $\zeta \in \mathbb{C}$. Substitution into Eq. (8.2) leads to the asymptotic form for S_z:

$$S_z \sim \frac{e^{-2Wr/a}}{r}. \tag{8.4}$$

If the decay rate of the fundamental mode in PCF also follows the same functional form, it should be possible to fit expressions of the form $S_z = \alpha \exp(-2\gamma r)/r$ (with α, γ as fitting parameters) to the axial Poynting vector. An example of this fitting, which has been done by eye only, is shown in Fig. 8.4.

8.3 Bend loss

In order to understand bend loss in bandgap-guiding PCF, it is first necessary to consider its origin in more general terms. A simple approach using geometrical optics is presented in this section, which is sufficient to demonstrate the origin of resonant coupling between the fundamental mode and radiation modes. To first order, this analysis correctly reproduces the result of more rigorous approaches [139, 140].

Consider the two equal lengths of fibre shown in Fig. 8.5, one of which has been bent through an angle $\theta = L/R$. The distance along the fibre for a light ray along the dotted line in each case is L, but in the bent fibre the path length for a ray at radial distance $R + r$ is $(R + r)\theta$. The path is therefore either lengthened or shortened relative to the dotted line by $\delta z = r\theta$.

The z-dependence of fields in a uniform, straight fibre takes the form $e^{i\beta z}$. We can take the bend into account by treating the fibre as straight but introducing an r-dependence into β such that the phase shift over the length of fibre shown

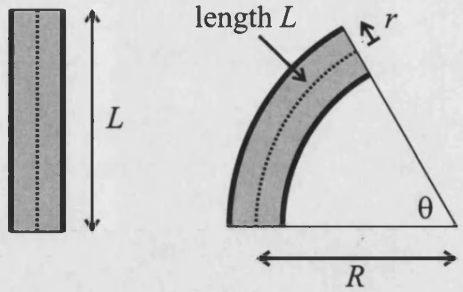


Figure 8.5: Section of straight fibre of length L, and the same fibre bent through angle $\theta = L/R$ with bend radius R and distance r from the axis of the fibre marked. The dotted line shows $r = 0$, i.e. the path with length L in both the straight and bent fibre sections.

in Fig. 8.5, βz, is unchanged:

$$\delta(\beta z) = \beta\delta z + z\delta\beta \tag{8.5a}$$

$$= \beta r\theta + R\theta\delta\beta, \tag{8.5b}$$

and hence $\delta\beta = -\beta r/R$ if $\delta(\beta z) = 0$. It is more convenient, however, to consider β for the mode to be constant and consider instead the bend to cause a local change in refractive index given by

$$n(r) = n_{\mathrm{s}}(r)\left(1 + \frac{r}{R}\right), \tag{8.6}$$

where $n_{\mathrm{s}}(r)$ is the refractive index of the straight fibre before bending[3]. Figs. 8.6a and 8.6b show how this variation of refractive index can allow light to escape to the *outside* of a bend in a conventional optical fibre. When the fibre is straight, the effective index n_{fm} of the fundamental mode is greater than that of any state in the cladding, and light is confined to the core. Bending the fibre skews the index profile in the way described by Eq. (8.6) and, if the index is raised enough on the outside of the bend to match n_{fm}, resonant tunnelling can occur and light escapes from the core at the radiation caustic (marked in Fig. 8.6b).

In a bandgap-guiding PCF, the same variation in refractive index with r applies, but the fundamental mode is bounded by cladding states at both edges of the bandgap: there exist cladding states with an effective index higher than n_{fm} as well as lower. In this case, as shown in Fig. 8.6d, skewing the index profile by bending can allow resonant tunnelling to either the outside or the inside of the bend. As discussed below, bend loss tends to be dominated by the nearest band

[3]There is also a correction due to the elasto-optic effect [131] but this is neglected here.

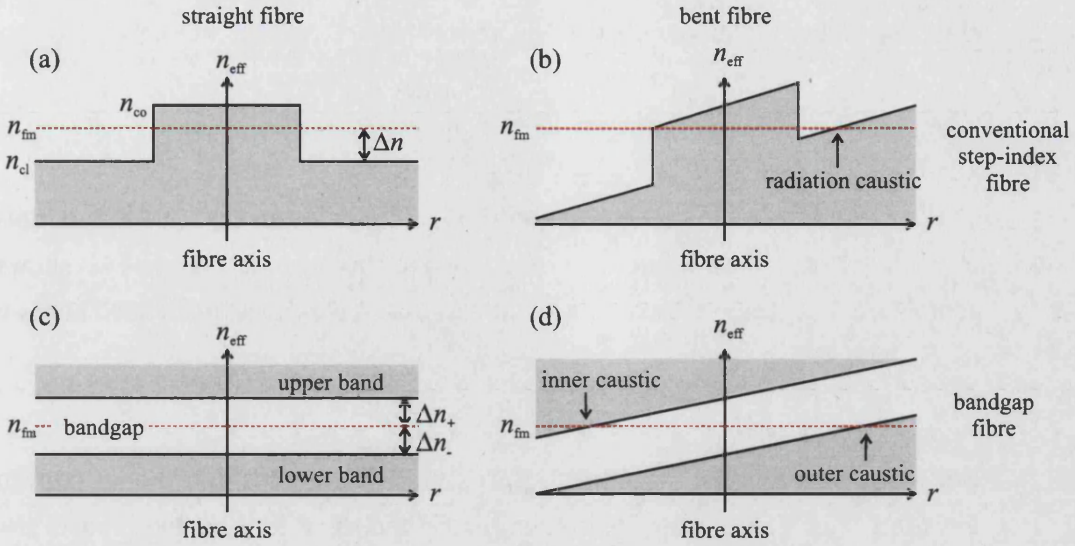


Figure 8.6: Variation of effective index n_{eff} with distance r along the direction perpendicular to the fibre axis. The index n_{fm} of the fundamental core-guided mode is marked as a dashed red horizontal line. A conventional step-index fibre is shown when (a) straight and (b) bent; the mismatch Δn between n_{fm} and the cladding index n_{cl} is marked, and the core refractive index is n_{co}. The equivalent diagrams for a bandgap-guiding PCF are shown in (c) and (d); grey areas mark cladding bands. Note that light can escape either to the inside or outside of the bend in (d), but only to the outside in (b).

edge, i.e. the upper band edge if $\Delta n_+ < \Delta n_-$ ($\Delta n_\pm$ defined in Figs. 8.3 and 8.6) and the lower band edge otherwise.

8.3.1 Quantifying bend loss

Love [141] uses the term 'critical bend radius' R_{c} to describe the approximate bending radius at which bend loss becomes significant, for conventional optical fibres. Although in practice bend loss has a gradual onset (see Fig. 8.8 on page 158), the concept of critical bend radius is useful in comparing the bend loss of a fibre at different frequencies. The derivation of an approximate expression for R_{c} from Ref. 141 is outlined here and it is then extended to make it more applicable to bandgap-guiding PCF.

The characteristic length scale z_{c} for coupling between the fundamental mode (with propagation constant β) and cladding modes in a conventional fibre with

cladding index n_{cl} is given by

$$z_{\mathrm{c}} = \frac{2\pi}{|\beta - k_0 n_{\mathrm{cl}}|}. \tag{8.7}$$

If the field in the fibre decays as $h^z \sim e^{-\gamma r}/\sqrt{r}$ (see Eq. (8.4) and note that $S_z \sim |h^z|^2$), there is a characteristic field decay rate γ. Ref. 141 shows by a geometrical argument involving comparisons of length scales that this is related to R_{c} by

$$R_{\mathrm{c}} \approx \gamma \frac{z_{\mathrm{c}}^2}{2}. \tag{8.8}$$

Eq. (8.8) can be used to obtain approximate values for R_{c}. For a conventional fibre, for which Eq. (8.4) gives an explicit expression for the decay rate $\gamma = W/a$, the critical bend radius is given by

$$R_{\mathrm{c}} \approx \frac{W}{2a}\left(\frac{2\pi}{\beta - k_0 n_{\mathrm{cl}}}\right)^2 \tag{8.9a}$$

$$\approx \frac{W}{2a}\left(\frac{4\pi k_0 a^2 n_{\mathrm{co}}}{W^2}\right)^2 \tag{8.9b}$$

$$\approx \frac{8\pi^2 k_0^2 a^3 n_{\mathrm{co}}^2}{W^3}, \tag{8.9c}$$

where Eq. (8.9b) follows from Eq. (8.9a) by multiplying the numerator and denominator of the bracketed fraction by $(\beta + k_0 n_{\mathrm{cl}})$ and employing the weak guidance approximation $\beta \cong k_0 n_{\mathrm{co}} \cong k_0 n_{\mathrm{cl}}$. This derivation and an equivalent of Eq. (8.9c) may be found in Ref. 141 but the remainder of this section is the work of the present author.

In bandgap-guiding PCF, the characteristic length scale z_{c} of Eq. (8.7) may be defined for coupling either to the cladding states with lower effective index than that of the fundamental mode (as in conventional fibres), or to those with higher effective index. Using Eqs. (8.7) and (8.8), the critical bend radius for coupling to the upper band edge R_{c}^+ or to the lower band edge R_{c}^- is given by

$$R_{\mathrm{c}}^{\pm} \approx \frac{2\pi^2 \gamma}{(k_0 \Delta n_{\pm})^2}, \tag{8.10}$$

where $\Delta n_{\pm}$ is the offset between the fundamental mode and the relevant band edge, as defined in Figs. 8.3 and 8.6. To find $R_{\mathrm{c}}^{\pm}$, the decay rate γ must be found. One way to determine γ is to use supercell calculations, plotting fields and fitting

an exponential decay curve, as described above in Sec. 8.2.3. These values of γ may be substituted directly into Eq. (8.10) to obtain $R_c^{\pm}$.

Another possible way to obtain R_c is to assume that the decay rate in PCF is dominated by the offset between the fundamental mode and the nearest band edge (with effective index n_{edge}), and that Eq. (8.9c) remains valid given an appropriate redefinition of W, viz. $W = k_0\Lambda|n_{fm}^2 - n_{edge}^2|^{1/2}$. Substitution into Eq. (8.9c) then gives[4]:

$$R_c \approx \frac{8\pi^2 n_0^2}{k_0 \left|n_{fm}^2 - n_{edge}^2\right|^{3/2}}. \tag{8.11}$$

Eq. (8.11) shows that the critical bend radius is directly related to the offset between the fundamental mode and the nearest band edge: the smaller the offset, the greater the critical bend radius (i.e. the more susceptible the fibre is to bend loss at that particular frequency). This allows estimates of the critical bend radius to be made directly from plots such as Fig. 8.3.

In summary, there are two methods of estimating the critical bend radius. The first method requires determination of the decay rate of the fundamental mode into the PCF cladding. The second method assumes that the decay rate is dominated by the offset between the fundamental mode and the nearest band edge; this has the advantage of being more easily calculated because no fitting of decay curves is needed, but it relies on an assumption to calculate the decay rate, and so may be less accurate. In Sec. 8.4 below, the two methods are compared with experimental data.

8.4 Comparison with experiment

The approximate expression of Eq. (8.11) suggests that a fibre is more susceptible to bend loss at frequencies where the offset between the fundamental mode and the nearest band edge is small. At frequencies where the offset is larger, the fibre should be less susceptible to bend loss. Examining the position of the fundamental mode in Fig. 8.3, it is clear that there is an alternating pattern in the bandstructure: Δn_- is in general greater in odd-numbered bandgaps than in even-numbered bandgaps. Although susceptibility to bend loss is determined by the smaller of Δn_- and Δn_+, the dispersion of the cladding states is such that,

[4]Note that it is not valid here to define $R_c^{\pm}$ because the decay rate is a single quantity assumed to be dominated by the nearest band edge.

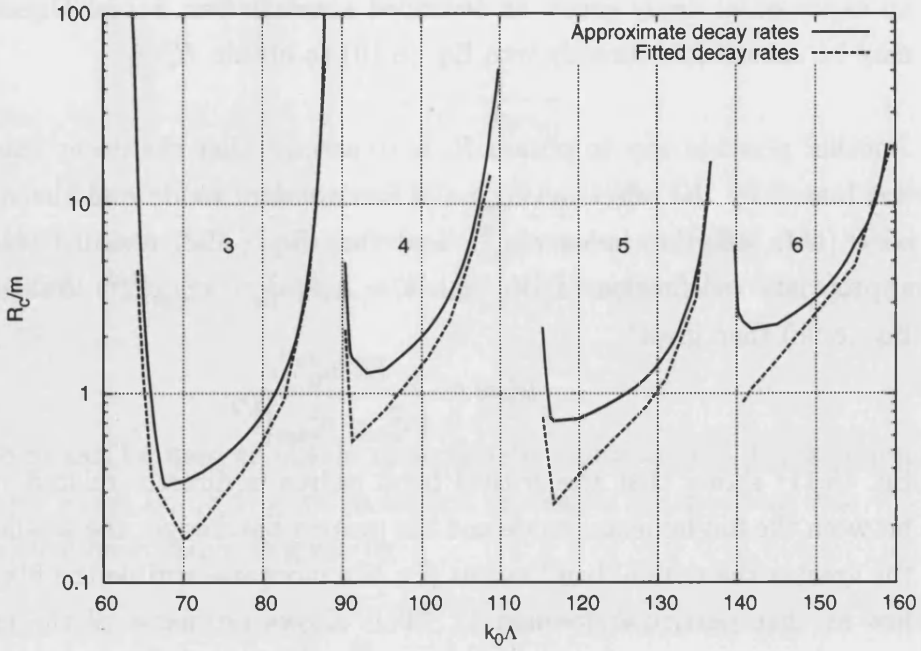


Figure 8.7: Critical bend radius R_c calculated by two methods. The solid line shows the approximation using Eq. (8.11), and the dotted line shows the result of fitting exponential decay rates to the fields obtained using the plane-wave method and using Eq. (8.10). The numbering labels the bandgaps.

over much of frequency range of each bandgap, Δn_- is smaller; Δn_+ is smaller only at the far low-frequency edge of each gap. As a result, the alternating pattern in Δn_- should manifest itself in the bend loss characteristics of the fibre as calculated using Eq. (8.11).

In Fig. 8.7, the critical bend radius is plotted as calculated by the two methods presented in Sec. 8.3.1. The approximate method using Eq. (8.11) shows the alternating pattern between bandgaps, with the odd-numbered bandgaps less susceptible to bend loss and the even-numbered bandgaps more susceptible. The more accurate method based on the decay rates obtained from supercell calculations also displays the alternating pattern, although the results are slightly different quantitatively: in general, the more accurate method predicts a greater susceptibility to bend loss than does the approximate method.

The general 'U' shape seen in Fig. 8.7 across each bandgap is a result of the increase in the offset between the fundamental mode and band edges in the middle of the bandgap relative to that at the edges of the gap. Fig. 8.7 also shows that the gradient of the R_c curves is greater at the low-frequency edge of each

bandgap than at the high-frequency edge. This suggests that, as a straight fibre is gradually bent, the frequency range over which bend loss is apparent will encroach on the low-loss transmission regions more quickly at the high-frequency edge. This asymmetry between the two bandgap edges is a result of the dispersion of the bands visible in Fig. 8.3. As the edge of a bandgap is approached, Δn_+ at the low-frequency edge decreases much more rapidly than Δn_- at the high-frequency edge.

Transmission spectra obtained from fibres having the form of that of Fig. 8.1 are shown in Fig. 8.8. The solid line in Fig. 8.8 shows the spectrum obtained when the fibre is straight, and shows clearly the low-loss regions corresponding to the bandgaps seen in Fig. 8.3. When the fibre is bent through a radius of $R = 15\,\mathrm{cm}$, the transmission through bandgaps 4 and 6 is greatly reduced, but that through bandgaps 3 and 5 is barely affected. With a bend radius of $R = 7.5\,\mathrm{cm}$, there is transmission in bandgap 3 but most light is lost through bandgaps 4,5 and 6. Also visible is the increased bend loss susceptibility of high-frequency bandgap edges: in all of the gaps shown, bend loss is more apparent at the high-frequency edges than at the low-frequency edges.

Fig. 8.8 suggests that, for example, R_c for bandgap 3 is less than 7.5 cm, because there is still significant transmission at a radius of $R = 7.5\,\mathrm{cm}$. Bandgap 5 has a critical bend radius of $\approx 7.5\,\mathrm{cm}$. This suggests that the estimates in Fig. 8.7 are too large. The method based on calculated decay rates is slightly more accurate than the approximate method, but for both methods the accuracy of the calculations is limited by the original derivation of R_c. As discussed in Sec. 8.3.1, both the definition of critical bend radius and the derivation of Eq. (8.8) are approximate and, as a result, exact quantitative agreement with experiment can not be expected.

Despite the lack of quantitative predictions of R_c, both methods presented in Sec. 8.3.1 correctly predict the alternating pattern of bend loss susceptibility across the bandgaps seen in Fig. 8.8. They also correctly predict greater susceptibility to bend loss at the high-frequency bandgap edges, and produce estimates of the critical bend radius of the correct order of magnitude.

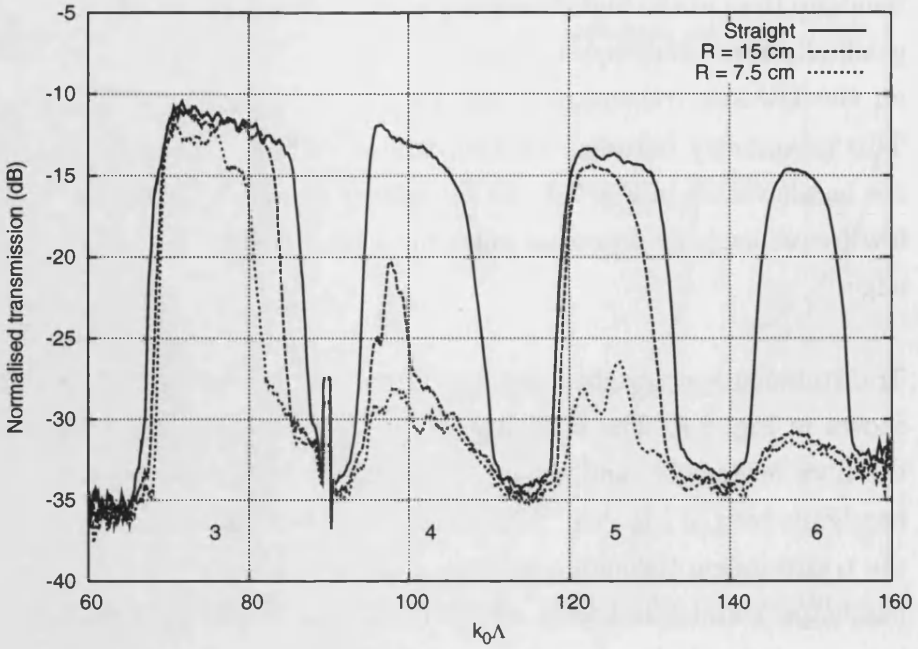


Figure 8.8: Experimental transmission spectrum for 2 m of the fibre of Fig. 8.1, when straight and when bent in a single turn around a cylinder of radius R. Two values of R are shown. The numbering on the plot labels the bandgaps for comparison with Fig. 8.3. The spike at $k_0\Lambda \approx 90$ is the supercontinuum pump source. Data courtesy of Dr Feng Luan.

8.5 Understanding the bandstructure

The pattern of bend loss susceptibility alternating across bandgaps, calculated and shown in Fig. 8.7 and confirmed experimentally in Fig. 8.8, is a result of the particular form of the DOS shown in Fig. 8.3. In this section, the structure of the DOS is analysed and the reason for the alternating pattern is explained by reference to a simple model.

The anti-resonant reflecting optical waveguide (ARROW) model [128, 138, 142, 143] describes how photonic bands are formed from the coupled resonances of individual resonators (in our case circular rods). For effective indices greater than the background index n_0 ($\beta > k_0 n_0$), these resonances are the waveguide modes of the individual rods. In the region $\beta < k_0 n_0$, the resonances are phase-matched to radiation modes and are therefore 'leaky', but they can still be identified as modes of the rods[5].

[5] A more detailed examination of the effect of leaky modes in ARROW fibres can be found in Ref. 143.

Because the index contrast in the structure previously described is low, it is possible to use the scalar approximation in which the vector equation (2.12) is replaced by its scalar equivalent, neglecting the small term involving $\nabla_t \ln n^2$:

$$\left(\nabla_t^2 + n^2 k_0^2\right) \Psi = \beta^2 \Psi, \tag{8.12}$$

where Ψ can be h^x or h^y [11]. The solutions of this equation for a circular rod can be designated LP_{lm}, where l and m are integers describing respectively the azimuthal and radial variation of the field [11]. In Fig. 8.3, the bands formed from coupled rod modes are labelled using the LP_{lm} notation to identify each rod mode from which the bands are formed.

The width of the bands seen in Fig. 8.3 is determined by the overlap between the modes of adjacent rods in the structure. Where $\beta > k_0 n_0$, the fields of each rod are strongly confined and there is little overlap. Consequently, the bands are narrow. As β decreases and the mode becomes leaky at $\beta = k_0 n_0$, the strength of coupling increases and the bands broaden, eventually closing the bandgap completely. However, it is clear from Fig. 8.3 that the rate at which broadening occurs as β decreases is not constant for all modes. In particular, the LP_{12} and LP_{13} bands, which define respectively the high-frequency edges of bandgaps 3 and 5, broaden comparatively slowly relative to LP_{03} and LP_{04}, which determine the high-frequency edges of bandgaps 4 and 6. These high-frequency edges determine Δn_- for bandgaps 3-6, and therefore the alternating bend loss behaviour across these bandgaps can be attributed to the way in which the individual rod modes couple together and broaden at different rates below $\beta = k_0 n_0$.

To understand why the rod modes broaden differently, it is necessary to consider the rate at which the fields of the rod modes decay away from the rod, which in turn determines how strongly the modes couple together[6]. If the mode decays slowly outside the rod, it will couple more strongly to adjacent modes than if it decays quickly and the field is strongly confined to the rod. The modal fields can be determined from Eq. (8.12). If we substitute the separated field $\Psi(\varrho, \phi) = R(\varrho) e^{il\phi}$ (where ϱ,ϕ are local polar coordinates for the rod) into

[6]A similar approach is used in the tight-binding approximation of solid-state physics. In this approximation, electronic bands are formed from the coupled states of single atoms, and the width of the bands so formed is determined by the strength of coupling between atoms [75].

Eq. (8.12), the resulting equation for the radial function $R(\varrho)$ is [11]:

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}\varrho^2} + \frac{1}{\varrho}\frac{\mathrm{d}}{\mathrm{d}\varrho} + k_0^2 n^2(\varrho) - \frac{l^2}{\varrho^2}\right) R(\varrho) = \beta^2 R(\varrho). \tag{8.13}$$

The modes of the rod can be found by solving Eq. (8.13) piecewise inside and outside the rod, and using the appropriate boundary conditions to determine β. However, the decay rates outside the rod can be obtained more simply at cutoff (where $\beta = n_0 k_0$). In this case, the modal fields are given by solutions to

$$\left(\varrho^2\frac{\mathrm{d}^2}{\mathrm{d}\varrho^2} + \varrho\frac{\mathrm{d}}{\mathrm{d}\varrho} - l^2\right) R(\varrho) = 0, \tag{8.14}$$

where Eq. (8.13) has been multiplied by ϱ^2 and the substitutions $\beta = n_0 k_0$ and $n(\varrho) = n_0$ have been made. Eq. (8.14) can then be re-written as:

$$\varrho\frac{\mathrm{d}}{\mathrm{d}\varrho}\left(\varrho\frac{\mathrm{d}R}{\mathrm{d}\varrho}\right) = l^2 R. \tag{8.15}$$

The solutions to Eq. (8.15) that are well-behaved as $\varrho \to \infty$ take the general form $R(\varrho) \propto \varrho^{-l}$ (constant for $l = 0$), as can be confirmed easily by substitution. Modes with $l = 0$ therefore do not decay at all outside the rod at cutoff, and modes with $l > 0$ decay progressively more quickly as l increases. We therefore expect the bands formed from $l = 0$ modes to be broader than those formed from $l = 1$ modes, and those from higher-l modes to be narrower still.

Because the $l = 0$ modes broaden more than the $l = 1$ modes, and the $l = 0$ modes define the edges of the even-numbered bandgaps, the offset Δn_- is smaller in these gaps and consequently the susceptibility to bend loss is greater. The odd-numbered bandgap edges are defined by $l = 1$ modes, which broaden more slowly and hence give rise to a greater Δn_- and less susceptibility to bend loss. In comparison, higher-l modes appear to have little effect: the locus of the fundamental mode is barely perturbed by these modes, and the bands formed from them are very narrow. This can be explained by the fast decay of the rod mode fields for $l > 1$, leading to very weak coupling [144]. In Fig. 8.8 the dominant $l = 0$ and $l = 1$ bands mask any effects of the $l = 3$ and $l = 4$ modes, and the narrow features resulting from $l = 5$ and $l = 6$ modes are not visible in the spectrum.

8.6 Bend loss in air-silica PCF

The methods developed in Sec. 8.3.1 to quantify bend loss have been shown to be successful in describing the bend loss characteristics of a fabricated low contrast fibre. However, although they are strictly valid only for low contrast structures, the general ideas are equally applicable to air-silica fibres. Bend loss is known to be less significant in air-silica hollow-core PCF [137], so it is worthwhile to consider briefly whether the methods developed here correctly predict this.

In Sec. 6.2, a typical 7-cell hollow-core PCF design is presented and used as a test system for the plane-wave method in generalised curvilinear coordinates. Figs. 6.5, 6.6 and 6.7 (page 110 onwards) fully define this structure, and the mode trajectories are plotted in Fig. 6.10 on page 116. As Fig. 6.10 contains information on both the positions of the band edges and the locus of the fundamental mode, it contains the information on n_{fm} and n_{edge} necessary to calculate the critical bend radius with Eq. (8.11). The method of calculating decay rates presented in Sec. 8.2.3 can also be used to determine the decay rate of the fundamental mode, and therefore both of the two methods presented in Sec. 8.3.1 can be applied.

In order to obtain critical bend radii in metres, the pitch of the fibre must be known. In the remainder of this section, the pitch $\Lambda = 3.5\,\mu\mathrm{m}$ has been chosen such that the fibre would be suitable for guidance at 1550 nm, a commonly-used telecommunications wavelength that coincides with a low-loss window in silica [10]. This wavelength was also used in the experimental measurements of bend loss of a similar fibre in Ref. 137.

Fig. 8.9 shows the critical bend radius as calculated using the two methods of Sec. 8.3.1. Note that, unlike in Fig. 8.7, the method using calculated decay rates gives slightly larger values of R_{c} than the approximate method; the reason for this difference is not known. However, both methods give similar results and, as expected, the calculated critical bend radii are much less than those of the low contrast fibre. The minimum value of R_{c} in Fig. 8.9 is approximately 2 mm, compared with a range (dependent on bandgap) of approximately 0.3-2 m for the low contrast fibre. Although the pitches of the two fibres differ by a factor of ≈ 4, this difference is much less than the calculated difference in critical bend radii and so pitch alone does not account for the difference.

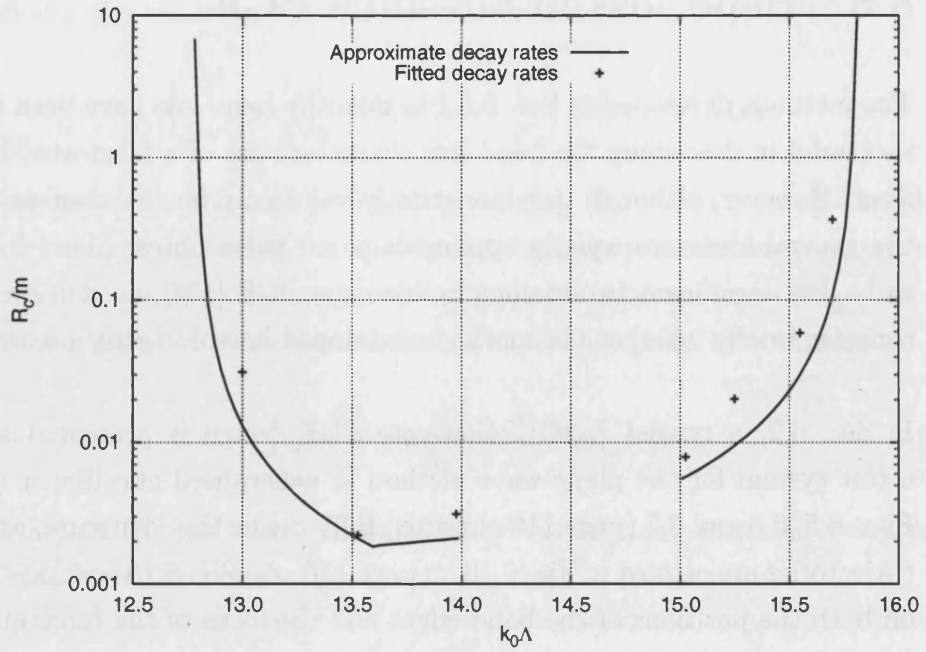


Figure 8.9: Calculated critical bend radii for the air-silica fibre presented in Sec. 6.2, with a pitch of $\Lambda = 3.5\,\mu$m. The critical bend radius has not been plotted in the region $14 < k_0\Lambda < 15$ because the fundamental mode has an anti-crossing with a surface mode in that frequency range, and the methods of Sec. 8.3.1 are applicable only to the fundamental mode.

The experimentally-measured bend loss of a fibre with similar structure and pitch to that described in this section is given in Ref. 137. Bend radii between 2 mm and 20 mm were used and, even at the tightest bend radius of 2 mm, the only bend loss observed was at the far high-frequency edge of the bandgap[7]. This suggests that the critical bend radius over much of the bandgap must be less than 2 mm, and that therefore the calculations presented in Fig. 8.9 represent an over-estimate of R_c. This over-estimation was also apparent in the results of Sec. 8.4 for low contrast fibres. However, the general finding that air-silica hollow-core PCF exhibits much less susceptibility to bend loss than low contrast fibres is correctly reproduced.

8.7 Conclusions

In this chapter the modelling of a low contrast bandgap-guiding PCF has been presented. Although in general the modelling of low contrast structures does not

[7]Maintaining such a small bend radius for a prolonged period makes the fibre likely to break, and therefore such fibres can be said to suffer no bend loss at any practical bend radius [137].

present any significant computational challenges for the fixed-frequency plane-wave method, the determination of field decay rates leading to a numerical estimate of critical bend radius is a novel use of the method. The fibre structure used in the plane-wave calculations was chosen to be a close match to a fabricated fibre, in order that comparisons could be drawn between the results of modelling and experimental data.

Two methods of estimating the bend loss characteristics of a PCF structure have been developed based on a previous analysis of bend loss in conventional optical fibres. One method uses an approximation to the decay rate of the fundamental mode into the fibre cladding, and suggests that the offset between the fundamental mode and nearest band edge (easily obtained directly from plane-wave calculations) determines the critical bend radius. Another method that has been developed uses calculated decay rates (again obtained using the plane-wave method), and is shown to be in better quantitative agreement with bend loss measurements on the fabricated low contrast fibre. Both methods correctly reproduce an alternating pattern of bend loss susceptibility across the bandgaps of the PCF structure and predict critical bend radii of the correct order of magnitude. However, as both methods ultimately rely on approximations, it is clear that a more detailed theory is required in order to obtain better quantitative agreement between calculations and experiment.

The alternating pattern of bend loss susceptibility across the bandgaps of the low contrast fibre is accompanied by alternation of the calculated depth of bandgaps seen in the density of states. A simple explanation has been presented that shows how the depth of the bandgaps is influenced by the azimuthal order (l-value) of the modes of the rods which, when coupled together, form the bands seen in the DOS. This explains the origin of the pattern of bend loss seen in experimental spectra.

The methods developed in this chapter to calculate decay rates and estimate bend loss susceptibility in low contrast fibres are equally applicable to hollow-core PCF. In Sec. 8.6, results are presented that show the application of these methods to a typical air-silica hollow-core fibre. The calculations correctly predict that the air-silica fibre is much less susceptible to bend loss, in good agreement with experimental findings. Although hollow-core bandgap fibres are more complicated

to analyse[8], it is also possible that, in a similar way to that of Sec. 8.5, identification of the specific modes defining band edges in hollow-core PCF could lead to better understanding of the origin of specific features in their densities of states.

[8]In particular, the non-circular and connected high-index resonators of hollow-core PCF (see, e.g., Fig. 7.3a on page 130) are more difficult to analyse than circular rods.

Chapter 9

Conclusions

The work presented in this thesis has primarily concerned improvements and applications of the fixed-frequency plane-wave method for the modelling of photonic crystal fibre. The conclusions drawn from this work, together with possible areas of further research arising from it, are described in this chapter.

In Chapter 4, a simple 1D 'PCF' structure was presented and analysed using the fixed-frequency plane-wave method. The results of this analysis, which can be generalised to 2D, showed that the convergence of the plane-wave method depends upon the number of real-space sampling points in the areas of the unit cell containing the sharp interfaces between regions of different dielectric function. In the conventional plane-wave method, there are two ways in which the number of sampling points at interfaces can be increased: either the interface can be smoothed, making it wider so that it is then described by more sampling points, or the density of sampling points can be increased. Although both of these actions enhance plane-wave convergence, they introduce new problems: smoothing the dielectric function causes an unwanted perturbation of the results away from their true values, and using a greater density of sampling points (or, equivalently, greater FFT grid sizes) increases the total computation time and storage requirement.

One finding of Chapter 4 was that the fixed-frequency plane-wave method is not (at least in its conventional form) a practical method for the modelling of 2D structures containing metallic dielectrics. For this reason metallic structures

were not considered further in this thesis. However, the modelling of structures containing metals remains an interesting possibility and is worthy of further consideration in future.

The reformulation of the plane-wave method in generalised curvilinear coordinates (GCCs), presented in detail in Chapter 5, provides a way to increase the density of real-space sampling points in particular regions of the unit cell *without* increasing the smoothing or the FFT grid size. This is achieved by the use of a coordinate transformation chosen such that it gives a position-dependent sampling point density, while at the same time retaining uniformity in one coordinate system so that FFTs can still be used to allow rapid calculation. Although the method can be used with any invertible coordinate transformation, the results of Chapter 4 suggest that a transformation that increases the density of sampling points at dielectric interfaces would be beneficial in improving the plane-wave convergence. A method to create such transformations, based on the minimisation of a fictitious energy functional, was also given. In Chapter 6 this method was shown to create coordinate transformations with the desired increase in the grid point density at dielectric interfaces.

Chapter 6 showed the application of the plane-wave method in GCCs to two example structures, firstly a cladding structure comprising a simple array of air holes in glass (for which exact numerical solutions were available), and then a supercell structure representing a realistic geometry for an air-silica hollow-core PCF. In both cases the GCC method was shown to provide a considerable improvement in plane-wave convergence over the conventional plane-wave method: i.e. for a given FFT grid size the GCC method shows better convergence, with the improvement related to the amount of distortion applied. However, the improvement brought about by the method comes at a cost of worsening the convergence of the linear solver (GMRES), with greater grid distortions causing increasingly slow convergence. Consequently, the GCC method is generally not faster than the conventional plane-wave method when carrying out similarly-converged calculations. It does, though, provide a way to investigate structures with very little dielectric smoothing with a much greater certainty in the level of plane-wave convergence. This may prove to be useful in understanding, for example, the guidance properties of Kagomé fibre structures (see, e.g., Ref. 33) at high frequency, which preliminary investigations suggest may be interesting—although

not necessarily well converged with the conventional fixed-frequency plane-wave method [145].

In order to make the GCC method faster, further development of preconditioning methods to overcome the poor GMRES convergence is clearly desirable. Similarly, the use of distortions greater than those presented in Chapter 6 could be beneficial, but better preconditioning would be necessary before using such distortions because they would otherwise require prohibitively long calculation times. The GCC method has been shown to give large potential improvements (and actual improvements, when considering structures with a very narrow dielectric smoothing) over the conventional plane-wave method. However, it is clear that with further preconditioner enhancements the method could be improved considerably. With such enhancements, it is possible that the GCC method may provide the necessary level of improvement to consider carrying out calculations on structures containing metals as well as the more usual air/glass PCF structures.

As previously noted, the plane-wave method in GCCs can be used with any coordinate transformation. The fictitious energy method described in Chapter 5 has been shown to be successful, but it is not necessarily the optimal method. There is considerable freedom in the choice of energy functional, and one potentially helpful direction of future work would be to undertake a more thorough study of how the choice of energy functional affects the eventual convergence of the plane-wave method.

Chapter 7 demonstrated the application of the plane-wave method to PCF structures containing a large refractive index contrast. Because the large index contrast in these structures causes slow convergence of GMRES, it was first necessary to develop an improved method of preconditioning. The method demonstrated in this chapter involved using an exact inverse of the matrix corresponding to a smoothed dielectric structure, determined iteratively, to provide an approximate inverse of the matrix describing the required dielectric structure. The improvement brought about by using this preconditioner was discussed, and the plane-wave method was then used to investigate designs for a realistic PCF structure suitable for the guidance of light in the mid- to far-infrared. An 'optimal' structure was presented. At the time of writing, this structure has not yet been fabricated; the fabrication of such structures experimentally is a logical direction for future work in this area. As the core design demonstrated that anti-resonance

(for example) does not provide an explanation for its favourable properties, another potential direction for future work would be a more thorough analysis of core wall structure that is applicable to a range of refractive indices and cladding designs.

Finally, Chapter 8 described the use of the fixed-frequency plane-wave method to model a fibre structure with low index contrast. Two methods of estimating the susceptibility of a PCF structure to bend loss were then developed, and the estimates obtained were compared with measurements carried out on a fabricated fibre. The experimental results were well reproduced semi-quantitatively by the theoretical estimates, both for the fabricated low contrast fibre and for an example air-silica fibre which, in agreement with experiment, was shown to be considerably less susceptible to bend loss than the low contrast fibre. However, there remains much scope for a more thorough theoretical investigation into bend loss which could lead to more accurate quantitative predictions.

Chapter 8 also explained the reason for an alternating bend loss susceptibility across bandgaps in the low contrast fibre by reference to features in its band-structure. These features arise from the coupled modes of the individual rods of the fibre cladding. A natural direction for further work arises from this analysis. If the high-index rods of the PCF cladding are replaced by high-index rings with low-index centres, this will suppress modes with high radial order and leave a succession of bandgaps with edges defined by modes with $l = 1, 2, 3, \ldots$ and $m = 1$ (notation as used in Chapter 8). Because these gaps are not bounded by the $l = 0$ modes, which have been shown to be a cause of large bend loss, the resulting structure could be expected to be more resistant to bend loss. This structure has now been investigated both experimentally and by modelling in Ref. 146, and a range of different cladding geometries are now being studied using the principles described in this chapter.

Appendix A

Fields of the 1D 'PCF'

In this appendix, complete expressions are given for the fields of the 1D 'PCF' model of Chapter 4.

A.1 General structure

In the unit cell centred at $x = 0$, the fields of the 1D periodic array of dielectric slabs defined by Eq. (4.1) take the general form:

$$h(x) = \begin{cases} A_+ e^{k_1 x} + A_- e^{-k_1 x} & -\Lambda/2 < x \leq -d/2 \\ B\cos(k_2 x) + C\sin(k_2 x) & -d/2 < x \leq d/2 \\ D_+ e^{k_1 x} + D_- e^{-k_1 x} & d/2 < x \leq \Lambda/2, \end{cases} \tag{A.1}$$

where $k_1 = (\beta^2 - k_0^2 n_{\text{air}}^2)^{1/2}$, $k_2 = (k_0^2 n_{\text{dielec}}^2 - \beta^2)^{1/2}$ and $A_\pm$, B, C and $D_\pm$ are in general all complex quantities. The even symmetry of the dielectric function about $x = 0$ allows the fields to be decomposed into either even or odd solutions, such that $h(x) = h(-x)$ and $h(x) = -h(-x)$ respectively. Relabelling the coefficients and fixing the overall normalisation, these take the form (where $E_\pm$ are complex coefficients):

$$h^{\text{even}}(x) = \begin{cases} E_- e^{k_1 x} + E_+ e^{-k_1 x} & -\Lambda/2 < x \leq -d/2 \\ \cos(k_2 x) & -d/2 < x \leq d/2 \\ E_+ e^{k_1 x} + E_- e^{-k_1 x} & d/2 < x \leq \Lambda/2 \end{cases} \tag{A.2}$$

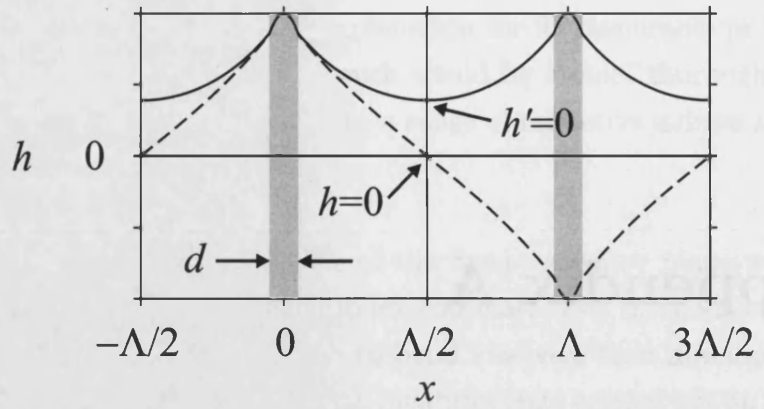


Figure A.1: Fields of the periodic 1D dielectric slab waveguide at the centre (solid line) and edge (dashed line) of the Brillouin zone, illustrated by an even TM mode.

or

$$h^{\text{odd}}(x) = \begin{cases} -E_- e^{k_1 x} - E_+ e^{-k_1 x} & -\Lambda/2 < x \leq -d/2 \\ \sin(k_2 x) & -d/2 < x \leq d/2 \\ E_+ e^{k_1 x} + E_- e^{-k_1 x} & d/2 < x \leq \Lambda/2. \end{cases} \tag{A.3}$$

For a given choice of odd/even modes, there are three unknown quantities: E_+, E_- and β. In order to determine them it is necessary to apply the relevant boundary conditions to h and h' at the interfaces between the three regions shown above. We apply the boundary conditions on h and h' at $x = d/2$ in order to determine E_+ and E_-; these coefficients are independent of whether the mode is at the centre or edge of the Brillouin zone. In order to determine β, we either choose $h' = 0$ at $x = \Lambda/2$ to obtain solutions at the centre of the Brillouin zone, or $h = 0$ at $x = \Lambda/2$ to obtain solutions at the edge of the Brillouin zone, as illustrated in Fig. A.1. The boundary conditions and resulting solutions are given in full in the following sections.

A.2 TE modes

The boundary conditions applied to obtain TE mode solutions at the Brillouin zone centre are:

$$h(d/2 - \epsilon) = h(d/2 + \epsilon) \tag{A.4a}$$

$$h'(d/2 - \epsilon) = h'(d/2 + \epsilon) \tag{A.4b}$$

$$h'(\Lambda/2) = 0, \tag{A.4c}$$

where $\epsilon \to 0$. At the Brillouin zone edge, the boundary conditions are:

$$h(d/2-\epsilon) = h(d/2+\epsilon) \tag{A.5a}$$
$$h'(d/2-\epsilon) = h'(d/2+\epsilon) \tag{A.5b}$$
$$h(\Lambda/2) = 0. \tag{A.5c}$$

The values E_+ and E_- are obtained by solving simultaneously Eqs. (A.4a) and (A.4b) using fields of the appropriate symmetry. Note that Eqs. (A.5a) and (A.5b) are identical to Eqs. (A.4a) and (A.4b), and therefore the values of $E_\pm$ are independent of whether the mode is at the centre or the edge of the Brillouin zone. $E_\pm$ is given by

$$E_\pm^{\text{even}} = \frac{1}{2}\left\{\cos\left(\frac{k_2 d}{2}\right) \mp \frac{k_2}{k_1}\sin\left(\frac{k_2 d}{2}\right)\right\}\exp\left(\mp\frac{k_1 d}{2}\right) \tag{A.6}$$
$$E_\pm^{\text{odd}} = \frac{1}{2}\left\{\sin\left(\frac{k_2 d}{2}\right) \pm \frac{k_2}{k_1}\cos\left(\frac{k_2 d}{2}\right)\right\}\exp\left(\pm\frac{k_1 d}{2}\right). \tag{A.7}$$

Solving Eq. (A.4c) together with the values of $E_\pm$ then provides the mode condition at the centre of the Brillouin zone:

$$\tanh\left(\frac{k_1(\Lambda-d)}{2}\right) = \frac{k_2}{k_1}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(even modes)} \tag{A.8}$$
$$\tanh\left(\frac{k_1(\Lambda-d)}{2}\right) = -\frac{k_1}{k_2}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(odd modes)}, \tag{A.9}$$

or at the edge of the Brillouin zone, solving Eq. (A.5c):

$$\coth\left(\frac{k_1(\Lambda-d)}{2}\right) = \frac{k_2}{k_1}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(even modes)} \tag{A.10}$$
$$\coth\left(\frac{k_1(\Lambda-d)}{2}\right) = -\frac{k_1}{k_2}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(odd modes)}. \tag{A.11}$$

A.3 TM modes

The TM mode fields are also given by Eqs. (A.2) and (A.3), but the different boundary condition on h' for TM modes results in a modification to the expressions for $E_\pm$ and the mode conditions. The boundary conditions applied to obtain

TM mode solutions at the Brillouin zone centre are:

$$h(d/2-\epsilon) = h(d/2+\epsilon) \tag{A.12a}$$
$$h'(d/2-\epsilon)/n^2_{\text{dielec}} = h'(d/2+\epsilon)/n^2_{\text{air}} \tag{A.12b}$$
$$h'(\Lambda/2) = 0. \tag{A.12c}$$

At the Brillouin zone edge, the boundary conditions are:

$$h(d/2-\epsilon) = h(d/2+\epsilon) \tag{A.13a}$$
$$h'(d/2-\epsilon)/n^2_{\text{dielec}} = h'(d/2+\epsilon)/n^2_{\text{air}} \tag{A.13b}$$
$$h(\Lambda/2) = 0. \tag{A.13c}$$

The expressions for $E_\pm$ are:

$$E^{\text{even}}_\pm = \frac{1}{2}\left\{\cos\left(\frac{k_2 d}{2}\right) \mp \frac{k_2 n^2_{\text{air}}}{k_1 n^2_{\text{dielec}}}\sin\left(\frac{k_2 d}{2}\right)\right\}\exp\left(\mp\frac{k_1 d}{2}\right) \tag{A.14}$$
$$E^{\text{odd}}_\pm = \frac{1}{2}\left\{\sin\left(\frac{k_2 d}{2}\right) \pm \frac{k_2 n^2_{\text{air}}}{k_1 n^2_{\text{dielec}}}\cos\left(\frac{k_2 d}{2}\right)\right\}\exp\left(\pm\frac{k_1 d}{2}\right), \tag{A.15}$$

and the mode conditions at the Brillouin zone centre become

$$\tanh\left(\frac{k_1(\Lambda-d)}{2}\right) = \frac{k_2 n^2_{\text{air}}}{k_1 n^2_{\text{dielec}}}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(even modes)} \tag{A.16}$$
$$\tanh\left(\frac{k_1(\Lambda-d)}{2}\right) = -\frac{k_1 n^2_{\text{dielec}}}{k_2 n^2_{\text{air}}}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(odd modes).} \tag{A.17}$$

At the edge of the Brillouin zone, the field coefficients $E_\pm$ are again unchanged, and the mode conditions are:

$$\coth\left(\frac{k_1(\Lambda-d)}{2}\right) = \frac{k_2 n^2_{\text{air}}}{k_1 n^2_{\text{dielec}}}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(even modes)} \tag{A.18}$$
$$\coth\left(\frac{k_1(\Lambda-d)}{2}\right) = -\frac{k_1 n^2_{\text{dielec}}}{k_2 n^2_{\text{air}}}\tan\left(\frac{k_2 d}{2}\right) \quad \text{(odd modes).} \tag{A.19}$$

Appendix B

Reciprocal-space matrix equation in curvilinear coordinates

In this appendix the structure of the reciprocal-space matrix equation is derived explicitly in generalised curvilinear coordinates. The reciprocal-space components corresponding to the operations required by the fixed-frequency plane-wave method are also derived.

Equivalents of the matrix elements given here, derived for use in electronic structure calculations, are published without proof in Ref. 100.

B.1 Structure of the matrix equation

Showing the h^x and h^y components explicitly, Eq. (2.18) is

$$\begin{pmatrix} R_{xx} & R_{yx} \\ R_{xy} & R_{yy} \end{pmatrix} \begin{pmatrix} h^x \\ h^y \end{pmatrix} = \beta^2 \begin{pmatrix} h^x \\ h^y \end{pmatrix}, \tag{B.1}$$

where R_{ij} are given by:

$$R_{xx} = \nabla_t^2 + n^2 k_0^2 - \left(\frac{\partial}{\partial y} \ln n^2\right) \frac{\partial}{\partial y} \tag{B.2a}$$

$$R_{yy} = \nabla_t^2 + n^2 k_0^2 - \left(\frac{\partial}{\partial x} \ln n^2\right) \frac{\partial}{\partial x} \tag{B.2b}$$

$$R_{yx} = \left(\frac{\partial}{\partial y} \ln n^2\right) \frac{\partial}{\partial x} \tag{B.2c}$$

$$R_{xy} = \left(\frac{\partial}{\partial x} \ln n^2\right) \frac{\partial}{\partial y}. \tag{B.2d}$$

Substituting the plane-wave expansion of Eq. (5.9) gives

$$\sum_{\mathbf{G}'} \begin{pmatrix} R_{xx} & R_{yx} \\ R_{xy} & R_{yy} \end{pmatrix} \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}'} \\ h^y_{\mathbf{k},\mathbf{G}'} \end{pmatrix} g^{-1/4} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}} = \beta^2 \sum_{\mathbf{G}'} \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}'} \\ h^y_{\mathbf{k},\mathbf{G}'} \end{pmatrix} g^{-1/4} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \tag{B.3}$$

We now multiply both sides of Eq. (B.3) by $\chi^*_{\mathbf{k},\mathbf{G}} = g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}}$ and integrate over the area of the unit cell:

$$\int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} \sum_{\mathbf{G}'} \begin{pmatrix} R_{xx} & R_{yx} \\ R_{xy} & R_{yy} \end{pmatrix} \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}'} \\ h^y_{\mathbf{k},\mathbf{G}'} \end{pmatrix} g^{-1/4} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}$$
$$= \beta^2 \int \mathrm{d}^2x\, g^{-1/2} \sum_{\mathbf{G}'} \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}'} \\ h^y_{\mathbf{k},\mathbf{G}'} \end{pmatrix} e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \tag{B.4}$$

Transforming the area element using $g^{-1/2}\mathrm{d}^2x = \mathrm{d}^2\xi$ and evaluating the RHS yields

$$\frac{1}{\Omega} \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} \sum_{\mathbf{G}'} \begin{pmatrix} R_{xx} & R_{yx} \\ R_{xy} & R_{yy} \end{pmatrix} \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}'} \\ h^y_{\mathbf{k},\mathbf{G}'} \end{pmatrix} g^{-1/4} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}$$
$$= \beta^2 \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}} \\ h^y_{\mathbf{k},\mathbf{G}} \end{pmatrix}. \tag{B.5}$$

Performing the matrix-vector multiplication on the LHS gives:

$$\begin{pmatrix} [\nabla_t^2 h^x]_{\mathbf{G}} + [n^2 k_0^2 h^x]_{\mathbf{G}} - [(\frac{\partial}{\partial y} \ln n^2)(\frac{\partial h^x}{\partial y})]_{\mathbf{G}} + [(\frac{\partial}{\partial y} \ln n^2)(\frac{\partial h^y}{\partial x})]_{\mathbf{G}} \\ [\nabla_t^2 h^y]_{\mathbf{G}} + [n^2 k_0^2 h^y]_{\mathbf{G}} + [(\frac{\partial}{\partial x} \ln n^2)(\frac{\partial h^x}{\partial y})]_{\mathbf{G}} - [(\frac{\partial}{\partial x} \ln n^2)(\frac{\partial h^y}{\partial x})]_{\mathbf{G}} \end{pmatrix}$$
$$= \beta^2 \begin{pmatrix} h^x_{\mathbf{k},\mathbf{G}} \\ h^y_{\mathbf{k},\mathbf{G}} \end{pmatrix}. \tag{B.6}$$

The definitions and derivations of the matrix elements appearing on the LHS of this equation are given explicitly in the following section.

B.2 Matrix elements

The LHS of Eq. (B.6) contains matrix elements for the Laplacian, gradient and n^2 ('potential') operators which will be derived in turn below.

B.2.1 Laplacian operator

The matrix element of the reciprocal-space Laplacian operator is

$$\left[\nabla_{\mathrm{t}}^2 h^i\right]_{\mathbf{G}} = \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} \sum_{j=1}^{2} \frac{\partial}{\partial x^j} \frac{\partial}{\partial x^j} g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \tag{B.7}$$

To evaluate this matrix element, it is convenient to use Green's first theorem which, for arbitrary functions ϕ and ψ in three dimensions is

$$\oint_S \phi \nabla\psi \cdot \mathrm{d}\mathbf{S} = \int_V \left[\phi\nabla^2\psi + (\nabla\phi)\cdot(\nabla\psi)\right] \mathrm{d}V, \tag{B.8}$$

where S is the closed surface bounding the volume V. Applied in 2D to Eq. (B.7) with $\phi \equiv g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}}$ and $\psi \equiv g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}$ this gives

$$\begin{aligned}\left[\nabla_{\mathrm{t}}^2 h^i\right]_{\mathbf{G}} &= \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} \sum_{j=1}^{2} \frac{\partial}{\partial x^j} \frac{\partial}{\partial x^j} g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}} \\ &= -\int \mathrm{d}^2x \sum_{k=1}^{2} \left(\frac{\partial}{\partial x^k} g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}}\right) \left(\frac{\partial}{\partial x^k} g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}\right),\end{aligned} \tag{B.9}$$

where the surface integral term vanishes owing to periodic boundary conditions. We use the chain rule to change variables in the derivative:

$$\begin{aligned}\left[\nabla_{\mathrm{t}}^2 h^i\right]_{\mathbf{G}} = -\int \mathrm{d}^2x \sum_{k=1}^{2} &\left(\sum_{p=1}^{2} \frac{\partial\xi^p}{\partial x^k} \frac{\partial}{\partial\xi^p} g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}}\right) \\ &\times \left(\sum_{q=1}^{2} \frac{\partial\xi^q}{\partial x^k} \frac{\partial}{\partial\xi^q} g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}\right).\end{aligned} \tag{B.10}$$

The derivatives are then evaluated:

$$[\nabla_{\mathrm{t}}^2 h^i]_{\mathbf{G}} = -\int \mathrm{d}^2x \sum_{k=1}^{2}\sum_{p=1}^{2} \frac{\partial \xi^p}{\partial x^k}\left\{-\frac{1}{4g}\frac{\partial g}{\partial \xi^p} - i(k_p + G_p)\right\} g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}}$$
$$\times \sum_{q=1}^{2} \frac{\partial \xi^q}{\partial x^k} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'}\left\{-\frac{1}{4g}\frac{\partial g}{\partial \xi^q} + i(k_q + G'_q)\right\} g^{-1/4} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \quad \text{(B.11)}$$

Collecting the factors of $g^{-1/4}$ and using $\mathrm{d}^2x = g^{1/2}\mathrm{d}^2\xi$ to give the change of variables, then substituting the vector potential of Eq. (5.4), gives

$$[\nabla_{\mathrm{t}}^2 h^i]_{\mathbf{G}} = -\int \mathrm{d}^2\xi \sum_{p,q=1}^{2}\sum_{k=1}^{2} \frac{\partial \xi^p}{\partial x^k}\frac{\partial \xi^q}{\partial x^k}(-A_p - ik_p - iG_p)$$
$$\times \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'}(-A_q + ik_q + iG'_q) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \quad \text{(B.12)}$$

Finally, noting that the inverse metric tensor g^{pq} is given by[1]:

$$g^{pq} = \sum_k \frac{\partial \xi^p}{\partial x^k}\frac{\partial \xi^q}{\partial x^k}, \quad \text{(B.13)}$$

the matrix element for the Laplacian can be written:

$$[\nabla_{\mathrm{t}}^2 h^i]_{\mathbf{G}} = -\sum_{p,q=1}^{2} \int \mathrm{d}^2\xi \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'}(k_p + G_p - iA_p) g^{pq} (k_q + G'_q + iA_q) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \quad \text{(B.14)}$$

B.2.2 Gradient operator

The terms involving $\ln n^2$ appearing in Eq. (B.6) are all of the form $\left[\frac{\partial \ln n^2}{\partial x^k}\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}}$. For notational convenience we assume here that $\frac{\partial \ln n^2}{\partial x^k}$ has been evaluated in advance and stored (as it is not a function of h^i), and denote it $f(\boldsymbol{\xi})$ below. The required matrix element is given by:

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} f(\boldsymbol{\xi}) \frac{\partial}{\partial x^j} g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \quad \text{(B.15)}$$

[1] It can be shown easily using the chain rule that $\sum_k g^{kq} g_{kp} = \delta^q_p$ as required when using this definition of g^{pq} and that of Eq. (5.3) for g_{pq}.

We use the chain rule to change variables in the derivative:

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} f(\boldsymbol{\xi}) \sum_{p=1}^{2} \frac{\partial \xi^p}{\partial x^j}\frac{\partial}{\partial \xi^p} g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}, \tag{B.16}$$

then evaluate the derivative:

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} f(\boldsymbol{\xi}) \sum_{p=1}^{2} \frac{\partial \xi^p}{\partial x^j} g^{-1/4} \times \left\{-\frac{1}{4g}\frac{\partial g}{\partial \xi^p} + i(k_p + G'_p)\right\} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \tag{B.17}$$

Collecting the factors of $g^{-1/4}$ and using $\mathrm{d}^2x = g^{1/2}\mathrm{d}^2\xi$ gives the change of variables:

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = \int \mathrm{d}^2\xi\, e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}} f(\boldsymbol{\xi}) \sum_{p=1}^{2} \frac{\partial \xi^p}{\partial x^j} \times \left\{-\frac{1}{4g}\frac{\partial g}{\partial \xi^p} + i(k_p + G'_p)\right\} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \tag{B.18}$$

Now using Eq. (5.4) gives the matrix element:

$$\left[f(\boldsymbol{\xi})\frac{\partial h^i}{\partial x^j}\right]_{\mathbf{G}} = i\sum_{p=1}^{2} \int \mathrm{d}^2\xi\, f(\boldsymbol{\xi}) \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} \frac{\partial \xi^p}{\partial x^j}(k_p + G'_p + iA_p) e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \tag{B.19}$$

B.2.3 'Potential' operator

The matrix element for the 'potential' operator giving n^2 is

$$\left[n^2 h^i\right]_{\mathbf{G}} = \int \mathrm{d}^2x\, g^{-1/4} e^{-i(\mathbf{k}+\mathbf{G})\cdot\boldsymbol{\xi}}\, n^2(\boldsymbol{\xi})\, g^{-1/4} \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'} e^{i(\mathbf{k}+\mathbf{G}')\cdot\boldsymbol{\xi}}. \tag{B.20}$$

Collecting the factors of $g^{1/4}$ gives

$$\left[n^2 h^i\right]_{\mathbf{G}} = \int \mathrm{d}^2\xi\, n^2(\boldsymbol{\xi}) \sum_{\mathbf{G}'} h^i_{\mathbf{k},\mathbf{G}'}\, e^{i(\mathbf{G}'-\mathbf{G})\cdot\boldsymbol{\xi}}. \tag{B.21}$$

Appendix C

Gradient of the fictitious energy functional

In this appendix, expressions for the gradient of the fictitious energy with respect to the transformation coefficients $\mathbf{x_G}$ are derived in order that they may be used in the BFGS minimisation algorithm. These have not been previously published, to the best of the author's knowledge.

Before determining the derivatives of the fictitious energy it is necessary to carry out some manipulation of derivatives of the metric tensor. Where noted below, one result presented has been previously published without proof by Gygi.

Throughout this appendix, $p, q, r, s, t, u, v \in 1, 2$.

C.1 Derivative of the metric tensor

Consider the derivative of the metric tensor given by

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = \frac{\partial}{\partial \frac{\partial x^r}{\partial \xi^s}} \sum_t g^{pt} \delta_t^q, \tag{C.1}$$

where the metric tensor on the RHS has been multiplied by the identity δ_t^q. Using the fact that g^{qu} and g_{qu} are inverses we can re-write the identity to give

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = \frac{\partial}{\partial \frac{\partial x^r}{\partial \xi^s}} \sum_{t,u} g^{pt} g^{qu} g_{tu}, \tag{C.2}$$

then differentiate using the chain rule:

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = \sum_{t,u} \left(g^{pt} g^{qu} \frac{\partial g_{tu}}{\partial \frac{\partial x^r}{\partial \xi^s}} + g^{pt} \frac{\partial g^{qu}}{\partial \frac{\partial x^r}{\partial \xi^s}} g_{tu} + \frac{\partial g^{pt}}{\partial \frac{\partial x^r}{\partial \xi^s}} g^{qu} g_{tu} \right). \tag{C.3}$$

This can be re-written by again using the fact that g^{qu} and g_{qu} are inverses to identify delta functions:

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = \sum_{t,u} \left(g^{pt} g^{qu} \frac{\partial g_{tu}}{\partial \frac{\partial x^r}{\partial \xi^s}} + \delta_u^p \frac{\partial g^{qu}}{\partial \frac{\partial x^r}{\partial \xi^s}} + \delta_t^q \frac{\partial g^{pt}}{\partial \frac{\partial x^r}{\partial \xi^s}} \right). \tag{C.4}$$

Now noting that the second and third terms on the RHS are the same as the LHS gives

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = -\sum_{t,u} g^{pt} g^{qu} \frac{\partial g_{tu}}{\partial \frac{\partial x^r}{\partial \xi^s}}, \tag{C.5}$$

and using the definition of g_{pq} of Eq. (5.3) gives

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = -\sum_{t,u,v} g^{pt} g^{qu} \frac{\partial}{\partial \frac{\partial x^r}{\partial \xi^s}} \frac{\partial x^v}{\partial \xi^t} \frac{\partial x^v}{\partial \xi^u}. \tag{C.6}$$

Performing the derivative,

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = -\sum_{t,u,v} g^{pt} g^{qu} \left(\delta_r^v \delta_s^t \frac{\partial x^v}{\partial \xi^u} + \delta_r^v \delta_s^u \frac{\partial x^v}{\partial \xi^t} \right). \tag{C.7}$$

Finally, collecting terms together yields

$$\frac{\partial g^{pq}}{\partial \frac{\partial x^r}{\partial \xi^s}} = -\sum_{t} (g^{ps} g^{qt} + g^{pt} g^{qs}) \frac{\partial x^r}{\partial \xi^t}, \tag{C.8}$$

which is stated without proof in Ref. 96. The relationship between an infinitesimal change in the transformation $\delta \mathbf{x}$ and the resulting change in g^{pq} is therefore

$$\delta g^{pq} = -\sum_{r,s,t} (g^{ps} g^{qt} + g^{pt} g^{qs}) \frac{\partial x^r}{\partial \xi^t} \frac{\partial}{\partial \xi^s} \delta x^r. \tag{C.9}$$

C.2 Derivative of the elastic energy

The elastic energy functional is defined by

$$E_{\mathrm{E}}[\mathbf{x}(\boldsymbol{\xi})] = \mu \int g^{-1} \mathrm{d}^2\xi, \tag{C.10}$$

where $g = \det g_{pq} = 1/\det g^{pq}$. The infinitesimal change in the elastic energy brought about by a change in the metric determinant is given by

$$\delta E_{\mathrm{E}} = \mu \int \delta\left(g^{-1}\right)\, \mathrm{d}^2\xi, \tag{C.11}$$

which, expanded in full (using the symmetry of g^{pq} and $\det g^{pq} = g^{11}g^{22} - g^{12}g^{12}$), is

$$\delta E_{\mathrm{E}} = \mu \int (g^{11}\delta g^{22} + g^{22}\delta g^{11} - 2g^{12}\delta g^{12})\, \mathrm{d}^2\xi. \tag{C.12}$$

Using Eq. (C.9) as derived previously gives

$$\delta E_{\mathrm{E}} = -2\mu \int \mathrm{d}^2\xi \sum_{r,s,t} \left\{g^{11}g^{2s}g^{2t} + g^{22}g^{1s}g^{1t} - g^{12}(g^{1s}g^{2t} + g^{1t}g^{2s})\right\} \frac{\partial x^r}{\partial \xi^t}\frac{\partial}{\partial \xi^s}\delta x^r. \tag{C.13}$$

Expanding the sum over t explicitly yields

$$\begin{aligned}\delta E_{\mathrm{E}} = -2\mu \int \mathrm{d}^2\xi \sum_{r,s} \Bigg\{ & (g^{11}g^{2s}g^{12} + g^{22}g^{1s}g^{11} - g^{12}g^{1s}g^{12} - g^{12}g^{11}g^{2s})\frac{\partial x^r}{\partial \xi^1} \\ & + (g^{11}g^{2s}g^{22} + g^{22}g^{1s}g^{12} - g^{12}g^{1s}g^{22} - g^{12}g^{12}g^{2s})\frac{\partial x^r}{\partial \xi^2}\Bigg\}\frac{\partial}{\partial \xi^s}\delta x^r.\end{aligned} \tag{C.14}$$

Rearranging gives

$$\begin{aligned}\delta E_{\mathrm{E}} = -2\mu \int \mathrm{d}^2\xi \sum_{r,s} \Bigg\{ & g^{1s}(g^{11}g^{22} - g^{12}g^{12})\frac{\partial x^r}{\partial \xi^1} \\ & + g^{2s}(g^{11}g^{22} - g^{12}g^{12})\frac{\partial x^r}{\partial \xi^2}\Bigg\}\frac{\partial}{\partial \xi^s}\delta x^r,\end{aligned} \tag{C.15}$$

and then substituting for $g^{-1} = g^{11}g^{22} - g^{12}g^{12}$:

$$\delta E_{\mathrm{E}} = -2\mu \int \mathrm{d}^2\xi \frac{1}{g} \sum_{r,s,t} g^{ts}\frac{\partial x^r}{\partial \xi^t}\frac{\partial}{\partial \xi^s}\delta x^r. \tag{C.16}$$

C.3 Derivative of the total fictitious energy

We can write the change in the total fictitious energy caused by a change in the transformation of $\delta\mathbf{x}$, using Eq. (C.16) for the elastic energy, as:

$$\delta E = \int \mathrm{d}^2\xi \sum_r \left(-\frac{\partial}{\partial x^r} |\nabla_t n^2| - \frac{2\mu}{g} \sum_{s,t} g^{ts} \frac{\partial x^r}{\partial \xi^t} \frac{\partial}{\partial \xi^s} \right) \delta x^r. \tag{C.17}$$

Eq. (5.1) relates $\delta\mathbf{x}$ to the set of $\mathbf{x_G}$ coefficients. An infinitesimal change of one coefficient $x^p_{\mathbf{G}}$ gives

$$\delta x^p = \delta x^p_{\mathbf{G}} e^{i\mathbf{G}\cdot\boldsymbol{\xi}}, \tag{C.18}$$

and substituting into Eq. (C.17):

$$\delta E = \int \mathrm{d}^2\xi \sum_r \left(-\frac{\partial}{\partial x^r} |\nabla_t n^2| - \frac{2\mu}{g} \sum_{s,t} g^{ts} \frac{\partial x^r}{\partial \xi^t} \frac{\partial}{\partial \xi^s} \right) \delta x^r_{\mathbf{G}} e^{i\mathbf{G}\cdot\boldsymbol{\xi}}. \tag{C.19}$$

Rearranging gives the required final result:

$$\frac{\partial E}{\partial x^r_{\mathbf{G}}} = -\int \mathrm{d}^2\xi \left(\frac{\partial}{\partial x^r} |\nabla_t n^2| + \frac{2i\mu}{g} \sum_{s,t} g^{ts} \frac{\partial x^r}{\partial \xi^t} G_s \right) e^{i\mathbf{G}\cdot\boldsymbol{\xi}}. \tag{C.20}$$

It is possible to use Eq. (C.20) to show that the elastic energy term is at an extremum when the grid is undistorted, i.e. $\mathbf{x_G} = \mathbf{0}$ or equivalently $\boldsymbol{\xi} \equiv \mathbf{x}$. In this case $g^{ts} = \delta^{ts}$, $g = 1$ and $\frac{\partial x^r}{\partial \xi^t} = \delta^r_t$. The derivative of the elastic energy is then given by:

$$\begin{aligned} \frac{\partial E_{\mathrm{E}}}{\partial x^r_{\mathbf{G}}} &= -2i\mu G_r \int e^{i\mathbf{G}\cdot\boldsymbol{\xi}} \, \mathrm{d}^2\xi && \text{(C.21a)} \\ &= -2i\mu\Omega G_r \delta_{\mathbf{G},\mathbf{0}} && \text{(C.21b)} \\ &= 0, && \text{(C.21c)} \end{aligned}$$

where δ is the Kronecker delta and Ω is the area of the unit cell.

Appendix D

Symmetry in reciprocal space

When interpreting the trajectories of modes guided in a defect in a PCF cladding, it is useful to be able to classify them according to their symmetry. In the first section of this appendix, a method is given to determine modal symmetry from the plane-wave coefficients describing the transverse magnetic field, without needing to examine the real-space representation of the field.

Symmetry considerations are also useful (although not essential) when using the adaptive grid generation method described in practice in Secs. 6.1.1 and 6.2.1. The way in which symmetry is used to reduce the computational effort of generating adaptive grids is described in the second section of this appendix.

D.1 Determination of field symmetry

The modal symmetries of waveguides with different structural symmetries were first investigated in detail by McIsaac in 1975 [147, 148]. Symmetry and degeneracy in PCFs can be analysed using the same methods, and have been the subject of much consideration, partly because of their use in simplifying calculations [67,112,149–151]. In Ref. 112, Guobin *et al.* apply the method of McIsaac to the modes of fibres with C_{6v} symmetry and label the eight possible modal symmetries (of which four consist of two degenerate pairs and four are non-degenerate)

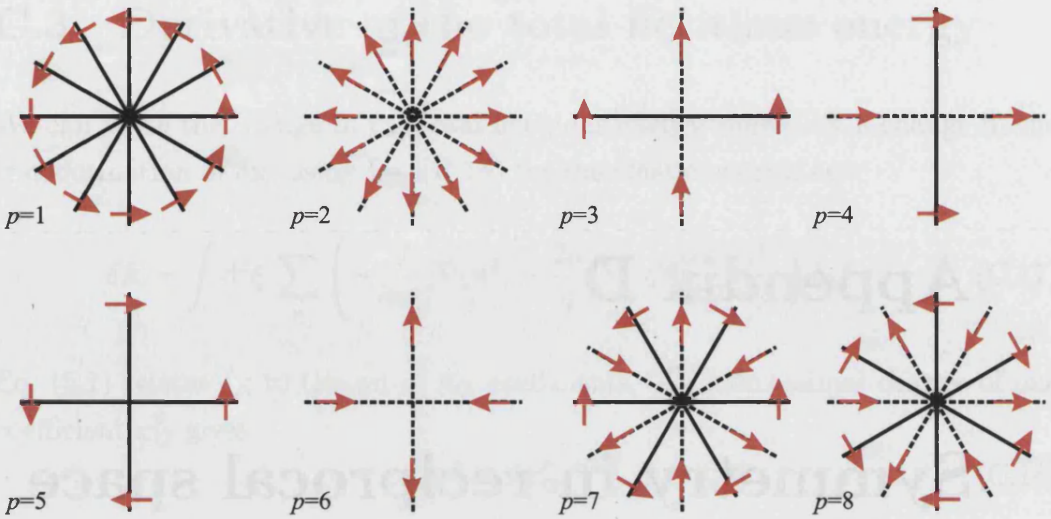


Figure D.1: The eight modal symmetries of a waveguide with C_{6v} symmetry. Dashed lines denote mirror lines, along which $\mathbf{h}_t$ is purely radial. Solid lines denote 'anti-mirror' lines, along which $\mathbf{h}_t$ is purely azimuthal. Red arrows show example directions of $\mathbf{h}_t$ along the symmetry axes. The degenerate pairs of modes are $p = 3, 4$ and $p = 5, 6$.

$p = 1 \ldots 8$. Fig. D.1 shows these eight symmetry types and the way in which the magnetic field is constrained by rotational and mirror symmetries[1].

In Table D.1 three features of the modal symmetry types are given: symmetry about $x = 0$, whether the field is even or odd under inversion, and whether the field vanishes at the origin. These properties can be verified easily by reference to the arrows representing the magnetic field vectors in Fig. D.1. It can be seen that the three properties shown uniquely identify each symmetry type, and therefore to determine symmetry types it is necessary only to establish these properties for a given field.

Fig. D.2 shows the first three 'stars' of reciprocal lattice vectors for the lattice defined by the primitive reciprocal lattice vectors $\mathbf{G}_1$ and $\mathbf{G}_2$. It is possible to determine the required properties described in Table D.1 by examination of components within the first 'star' only. Table D.2 shows the way in which the field components in the first 'star' (and at $\mathbf{G} = \mathbf{0}$) are related when the field possesses the properties of Table D.1.

[1]Note that the electric rather than magnetic field is used in Ref. 112, and so the mirror and anti-mirror lines are reversed relative to those here. We find it more convenient to use the magnetic field only.

p	Symmetry about $x = 0$	Inversion symmetry	$\mathbf{h}_t(\mathbf{x} = \mathbf{0})$
1	h^x even, h^y odd	even	$\mathbf{0}$
2	h^x odd, h^y even	even	$\mathbf{0}$
3,4	-	odd	$\neq \mathbf{0}$
5,6	-	even	$\mathbf{0}$
7	h^x odd, h^y even	odd	$\mathbf{0}$
8	h^x even, h^y odd	odd	$\mathbf{0}$

Table D.1: Selected properties of the symmetry types of Fig. D.1: symmetry about the y-axis, inversion symmetry, and whether the field vanishes at the origin. The field h^i is considered even about the y-axis if $h^i(-x, y) = h^i(x, y)$ and odd if $h^i(-x, y) = -h^i(x, y)$. Inversion symmetry is even if $\mathbf{h}_t(-\mathbf{x}) = -\mathbf{h}_t(\mathbf{x})$ and odd if $\mathbf{h}_t(-\mathbf{x}) = \mathbf{h}_t(\mathbf{x})$. Note that, although the two members of each degenerate pair separately satisfy symmetry conditions about $x = 0$, an arbitrary linear combination of the two does not.

To identify the symmetry type of a given mode using this procedure, it is necessary first to use the information in Table D.2 to determine whether h^x or h^y is even about $x = 0$, and then whether the field is even under inversion. If neither field is even about $x = 0$, the mode must be one of a degenerate pair: the two pairs can be distinguished by determining whether the component at $\mathbf{G} = \mathbf{0}$ is zero. Collecting this information and consulting Table D.1 then unambiguously identifies the required value of p for the mode.

D.2 Enforcing symmetry of adaptive grids

The mapping between $\boldsymbol{\xi}$-space and $\mathbf{x}$-space used in the GCC method must satisfy the same symmetries as the structure to which it is adapted. The structures considered in Chapter 6 belong to plane group P6mm (they have six-fold rotational symmetry and three mirror axes), and therefore the mapping must also have this symmetry[2]. By exploiting the relationships between $\mathbf{x}_\mathbf{G}$ components that are required by symmetry, the computational effort of generating adaptive grids can be reduced considerably.

In the same way that the reciprocal-space components of $\mathbf{h}_t$ must satisfy particular conditions as a result of symmetry (as described above in Sec. D.1), the components $\mathbf{x}_\mathbf{G}$ of the mapping are also constrained. The solid red arrows in Fig. D.3 show the directions and relative magnitudes of the components of $\mathbf{x}_\mathbf{G}$

[2]The directions of the vectors $(\mathbf{x} - \boldsymbol{\xi})$ under P6mm symmetry are the same as those of the magnetic field for $p = 2$ shown in Fig. D.1.

Property	Relationships between reciprocal-space components		
h^i even about $x = 0$	$h^i_{\mathbf{G}(1)} = h^i_{\mathbf{G}(3)}$	$h^i_{\mathbf{G}(4)} = h^i_{\mathbf{G}(6)}$	
h^i odd about $x = 0$	$h^i_{\mathbf{G}(1)} = -h^i_{\mathbf{G}(3)}$	$h^i_{\mathbf{G}(4)} = -h^i_{\mathbf{G}(6)}$	$h^i_{\mathbf{G}(2)} = h^i_{\mathbf{G}(5)} = 0$
Even inversion	$\mathbf{h}_{\mathbf{G}(1)} = -\mathbf{h}_{\mathbf{G}(4)}$	$\mathbf{h}_{\mathbf{G}(2)} = -\mathbf{h}_{\mathbf{G}(5)}$	$\mathbf{h}_{\mathbf{G}(3)} = -\mathbf{h}_{\mathbf{G}(6)}$
Odd inversion	$\mathbf{h}_{\mathbf{G}(1)} = \mathbf{h}_{\mathbf{G}(4)}$	$\mathbf{h}_{\mathbf{G}(2)} = \mathbf{h}_{\mathbf{G}(5)}$	$\mathbf{h}_{\mathbf{G}(3)} = \mathbf{h}_{\mathbf{G}(6)}$
$\mathbf{h}_t(\mathbf{x} = \mathbf{0}) = \mathbf{0}$	$\mathbf{h}_{\mathbf{G}(0)} = \mathbf{0}$		
$\mathbf{h}_t(\mathbf{x} = \mathbf{0}) \neq \mathbf{0}$	$\mathbf{h}_{\mathbf{G}(0)} \neq \mathbf{0}$		

Table D.2: Relationships between reciprocal-space field components in the first 'star' of Fig. D.2 for the set of properties given in Table D.1. The labels $\mathbf{G}(1)$ etc. refer to the numbering in Fig. D.2, with $\mathbf{h}_{\mathbf{G}(0)}$ the component at $\mathbf{G} = \mathbf{0}$.

in the second $\mathbf{G}$-vector 'star': note that, as a result of satisfying all mirror and rotational symmetries, the directions are purely radial and the magnitudes are all equal. The same conditions apply to the components within each 'star'. There is therefore only one degree of freedom in the set of components of each 'star' and, once one vector in the 'star' is chosen (e.g. those marked with numbers in Fig. D.3), all others can be calculated. This reduces the 38 $\mathbf{x}_\mathbf{G}$ components of the first three 'stars' to only 3 independent variables $x^1_{\mathbf{G}(1)}$, $x^1_{\mathbf{G}(2)}$ and $x^1_{\mathbf{G}(3)}$. The components at $\mathbf{G} = \mathbf{0}$ must be zero. As the same procedure can be extended easily to a larger number of 'stars', it is ideal for reducing the computational effort of creating adaptive grids for supercells (where ≈ 200 independent variables are needed).

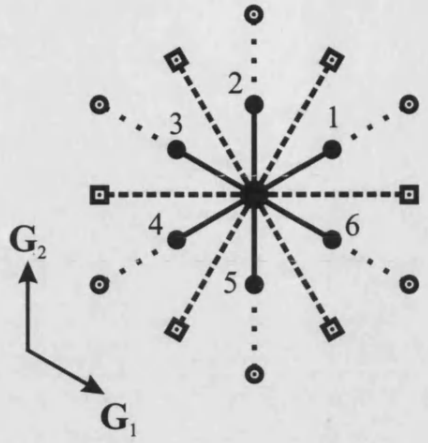


Figure D.2: First three **G**-vector 'stars' of the reciprocal lattice, shown respectively with solid lines ending in filled circles, dashes ending in squares, and shorter dashes ending in unfilled circles. The larger circle at the centre shows $\mathbf{G} = \mathbf{0}$. The primitive reciprocal lattice vectors are marked, and the six components of the first 'star' are labelled and referred to in Table D.2.

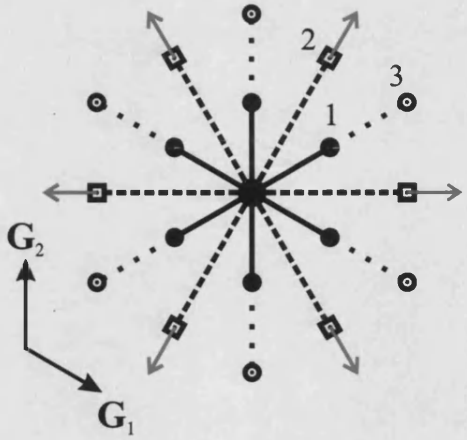


Figure D.3: First three **G**-vector 'stars' of the reciprocal lattice using the same line/dash notation as Fig. D.2. The solid red arrows represent the components $\mathbf{x}_\mathbf{G}$ of the second 'star'. The numbering is referred to in the text.

Appendix E

Density of states

The density of states (DOS) is a useful way in which to present the bandstructure of PCF cladding structures. Although calculation of the DOS is described in Refs. 45 and 74, its definition and method of calculation using the fixed-frequency plane-wave method is presented here for completeness.

We define the DOS, $\rho(k_0\Lambda, \beta\Lambda)$, such that $\rho(k_0\Lambda, \beta\Lambda)\Lambda \mathrm{d}\beta$ is proportional to the number of cladding states in the range $\beta\Lambda$ to $(\beta + \mathrm{d}\beta)\Lambda$ at a particular frequency $k_0\Lambda$. It is usual to normalise the DOS relative to its value in vacuum, $\rho_v = \sqrt{3}\beta\Lambda/2\pi$ [74]. The DOS then provides immediate information about the bandstructure of a PCF cladding at a particular value of k_0 and β. Large values of ρ indicate enhancement of the DOS relative to vacuum, and conversely a photonic bandgap is defined by $\rho(k_0\Lambda, \beta\Lambda) = 0$. This definition of the DOS makes it an ideal way to trace and identify the features within a bandstructure (see, e.g., Chapter 8 and Refs. 45, 52, 146, 152, 153).

When using the fixed-frequency plane-wave method to calculate the states of a PCF cladding, it is necessary to specify the frequency k_0 and also the Bloch wavevector $\mathbf{k}$ (see Chapter 3). The resulting β values are therefore a function of both k_0 and $\mathbf{k}$, and calculation at a finite set of values of $\mathbf{k}$ followed by integration over the Brillouin zone is required when calculating the DOS[1]. If the ith value of

[1]This approach is also common in plane-wave electronic structure calculations, to determine the electronic potential or total energy of a structure [72].

β calculated at Bloch wavevector $\mathbf{k}$ is labelled $\beta_{\mathbf{k}}^{(i)}$, the DOS is given by

$$\rho(\beta\Lambda, k_0\Lambda) = \frac{1}{\rho_v} \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_i \delta(\beta\Lambda - \beta_{\mathbf{k}}^{(i)}\Lambda), \tag{E.1}$$

where $w_{\mathbf{k}}$ are a set of weights normalised such that $\sum_{\mathbf{k}} w_{\mathbf{k}} = 1$, and δ is the Dirac delta function. The weights associated with each discrete $\mathbf{k}$-point are chosen such that summation over the irreducible wedge of the first Brillouin zone approximates to an integral over the entire first Brillouin zone. Although there exist methods to make efficient choices of $\mathbf{k}$-points (developed for electronic structure calculations [154, 155]), we find in general that a uniform grid (typically of $\approx$ 50-100 points in the irreducible wedge) produces well-converged results.

In practice the δ-function of Eq. (E.1) is replaced by a function of non-zero width and unit area to obtain a smooth DOS. The choice of the width of this function is 'cosmetic': a more smoothed DOS masks the discrete nature of the $\mathbf{k}$-point sampling but can cause blurring of fine features. A typical smoothing function is a Gaussian with full-width half-maximum 0.2 in units of $\beta\Lambda$.

References

[1] E. Yablonovitch, "Inhibited spontaneous emission in solid-state physics and electronics", Phys. Rev. Lett. **58**, 2059 (1987).

[2] S. John, "Strong localization of photons in certain disordered dielectric superlattices", Phys. Rev. Lett. **58**, 2486 (1987).

[3] J.D. Joannopoulos, R.D. Meade and J.N. Winn, *Photonic Crystals: Molding the Flow of Light*, Princeton University Press (1995).

[4] C.M. Soukoulis, "The history and a review of the modelling and fabrication of photonic crystals", Nanotechnology **13**, 420 (2002).

[5] J.D. Joannopoulos, P.R. Villeneuve and S. Fan, "Photonic crystals: putting a new twist on light", Nature **386**, 143 (1997).

[6] G. Parker and M. Charlton, "Photonic crystals", Physics World **13**, 29 (Aug 2000).

[7] L. Venema (Ed.), "Photonic technologies", Nature **424**, 809 (2003).

[8] P. St.J. Russell, "Photonic crystal fibers", Science **299**, 358 (2003).

[9] J.C. Knight, "Photonic crystal fibres", Nature **424**, 847 (2003).

[10] J.M. Senior, *Optical Fiber Communications: Principles and Practice*, Prentice Hall, 2nd edition (1992).

[11] A.W. Snyder and J.D. Love, *Optical Waveguide Theory*, Chapman and Hall (1983).

[12] J.C. Knight, T.A. Birks, P. St.J. Russell and D.M. Atkin, "All-silica single-mode optical fiber with photonic crystal cladding", Opt. Lett. **21**, 1547 (1996).

[13] T.A. Birks, J.C. Knight and P. St.J. Russell, "Endlessly single-mode photonic crystal fiber", Opt. Lett. **22**, 961 (1997).

[14] J.C. Knight, J. Broeng, T.A. Birks and P. St.J. Russell, "Photonic band gap guidance in optical fibers", Science **282**, 1476 (1998).

[15] R.F. Cregan, B.F. Mangan, J.C. Knight, T.A. Birks, P. St.J. Russell, P.J. Roberts and D.C. Allan, "Single-mode photonic band gap guidance of light in air", Science **285**, 1537 (1999).

[16] P. Yeh and A. Yariv, "Bragg reflection waveguides", Opt. Commun. **19**, 427 (1976).

[17] P. Yeh, A. Yariv and E. Marom, "Theory of Bragg fiber", J. Opt. Soc. Am. **68**, 1196 (1978).

[18] B. Temelkuran, S.D. Hart, G. Benoit, J.D. Joannopoulos and Y. Fink, "Wavelength-scalable hollow optical fibres with large photonic bandgaps for CO_2 laser transmission", Nature **420**, 650 (2002).

[19] S.D. Hart, G.R. Maskaly, B. Temelkuran, P.H. Prideaux, J.D. Joannopoulos and Y. Fink, "External reflection from omnidirectional dielectric mirror fibers", Science **296**, 510 (2002).

[20] J.C. Knight, T.A. Birks, R.F. Cregan, P. St.J. Russell and J.P. de Sandro, "Large mode area photonic crystal fibre", Electron. Lett. **34**, 1347 (1998).

[21] J. Limpert, T. Schreiber, S. Nolte, H. Zellmer, A. Tunnermann, R. Iliew, F. Lederer, J. Broeng, G. Vienne, A. Petersson and C. Jakobsen, "High-power air-clad large-mode-area photonic crystal fiber laser", Opt. Express **11**, 818 (2003).

[22] J.C. Knight, J. Arriaga, T.A. Birks, A. Ortigosa-Blanch, W.J. Wadsworth and P. St.J. Russell, "Anomalous dispersion in photonic crystal fiber", IEEE Phot. Tech. Lett. **12**, 807 (2000).

[23] A. Ferrando, E. Silvestre, J.J. Miret and P. Andrés, "Nearly zero ultraflattened dispersion in photonic crystal fibers", Opt. Lett. **25**, 790 (2000).

[24] W.H. Reeves, J.C. Knight, P. St.J. Russell and P.J. Roberts, "Demonstration of ultra-flattened dispersion in photonic crystal fibers", Opt. Express **10**, 609 (2003).

[25] L.P. Shen, W.P. Huang, G.X. Chen and S.S. Jian, "Design and optimization of photonic crystal fibers for broad-band dispersion compensation", IEEE Phot. Tech. Lett. **15**, 540 (2003).

[26] K. Saitoh, M. Koshiba, T. Hasegawa and E. Sasaoka, "Chromatic dispersion control in photonic crystal fibers: application to ultra-flattened dispersion", Opt. Express **11**, 843 (2003).

[27] J.K. Ranka, R.S. Windeler and A.J. Stentz, "Visible continuum generation in air-silica microstructure optical fibers with anomalous dispersion at 800 nm", Opt. Lett. **25**, 27 (2000).

[28] W.J. Wadsworth, A. Ortigosa-Blanch, J.C. Knight, T.A. Birks, T.P.M. Man and P. St.J. Russell, "Supercontinuum generation in photonic crystal fibers and optical fiber tapers: a novel light source", J. Opt. Soc. Am. B **19**, 2148 (2002).

[29] I. Hartl, X.D. Li, C. Chudoba, R.K. Ghanta, T.H. Ho, J.G. Fujimoto, J.K. Ranka and R.S. Windeler, "Ultrahigh-resolution optical coherence tomography using continuum generation in an air-silica microstructure optical fiber", Opt. Lett. **26**, 608 (2001).

[30] W.J. Wadsworth, N. Joly, J.C. Knight, T.A. Birks, F. Biancalana and P. St.J. Russell, "Supercontinuum and four-wave mixing with Q-switched pulses in endlessly single-mode photonic crystal fibres", Opt. Express **12**, 299 (2004).

[31] R. Holzwarth, T. Udem, T.W. Hänsch, J.C. Knight, W.J. Wadsworth and P. St.J. Russell, "Optical frequency synthesizer for precision spectroscopy", Phys. Rev. Lett. **85**, 2264 (2000).

[32] T. Udem, R. Holzwarth and T.W. Hänsch, "Optical frequency metrology", Nature **416**, 233 (2002).

[33] F. Benabid, J.C. Knight, G. Antonopoulos and P. St.J. Russell, "Stimulated Raman scattering in hydrogen-filled hollow-core photonic crystal fiber", Science **298**, 399 (2002).

[34] M.O. Scully and M.S. Zubairy, *Quantum Optics*, Cambridge University Press (1997).

[35] F. Benabid, P.S. Light, F. Couny and P. St.J. Russell, "Electromagnetically-induced transparency grid in acetylene-filled hollow-core PCF", Opt. Express **13**, 5694 (2005).

[36] S.E. Harris, J.E. Field and A. Imamoğlu, "Nonlinear optical processes using electromagnetically induced transparency", Phys. Rev. Lett. **64**, 1107 (1990).

[37] S. Ghosh, J.E. Sharping, D.G. Ouzounov and A.L. Gaeta, "Resonant optical interactions with molecules confined in photonic band-gap fibers", Phys. Rev. Lett. **94**, 093902 (2005).

[38] K. Dholakia, G. Spalding and M. MacDonald, "Optical tweezers: The next generation", Physics World **15**, 31 (Oct 2002).

[39] F. Benabid, J.C. Knight and P. St.J. Russell, "Particle levitation and guidance in hollow-core photonic crystal fiber", Opt. Express **10**, 1195 (2002).

[40] P.J. Roberts, F. Couny, H. Sabert, B.J. Mangan, D. Williams, L. Farr, M.W. Mason, A. Tomlinson, T.A. Birks, J.C. Knight and P. St.J. Russell, "Ultimate low loss of hollow-core photonic crystal fibres", Opt. Express **13**, 236 (2004).

[41] G. Humbert, J.C. Knight, G. Bouwmans, P. St.J. Russell, D.P. Williams, P.J. Roberts and B.J. Mangan, "Hollow core photonic crystal fibers for beam delivery", Opt. Express **12**, 1477 (2004).

[42] S.O. Konorov, V.P. Mitrokhin, A.B. Fedotov, D.A. Sidorov-Biryukov, V.I. Beloglazov, N.B. Skibina, E. Wintner, M. Scalora and A.M. Zheltikov, "Hollow-core photonic-crystal fibres for laser dentistry", Phys. Med. Biol. **49**, 1359 (2004).

[43] R. Haynes, J. Bland-Hawthorn, M.C.J. Large, K.F. Klein and G.W. Nelson, "New age fibers: the children of the photonic revolution", Proc. SPIE **5494**, 586 (2004).

[44] J.D. Shephard, F. Couny, P. St.J. Russell, J.D.C. Jones, J.C. Knight and D.P. Hand, "Improved hollow-core photonic crystal fiber design for delivery of nanosecond pulses in laser micromachining applications", Appl. Opt. **44**, 4582 (2005).

[45] J.M. Pottage, D.M. Bird, T.D. Hedley, T.A. Birks, J.C. Knight, P. St.J. Russell and P.J. Roberts, "Robust photonic band gaps for hollow core guidance in PCF made from high index glass", Opt. Express **11**, 2854 (2003).

[46] D.G. Ouzounov, F.R. Ahmad, D. Müller, N. Venkataraman, M.T. Gallagher, M.G. Thomas, J. Silcox, K.W. Koch and A.L. Gaeta, "Generation of megawatt optical solitons in hollow-core photonic band-gap fibers", Science **301**, 1702 (2003).

[47] F. Luan, J.C. Knight, P. St.J. Russell, S. Campbell, D. Xiao, D.T. Reid, B.J. Mangan, D.P. Williams and P.J. Roberts, "Femtosecond soliton pulse delivery at 800 nm wavelength in hollow-core photonic bandgap fibers", Opt. Express **12**, 835 (2004).

[48] R.T. Bise, R.S. Windeler, K.S. Kranz, C. Kerbage, B.J. Eggleton and D.J. Trevor, "Tunable photonic band gap fiber", in *OSA Trends in Optics and Photonics Series* **70**, p466, Optical Society of America (2002).

[49] F. Luan, A.K. George, T.D. Hedley, G.J. Pearce, D.M. Bird, J.C. Knight and P. St.J. Russell, "All-solid photonic bandgap fiber", Opt. Lett. **29**, 2369 (2004).

[50] G. Bouwmans, L. Bigot, Y. Quiquempois, F. Lopez, L. Provino and M. Douay, "Fabrication and characterization of an all-solid 2D photonic bandgap fiber with a low-loss region (< 20 dB/km) around 1550 nm", Opt. Express **13**, 8452 (2005).

[51] T.A. Birks, P.J. Roberts, P. St.J. Russell, D.M. Atkin and T.J. Shepherd, "Full 2-D photonic bandgaps in silica/air structures", Electron. Lett. **31**, 1941 (1995).

[52] F. Couny, H. Sabert, P.J. Roberts, D.P. Williams, A. Tomlinson, B.J. Mangan, L. Farr, J.C. Knight, T.A. Birks and P. St.J. Russell, "Visualizing the photonic band gap in hollow core photonic crystal fibers", Opt. Express **13**, 558 (2005).

[53] J.C. Maxwell, *A Treatise on Electricity and Magnetism*, Clarendon Press, Oxford (1873).

[54] A. Taflove and S.C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, 3rd edition (2005).

[55] M. Qiu, "Analysis of guided modes in photonic crystal fibers using the finite-difference time-domain method", Microwave and Optical Tech. Lett. **30**, 327 (2001).

[56] M.D. Feit and J. J A Fleck, "Computation of mode properties in optical fiber waveguides by a propagating beam method", Appl. Opt. **19**, 1154 (1980).

[57] K. Saitoh and M. Koshiba, "Full-vectorial imaginary-distance beam propagation method based ona finite element scheme: application to photonic crystal fibers", IEEE J. Quant. Electron. **38**, 927 (2002).

[58] Z. Zhu and T.G. Brown, "Full-vectorial finite-difference analysis of microstructured optical fibers", Opt. Express **10**, 853 (2002).

[59] V. Dangui, M.J.F. Digonnet and G.S. Kino, "A fast and accurate numerical tool to model the modal properties of photonic-bandgap fibers", Opt. Express **14**, 2979 (2006).

[60] J. Jin, *The Finite-Element Method in Electromagnetics*, Wiley, New York (1993).

[61] F. Brechet, J. Marcou, D. Pagnoux and P. Roy, "Complete analysis of the characteristics of propagation into photonic crystal fibers by the finite element method", Opt. Fiber Tech. **6**, 181 (2000).

[62] H.P. Uranus and H.J.W.M. Hoekstra, "Modelling of microstructured waveguides using a finite-element-based vectorial mode solver with transparent boundary conditions", Opt. Express **12**, 2795 (2004).

[63] M. Koshiba, "Full-vector analysis of photonic crystal fibers using the finite element method", IEICE Trans. Electron. **E85-C**, 881 (2002).

[64] N. Guan, S. Habu, K. Takenaga, K. Himeno and A. Wada, "Boundary element method for analysis of holey optical fibers", J. Lightwave Tech. **21**, 1787 (2003).

[65] X. Wang, J. Lou, C. Lu, C. Zhao and W.T. Ang, "Modeling of PCF with multiple reciprocity boundary element method", Opt. Express **12**, 961 (2004).

[66] T.P. White, R.C. McPhedran, C.M. de Sterke, L.C. Botten and M.J. Steel, "Confinement losses in microstructured optical fibers", Opt. Lett. **26**, 1660 (2001).

[67] T.P. White, B.T. Kuhlmey, R.C. McPhedran, D. Maystre, G. Renversez, C.M. de Sterke and L.C. Botten, "Multipole method for microstructured optical fibers: I. Formulation", J. Opt. Soc. Am. B **19**, 2322 (2002).

[68] S.G. Johnson and J.D. Joannopoulos, "Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis", Opt. Express **8**, 173 (2001).

[69] J.D. Jackson, *Classical Electrodynamics*, Wiley, New York, 3rd edition (1988).

[70] J. Lægsgaard, A. Bjarklev and S.E. Barkou Libori, "Chromatic dispersion in photonic crystal fibers: fast and accurate scheme for calculation", J. Opt. Soc. Am. B **20**, 443 (2003).

[71] S. Campbell, R.C. McPhedran, C.M. de Sterke and L.C. Botten, "Differential multipole method for microstructured optical fibers", J. Opt. Soc. Am. B **21**, 1919 (2004).

[72] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias and J.D. Joannopoulos, "Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients", Rev. Mod. Phys. **64**, 1045 (1992).

[73] C.T. Chan, Q.L. Yu and K.M. Ho, "Order-N spectral method for electromagnetic waves", Phys. Rev. B **51**, 16635 (1995).

[74] T.D. Hedley, *Modelling of Photonic Crystal Fibre*, Ph.D. thesis, University of Bath (2006).

[75] N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, Thomson (1976).

[76] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in FORTRAN 77*, Cambridge University Press (1992).

[77] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe and H. van der Vorst (Eds.), *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM (2006).

[78] G. Arfken, *Mathematical Methods for Physicists*, Academic Press, 3rd edition (1985).

[79] R.B. Lehoucq, D.C. Sorensen and C. Yang, *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, SIAM (1998).

[80] R. Barrett, M. Berry, T.F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM (1994).

[81] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS (1996).

[82] Y. Saad and M.H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems", SIAM J. Sci. Stat. Comp. **7**, 856 (1986).

[83] W.D. Joubert, G.F. Carey, N.A. Berner, A. Kalhan, H. Kohli, A. Lorber, R.T. McLay and Y. Shen, "PCG Reference Manual", (1996).

[84] W.J. Thompson, "Fourier series and the Gibbs phenomenon", Am. J. Phys. **60**, 425 (1992).

[85] H.S. Sözüer, J.W. Haus and R. Inguva, "Photonic bands: Convergence problems with the plane-wave method", Phys. Rev. B **45**, 13962 (1992).

[86] R.D. Meade, A.M. Rappe, K.D. Brommer, J.D. Joannopoulos and O.L. Alherhand, "Accurate theoretical analysis of photonic band-gap materials", Phys. Rev. B **48**, 8434 (1993).

[87] J.M. Ford, "A black box at the end of the rainbow: searching for the perfect preconditioner", Phil. Trans. R. Soc. Lond. A **361**, 2665 (2003).

[88] E.D. Palik (Ed.), *Handbook of Optical Constants of Solids II*, Academic Press (1991).

[89] E. Merzbacher, *Quantum Mechanics*, Wiley (1998).

[90] D.M. Bird, private communication.

[91] R. Holland, "Finite-difference solution of Maxwell's equations in generalized nonorthogonal coordinates", IEEE Trans. Nucl. Sci. **30**, 4589 (1983).

[92] A.J. Ward and J.B. Pendry, "Calculating photonic Green's functions using a nonorthogonal finite-difference time-domain method", Phys. Rev. B **58**, 7252 (1998).

[93] J.P. Plumey, G. Granet and J. Chandezon, "Differential covariant formalism for solving Maxwell's equations in curvilinear coordinates: oblique scattering from lossy periodic surfaces", IEEE Trans. Antennas Propag. **43**, 835 (1995).

[94] M. Skorobogatiy, S.A. Jacobs, S.G. Johnson and Y. Fink, "Geometric variations in high index-contrast waveguides, coupled mode theory in curvilinear coordinates", Opt. Express **10**, 1227 (2002).

[95] M. Skorobogatiy, "Modeling the impact of imperfections in high-index-contrast photonic waveguides", Phys. Rev. E **70**, 046609 (2004).

[96] F. Gygi, "Adaptive Riemannian metric for plane-wave electronic-structure calculations", Europhys. Lett. **19**, 617 (1992).

[97] F. Gygi, "Electronic-structure calculations in adaptive coordinates", Phys. Rev. B **48**, 11692 (1993).

[98] M.P. Hobson, G. Efstathiou and A.N. Lasenby, *General Relativity*, Cambridge University Press (2006).

[99] D.R. Hamann, "Application of adaptive curvilinear coordinates to the electronic structure of solids", Phys. Rev. B **51**, 7337 (1995).

[100] D.R. Hamann, "Generalized-gradient functionals in adaptive curvilinear coordinates", Phys. Rev. B **54**, 1568 (1996).

[101] J.M. Pérez-Jordá, "Variational plane-wave calculations in adaptive coordinates", Phys. Rev. B **58**, 1230 (1998).

[102] W. Huang, Y. Ren and R.D. Russell, "Moving mesh partial differential equations (MMPDEs) based on the equidistribution principle", SIAM J. Numer. Anal. **31**, 709 (1994).

[103] W. Cao, R. Carretero-González, W. Huang and R.D. Russell, "Variational mesh adaptation methods for axisymmetrical problems", SIAM J. Numer. Anal. **41**, 235 (2003).

[104] F. Gygi, "Ab initio molecular dynamics in adaptive coordinates", Phys. Rev. B **51**, 11190 (1995).

[105] F. Gygi and G. Galli, "Real-space adaptive-coordinate electronic-structure calculations", Phys. Rev. B **52**, R2229 (1995).

[106] N.A. Modine, G. Zumbach and E. Kaxiras, "Adaptive-coordinate real-space electronic-structure calculations for atoms, molecules, and solids", Phys. Rev. B **55**, 10289 (1997).

[107] G.E. Mase, *Continuum Mechanics*, McGraw Hill (1970).

[108] A.E. Green and W. Zerna, *Theoretical Elasticity*, Oxford University Press (1968).

[109] P.J. Roberts, private communication.

[110] N.A. Nicorovici and R.C. McPhedran, "Lattice sums for off-axis electromagnetic scattering by gratings", Phys. Rev. E **50**, 3143 (1994).

[111] N.A. Mortensen and M.D. Nielsen, "Modeling of realistic cladding structures for air-core photonic bandgap fibers", Opt. Lett. **29**, 349 (2004).

[112] R. Guobin, W. Zhi, L. Shuqin and J. Shuisheng, "Mode classification and degeneracy in photonic crystal fibers", Opt. Express **11**, 1310 (2003).

[113] J.D. Shephard, W.N. MacPherson, R.R.J. Maier, J.D.C. Jones, D.P. Hand, M. Mohebbi, A.K. George, P.J. Roberts and J.C. Knight, "Single-mode mid-IR guidance in a hollow-core photonic crystal fiber", Opt. Express **13**, 7139 (2005).

[114] J.S. Wang, E.M. Vogel and E. Snitzer, "Tellurite glass: a new candidate for fiber devices", Opt. Mater. **3**, 187 (1994).

[115] V.V. Ravi Kanth Kumar, A.K. George, J.C. Knight and P. St.J. Russell, "Tellurite photonic crystal fiber", Opt. Express **10**, 2641 (2003).

[116] T.M. Monro, Y.D. West, D.W. Hewak, N.G.R. Broderick and D.J. Richardson, "Chalcogenide holey fibres", Electron. Lett. **36**, 1998 (2000).

[117] L. Brilland, F. Smektala, G. Renversez, T. Chartier, J. Troles, T.N. Nguyen, N. Traynor and A. Monteville, "Fabrication of complex structures of holey fibres in chalcogenide glass", Opt. Express **14**, 1280 (2006).

[118] L.B. Shaw, J.S. Sanghera, I.D. Aggarwal and F.H. Hung, "As-S and As-Se based photonic band gap fiber for IR laser transmission", Opt. Express **11**, 3455 (2003).

[119] C.M. Smith, N. Venkataraman, M.T. Gallagher, D. Müller, J.A. West, N.F. Borrelli, D.C. Allan and K.W. Koch, "Low-loss hollow-core silica/air photonic bandgap fibre", Nature **424**, 657 (2003).

[120] J.A. West, C.M. Smith, N.F. Borrelli, D.C. Allen and K.W. Koch, "Surface Modes in air-core photonic-bandgap fibers", Opt. Express **12**, 1485 (2004).

[121] D.C. Allan, N.F. Borrelli, M.T. Gallagher, D. Müller, C.M. Smith, N. Venkataraman, J.A. West, P. Zhang and K.W. Koch, "Surface modes and loss in air-core photonic bandgap fibers", Proc. SPIE **5000**, 161 (2003).

[122] K. Saitoh, N.A. Mortensen and M. Koshiba, "Air-core photonic band-gap fibers: the impact of surface modes", Opt. Express **12** (2004).

[123] H.K. Kim, M.J.F. Digonnet, G.S. Kino, J. Shin and S. Fan, "Simulations of the effect of the core ring on surface and air-core modes in photonic bandgap fibers", Opt. Express **12**, 3436 (2004).

[124] H.K. Kim, J. Shin, S. Fan, M.J.F. Digonnet and G.S. Kino, "Designing air-core photonic-bandgap fibers free of surface modes", IEEE J. Quant. Electron. **40**, 551 (2004).

[125] M.J.F. Digonnet, H.K. Kim, J. Shin, S. Fan and G.S. Kino, "Simple geometric criterion to predict the existence of surface modes in air-core photonic-bandgap fibers", Opt. Express **12**, 1864 (2004).

[126] P.J. Roberts, D.P. Williams, B.J. Mangan, H. Sabert, F. Couny, W.J. Wadsworth, T.A. Birks, J.C. Knight and P. St.J. Russell, "Realizing low loss air core photonic crystal fibers by exploiting an antiresonant core surround", Opt. Express **13**, 8277 (2005).

[127] I. Zavorin, D.P. O'Leary and H. Elman, "Complete stagnation of GMRES", Lin. Alg. & Appl. **367**, 165 (2003).

[128] N.M. Litchinitser, A.K. Abeeluck, C. Headley and B.J. Eggleton, "Antiresonant reflecting photonic crystal optical waveguides", Opt. Lett. **27**, 1592 (2002).

[129] J. Jasapara, T.H. Her, R. Bise, R. Windeler and D.J. DiGiovanni, "Group-velocity dispersion measurements in a photonic bandgap fiber", J. Opt. Soc. Am. B **20**, 1611 (2003).

[130] A. Argyros, T.A. Birks, S.G. Leon-Saval, C.M.B. Cordeiro, F. Luan and P. St.J. Russell, "Photonic bandgap with an index step of one percent", Opt. Express **13**, 309 (2005).

[131] A. Argyros, T.A. Birks, S.G. Leon-Saval, C.M.B. Cordeiro and P. St.J. Russell, "Guidance properties of low-contrast photonic bandgap fibres", Opt. Express **13**, 2503 (2005).

[132] A. Wang, A.K. George and J.C. Knight, "Three-level neodymium fiber laser incorporating photonic bandgap fiber", Opt. Lett. **31**, 1388 (2006).

[133] J.C. Knight, F. Luan, G.J. Pearce, A. Wang, T.A. Birks and D.M. Bird, "Solid photonic bandgap fibres and applications", Jpn. J. Appl. Phys. **45**, 6059 (2006).

[134] H.F. Taylor, "Bending effects in optical fibers", J. Lightwave Tech. **2**, 617 (1984).

[135] A.A. Gambling, H. Matsumura and C.M. Ragdale, "Curvature and microbending losses in single-mode optical fibres", Opt. Quant. Elec. **11**, 43 (1979).

[136] J.C. Baggett, T.M. Monro, K. Furusawa, V. Finazzi and D.J. Richardson, "Understanding bending losses in holey optical fibers", Opt. Commun. **227**, 317 (2003).

[137] T.P. Hansen, J. Broeng, C. Jakobsen, G. Vienne, H.R. Simonsen, M.D. Nielsen, P.M.W. Skovgaard, J.R. Folkenberg and A. Bjarklev, "Air-guiding photonic bandgap fibers: spectral properties, macrobending loss, and practical handling", J. Lightwave Tech. **22**, 11 (2004).

[138] T.P. White, R.C. McPhedran, C.M. de Sterke, N.M. Litchinitser and B.J. Eggleton, "Resonance and scattering in microstructured optical fibers", Opt. Lett. **27**, 1977 (2002).

[139] D. Marcuse, "Influence of curvature on the losses of doubly clad fibers", Appl. Opt. **21**, 4208 (1982).

[140] H. Heiblum and J.H. Harris, "Analysis of curved optical waveguides by conformal transformation", IEEE J. Quant. Electron. **11**, 75 (1975).

[141] J.D. Love, "Application of a low-loss criterion to optical waveguides and devices", IEE Proc. J: Optoelectronics **136**, 225 (1989).

[142] N.M. Litchinitser, S.C. Dunn, B. Usner, B.J. Eggleton, R.C. McPhedran and C.M. de Sterke, "Resonances in microstructured optical waveguides", Opt. Express **11**, 1243 (2003).

[143] G. Renversez, P. Boyer and A. Sagrini, "Antiresonant reflecting optical waveguide microstructured fibers revisited: a new analysis based on leaky mode coupling", Opt. Express **14**, 5682 (2006).

[144] J. Lægsgaard, "Gap formation and guided modes in photonic bandgap fibres with high-index rods", J. Opt. A: Pure Appl. Opt. **6**, 798 (2004).

[145] F. Benabid and D.M. Bird, private communication.

[146] J.M. Stone, G.J. Pearce, F. Luan, T.A. Birks, J.C. Knight, A.K. George and D.M. Bird, "An improved photonic bandgap fiber based on an array of rings", Opt. Express **14**, 6291 (2006).

[147] P.R. McIsaac, "Symmetry-induced modal characteristics of uniform waveguides: I. Summary of results", IEEE Trans. Microwave Theory Tech. **23**, 421 (1975).

[148] P.R. McIsaac, "Symmetry-induced modal characteristics of uniform waveguides: II. Theory", IEEE Trans. Microwave Theory Tech. **23**, 429 (1975).

[149] M.J. Steel, T.P. White, C.M. de Sterke, R.C. McPhedran and L.C. Botten, "Symmetry and degeneracy in microstructured optical fibers", Opt. Lett. **26**, 488 (2001).

[150] B.T. Kuhlmey, T.P. White, G. Renversez, D. Maystre, L.C. Botten, C.M. de Sterke and R.C. McPhedran, "Multipole method for microstructured optical fibers: II. Implementation and results", J. Opt. Soc. Am. B **19**, 2331 (2002).

[151] H.P. Uranus, "A simple and intuitive procedure for evaluating mode degeneracy in photonic crystal fibers", Am. J. Phys. **74**, 211 (2006).

[152] T.A. Birks, D.M. Bird, T.D. Hedley, J.M. Pottage and P. St.J. Russell, "Scaling laws and vector effects in bandgap-guiding fibres", Opt. Express. **12**, 69 (2004).

[153] G. Antonopoulos, F. Benabid, T.A. Birks, D.M. Bird, J.C. Knight and P. St.J. Russell, "Experimental demonstration of the frequency shift of bandgaps in photonic crystal fibers due to refractive index scaling", Opt. Express **14**, 3000 (2006).

[154] S.L. Cunningham, "Special points in the two-dimensional Brillouin zone", Phys. Rev. B **10**, 4988 (1974).

[155] H.J. Monkhorst and J.D. Pack, "Special points for Brillouin-zone integrations", Phys. Rev. B **13**, 5188 (1976).