

Photonic Crystals: From Theory to Practice

by

Steven G. Johnson

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2001

© Massachusetts Institute of Technology, 2001. All Rights Reserved.

Author
Department of Physics
May 18, 2001

Certified by
John D. Joannopoulos
Francis Wright Davis Professor of Physics
Thesis Supervisor

Accepted by
Thomas J. Greytak
Associate Department Head for Education

Photonic Crystals: From Theory to Practice

by

Steven G. Johnson

Submitted to the Department of Physics on May 18, 2001,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis, we explore the design, computation, and analysis of photonic crystals, with a special emphasis on structures and devices that make a connection with practically realizable systems.

First, we analyze the properties of photonic-crystal slabs: 2d periodic dielectric structures that have a band gap for propagation in a plane and that use index-guiding to confine light in the third dimension. Such structures are more amenable to fabrication than photonic crystals with full 3d band-gaps, but retain or approximate many of the latter's desirable properties. We show how traditional band-structure analysis can be adapted to slab systems in the context of several representative structures, and describe the unique features that arise in this framework compared to ordinary photonic crystals. We study the possibility of lossless linear waveguides in such systems, and highlight their differences with both conventional waveguides and waveguides in true photonic crystals. Finally, we consider the creation of high- Q cavities in slabs, for which the lack of a complete gap entails unavoidable radiation losses. Two mechanisms for minimizing such losses are described and demonstrated: mode delocalization and the novel far-field multipole cancellation.

Next, we present a 3d periodic dielectric structure with a large, complete photonic band-gap. The structure is distinguished by a sequence of planar layers, identical except for a horizontal offset, and repeating every three layers to form an fcc lattice. The high symmetry of the layers means that complex devices could be formed by modifying only a single layer, and their similarity to common 2d photonic crystals allows the direct application of results and experience from those simpler systems.

Third, we present and demonstrate general criteria for crossing perpendicular waveguides without crosstalk, based on *a priori* principles of symmetry and resonance.

Finally, we describe a fully-vectorial, 3d algorithm to compute the definite-frequency eigenstates of Maxwell's equations in arbitrary periodic dielectric structures, including systems with anisotropy (birefringence) or magnetic materials, using preconditioned block-iterative eigensolvers in a planewave basis. Many different numerical techniques are compared and characterized. Our implementation is freely available on the Web.

Thesis Supervisor: John D. Joannopoulos

Title: Francis Wright Davis Professor of Physics

Biographical Note

Steven G. Johnson was born on September 8, 1973, amid the cornfields, or rather the suburbs, of Illinois—growing up in the promised land of Saint Charles after a fleeting golden age of Naperville. His ambitions for the future formed a continuous progression, starting from the goal of being Dopey of the Seven Dwarfs, to the closely-related career of astronaut, thence to mission specialist, and finally to scientist (after being reassured by his sister Teresa of how “scientific” he looked in his new glasses). Any photographs of him dressed as a ballerina are utter fabrications. Mrs. Fitz’ chemistry class at Saint Patrick’s grade school (motto: “If it isn’t green, it’s not from God”) was terribly exciting, showing him for the first time how to make quantitative predictions of natural events (here, the proportions of different chemicals that could combine to make new compounds, such as explosives). He sensibly shifted his sights to physics, however, after a high-school chemistry text explained that the laws of chemistry actually come from physicists, “who are very clever at figuring out those kinds of things.” His high-school, by the way, consisted of one year making friends for life at Saint Charles High School (motto: “All school buildings should fit comfortably inside the swimming pool”), followed by three years in search of a girlfriend at the Illinois Mathematics and Science Academy (IMSA, a.k.a. “High-Tech High”). He had wanted to go to MIT since his freshman year of high school, based in part on televised robot competitions, which turned out to be a good thing as he seems to have received a life sentence at this institution. Here, surrounded by hackers wearing capes and machines that go “ping,” Steven’s family, friends, and colleagues agree that his lifelong goal has been fulfilled: he is Dopey, after all.

Assuming that he’s not dead, Steven can best be reached at his “lifetime” email address of:

`stevenj@alum.mit.edu`

The following pages contain a less colloquial summary of Steven’s professional activities.

Education and Awards

Massachusetts Institute of Technology (MIT), 1995–2001: Doctoral program in physics. Recipient of a National Defense Science and Engineering Graduate Fellowship (NDSEG) and an MIT Karl Taylor Compton Fellowship; also offered National Science Foundation (NSF) and Hertz Foundation fellowships.

MIT, 1991–1995: B.S. Physics, B.S. Electrical Engineering and Computer Science, B.S. Mathematics. 5.0/5.0 GPA. Thesis on semi-classical analyses of quantum chaos. Recipient of MIT Orloff Award for scholarship in physics. Member of honor societies: Phi Beta Kappa (liberal arts), Sigma Pi Sigma (physics), Tau Beta Pi (engineering), and Eta Kappa Nu (electrical engineering).

Illinois Mathematics and Science Academy, 1988–1991. Eagle Scout. Voted “biggest I-think-I-know-it-all.”

Co-recipient of the 1999 *J. H. Wilkinson Prize for Numerical Software*, for work on the FFTW fast Fourier transform library (<http://www.fftw.org/>) with Dr. Matteo Frigo of MIT. This is awarded every four years at the Intl. Conf. on Industrial and Applied Math (ICIAM) by Argonne National Laboratory, the National Physical Laboratory (UK), and the Numerical Algorithms Group to the software that “best addresses all phases of the preparation of high-quality numerical software.” For the same work, was co-recipient of a 1999 *Laurels Award* in “Electronics” from *Aviation Week & Space Technology* for “significant contributions to the global field of aerospace in 1999.”

Research and Teaching Experience

OmniGuide Communications, Inc. (Cambridge, MA), Sep. 2000–Jan. 2001: lead theory and simulation research group in the investigation and theoretical characterization of a new class of hollow optical fibers. (Part-time work there continues to the present day.)

Femtosecond Technology Research Association (Tsukuba, Japan), Summer 2000: introduced researchers at FESTA to the theoretical study of photonic crystals, and performed calculations to characterize the structures fabricated at their laboratory.

AT&T Bell Laboratories, summers of 1993 and 1994 (MIT VI-A internship program): developed real-time visualization software for high-speed optics, built laser-alignment system for free-space optics, and developed program for 3d capacitance extraction.

MIT Mathematics Dept. and School of Mgmt., spring 1993 (MIT UROP program): developed numerical studies of options contracts and Black-Scholes equations.

Fermi Natl. Accelerator Laboratory, summers/winters 1989–1991, 1995: developed simulation software for liquid Helium cryogenics systems.

MIT Experimental Study Group, 1994–present: taught freshman physics courses to small groups of 1–3 students, for roughly 6–8 semesters up to the present date, in this optional alternative academic program for MIT freshmen.

Publications

- S. G. Johnson and J. D. Joannopoulos, “Block-iterative frequency-domain methods for Maxwell’s equations in a planewave basis,” *Optics Express* **8** (3), 173–190 (Focus Issue on Photonic Bandgap Calculations) (2001).
- E. Chow, S. Y. Lin, J. R. Wendt, S. G. Johnson, and J. D. Joannopoulos, “Quantitative analysis of bending efficiency in photonic-crystal waveguide bends at $\lambda = 1.55\mu\text{m}$ wavelengths,” *Optics Lett.* **26**, 286–288 (2001).
- S. Fan, S. G. Johnson, J. D. Joannopoulos, C. Manolatu, and H. A. Haus, “Waveguide branches in photonic crystals,” *J. Opt. Soc. Am. B* **18** (2), 162–165 (2001).
- S. G. Johnson and J. D. Joannopoulos, “Three-dimensionally periodic dielectric layered structure with omnidirectional photonic band gap,” *Appl. Phys. Lett.* **77**, 3490–3492 (2000).
- E. Chow, S. Y. Lin, S. G. Johnson, P. B. Villeneuve, J. D. Joannopoulos, J. R. Wendt, G. A. Vawter, W. Zubrzycki, H. Hou, and A. Alleman, “Three-dimensional control of light in a two-dimensional photonic crystal slab,” *Nature* **407**, 983–986 (2000).
- S. G. Johnson, P. R. Villeneuve, S. Fan, and J. D. Joannopoulos, “Linear waveguides in photonic-crystal slabs,” *Phys. Rev. B* **62**, 8212–8222 (2000).
- S. Y. Lin, E. Chow, S. G. Johnson, and J. D. Joannopoulos, “Demonstration of highly efficient waveguiding in a photonic crystal slab at the 1.5- μm wavelength,” *Opt. Lett.* **25**, 1297–1299 (2000).
- S. G. Johnson, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and L. A. Kolodziejski, “Guided modes in photonic crystal slabs,” *Phys. Rev. B* **60**, 5751–5758 (1999).
- C. Manolatu, S. G. Johnson, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, “High-Density integrated optics,” *J. Lightwave Tech.* **17**, no. 9 (Sep. 1999).
- S. G. Johnson, C. Manolatu, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, “Elimination of cross talk in waveguide intersections,” *Optics Letters* **23**, 1855–1857 (1998).
- P. R. Villeneuve, S. Fan, S. G. Johnson, and J. D. Joannopoulos, “Three-dimensional photon confinement in photonic crystals of low-dimensional periodicity,” *IEE Proc. Optoelec.* **145**, 384 (1998).
- M. Frigo and S. G. Johnson, “FFTW: An adaptive software architecture for the FFT,” *Proc. ICASSP 1998*, vol. 3, 1381 (1998).
- W. M. Soyars and S. G. Johnson, “Simulating the Tevatron liquid Helium satellite refrigerators,” *Advances in Cryogenic Engineering* **39**, 1231–1235 (1994).

Acknowledgments

I don't know half of you half as well as I should like; and I like less than half of you half as well as you deserve.

J. R. R. Tolkien, *The Fellowship of the Ring*

It would be impossible to express my thanks on this page to all those who have supported me, without whose help I could never have come so far. I will attempt, at least, to satisfy the barest demands of decency by saying a few words here.

No gratitude is sufficient to repay the unceasing love of my parents and family, who have stood behind me from my first steps through all the moments of skinned knees and shaken confidence. They read me my first book, let a young boy wander through a particle accelerator, and never failed to call when my studies overwhelmed me. They gave me the ground I could stand on whenever the path ahead seemed dim. No son (or brother) could ask for better.

I will always be a student, I hope, and I pray that my future studies will be blessed by teachers as gifted and inspiring as the ones I have had. From Mrs. Fitz and Mr. Cordier in grade school who made their classrooms the best place in the world, to my calculus teacher Mr. Choi who urged us to see the beauty that others overlook, to Dr. Barr who brought me poetry and philosophy, to Peter Dourmashkin who showed me that teaching is the highest form of learning, they and many others have opened new worlds before my eyes.

As a teacher, as an advisor, as a colleague, as a friend, John Joannopoulos sets a standard that I will always strive to attain. I have been truly lucky to have learned from one so generous and so wise in the ways of life as well as those of physics. Thanks for putting up with such an obstinate student, John.

Finally, I am grateful to MIT for supporting me with a Karl Taylor Compton fellowship, and to the United States government for a National Defense Science and Engineering Graduate Fellowship.

Table of Contents

1	Photonic Crystals and Other Obscure Topics	21
1.1	In the Beginning, There Was Maxwell	21
1.2	Periodic Surprises	22
1.3	Photonic Crystals	24
1.4	Things to Come.....	29
2	Band Gaps & Maxwell’s Equations: A Whirlwind Tour.....	35
2.1	Maxwell’s Equations	35
2.2	The Bloch-Floquet Theorem and Electromagnetism.....	37
2.3	The Origin of the Photonic Band Gap	39
3	Elimination of Crosstalk in Waveguide Intersections	43
3.1	Introduction.....	43
3.2	Eliminating Crosstalk by Symmetry.....	44
3.3	Intersecting Photonic-Crystal Waveguides.....	46
3.4	Intersecting Conventional Waveguides	49
3.5	Summary	50
4	Guided Modes in Photonic-Crystal Slabs	53
4.1	Introduction.....	53
4.2	Computational Method	56
4.3	Photonic-Crystal Slab Band Structures.....	57
4.4	Effects of Slab Thickness.....	63
4.5	Slabs with Solid Backgrounds (Sandwiches)	67
4.6	Slabs with Periodic Backgrounds	69
4.7	Slabs with Symmetry-Breaking Backgrounds.....	72
4.8	Summary	74
5	Linear Waveguides in Photonic-Crystal Slabs	77
5.1	Introduction.....	77
5.2	Computational Method	83
5.3	Reduced-Index Waveguides	84
5.4	Increased-Index Waveguides.....	90
5.5	Strip Waveguides in Photonic-Crystal Slabs	93
5.6	Waveguides in Other Directions.....	96
5.7	Estimating the Field Energy in the Dielectric.....	98
5.8	Summary	102
6	High-Q Cavities in Photonic-Crystal Slabs	103
6.1	Introduction.....	103
6.2	Computational Method	105
6.3	Mode Delocalization.....	106
6.4	Multipole Cancellation.....	110
6.5	Summary and Conclusions	115
7	Three-dimensionally periodic dielectric layered structure with omnidirectional photonic bandgap.....	117
7.1	Introduction.....	117
7.2	Other Photonic-Crystal Structures	119

7.3	Characterizing the New Structure	122
7.4	Possible Fabrication Methods	123
7.5	Summary	124
8	Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis	125
8.1	Introduction	125
8.2	The Maxwell Eigenproblem	128
8.3	Iterative Eigensolvers	138
8.4	Summary	152
9	Concluding Remarks	153
9.1	Looking Back	153
9.2	Looking Forward	154
	Bibliography	157

List of Figures

Figure 1.1: Schematic of a one-dimensional photonic crystal: alternating dielectric layers (represented by blue and yellow), with a band gap in one direction. Incident light is reflected for colors of light in the band gap, but it cannot be used to trap light in 3d—the light inevitably escapes “sideways.”27

Figure 1.2: Schematic of the photonic-crystal slab chapters, covering first the bulk systems, then waveguides formed by linear defects, then cavities formed by point defects....31

Figure 1.3: A portion of a novel 3d photonic crystal that we propose in Chapter 7. This structure has a large, complete band gap, can be constructed layer-by-layer using traditional lithography, and its behaviour is understandable in terms of much-simpler two-dimensional photonic-crystal structures.....32

Figure 3.1: Abstract diagram of symmetry requirements for waveguide crossing, showing waveguide mode profiles and resonant-cavity mode contours. By symmetry, the solid-line (red) modes cannot couple with the dashed-line (blue) modes and vice-versa...44

Figure 3.2: Waveguide intersections in a two-dimensional photonic crystal consisting of a square lattice of dielectric rods in air.46

Figure 3.3: TM transmission spectra for the four intersections shown in Fig. 3.2: (a) throughput from the input port to the output port and (b) crosstalk from the input port to one of the transverse ports.48

Figure 3.4: Intersection of two-dimensional, high index-contrast waveguides in air, both (a) without holes and (b) with air holes added to create a resonant cavity using a photonic band gap.49

Figure 3.5: TE transmission spectra for the two intersections shown in Fig. 3.4.....50

Figure 4.1: Two-dimensional photonic crystals. (a) Square lattice of dielectric rods in air, with lattice constant a and radius $0.2a$. (b) Triangular lattice of air holes in dielectric, with lattice constant a and radius $0.45a$. In both case, the dielectric constant of the high-index material is 12.54

Figure 4.2: Band diagrams for the photonic crystals from (a) Fig. 4.1(a) and (b) Fig. 4.1(b). The shaded region indicates the frequencies of states introduced when vertical propagation (*i.e.*, perpendicular to the plane of periodicity) is permitted.....55

Figure 4.3: Photonic-crystal slabs analogous to the two-dimensional structures from Fig. 4.1. (a) Square lattice of rods in air with height $2.0a$. (b) Triangular lattice of holes in a dielectric slab with thickness $0.6a$. (Other parameters are as in Fig. 4.1.) Both slabs are suspended in air.57

Figure 4.4: Projected band diagrams corresponding to the two slabs in Fig. 4.3. Whether states are even or odd with respect to the horizontal mirror plane of the slab is indicated by open or filled circles, respectively.59

Figure 4.5: Lowest-frequency states of the rod and hole structures from Fig. at M and K ,

respectively, depicted in the unit cell. The shading indicates the z components of the electric and magnetic fields, respectively, which happen to be positive throughout the unit cell for these states. Outlines of the rods/holes are shown in white. (a) Vertical cross-section. (b) Horizontal cross-section. In both cases the cross-sections bisect the holes or rods.	61
Figure 4.6: (a) Horizontal cross-section of E_z for the first odd (TM-like) excited state at M in the rod structure from Fig. 4.4(a). (b) Vertical cross-section of the H_z for the first TE-like (but odd) excited state at K in the hole structure shown in Fig. 4.4(b).	62
Figure 4.7: Gap size (as a percentage of mid-gap frequency) versus slab thickness for the two slabs from Fig. 4.3.	63
Figure 4.8: Side-view of the slabs from Fig. 4.3 with a uniform, low-index background above and below the slabs. (a) is the rod slab and (b) is the hole slab. The background has a dielectric constant of 2.0 and all other parameters are as in Fig. 4.3.	67
Figure 4.9: Projected band diagrams for the structures shown in Fig. 4.8.	68
Figure 4.10: Side-view of the slabs shown in Fig. 4.3 with a periodic background formed by “extruding” the structures with a low-index material (dielectric constant 2.0). All other parameters are as in Fig. 4.3.	69
Figure 4.11: Projected band diagrams for the structures shown in Fig. 4.10.	70
Figure 4.12: Side-view of the slabs shown in Fig. 4.10 (extruded backgrounds) with the upper background replaced by air. All other parameters are as in Fig. 4.10. This background breaks the mirror symmetry of the two photonic-crystal slabs.	72
Figure 4.13: Projected band diagrams for the structures shown in Fig. 4.12. The bands can no longer be segregated into even and odd modes, and there is no longer a band gap in the guided modes.	73
Figure 4.14: A solid substrate may lie below a slab without affecting the band structure by being sufficiently far from the slab. (a), (b), (c), and (d) depict such solid substrates below the symmetric and antisymmetric rod and hole structures from Figs. 10(a), 12(a), 10(b), and 12(b), respectively.	74
Figure 5.1: Photonic-crystal slabs. (a) Square lattice of dielectric rods in air with ϵ of 12.0, lattice constant a , radius $0.2a$, and height $2.0a$, with low-index (ϵ of 2.0) rods extending infinitely above and below. (b) Triangular lattice of air holes extending through both a high-index (ϵ of 12.0) finite-height dielectric slab and low-index (ϵ of 2.0) semi-infinite dielectric regions above and below. The holes have lattice constant a and radius $0.3a$, while the high-index slab is of thickness $0.5a$	79
Figure 5.2: Band structure for the rod slab from Fig. 5.1(a). Hollow and filled circles represent modes which are even and odd with respect to the horizontal ($z = 0$) mirror plane bisecting the slab. There is a band gap in the odd (TM-like) modes in the frequency range $0.3362\text{--}0.4292 c/a$	80
Figure 5.3: Band structure for the hole slab from Fig. 5.1(b). Hollow and filled circles denote even and odd modes as in Fig. 5.1. There is a band gap in the even (TE-like) modes in the frequency range $0.2708\text{--}0.3471 c/a$	81

- Figure 5.4: Linear defects, which give rise to waveguide modes, in the rod and hole photonic-crystal slabs from Fig. 5.1. The low-index material is not shown, but “extrudes” above and below the structures as in Fig. 5.1. (a) Reduced-index waveguides, created by decreasing/increasing the radii of a line of nearest-neighbor rods/holes, respectively. (b) Increased-index waveguides, created by increasing/decreasing the radii of a line of nearest-neighbor rods/holes, respectively. (c) Dielectric-strip waveguide surrounded by the rod slab, from which a line of nearest-neighbor rods has been removed and replaced with the waveguide.82
- Figure 5.5: Projected band structure for the reduced-index rod-slab waveguide from Fig. 5.4(a), showing the (odd-symmetry) guided bands for various defect rod radii (compared to the bulk radius of $0.20a$).85
- Figure 5.6: Field cross-sections for the rod-radius $0.14a$ guided mode at X ($0.3707 c/a$), showing the z component of the electric field. The contours of the dielectric function are shown in black. (a) Horizontal (xy) cross-section in the plane bisecting the slab ($z = 0$). (b) Vertical (yz) cross-section in the plane perpendicular to the waveguide and bisecting a row of rods.....86
- Figure 5.7: Projected band structure for the reduced-index hole-slab waveguide from Fig. 5.4(a), showing the (even-symmetry) guided bands for various defect hole radii (compared to the bulk radius of $0.30a$).87
- Figure 5.8: Diagram of how the reciprocal lattice is projected onto $\Gamma-K$ (vertical dashed line) for a nearest-neighbor line defect in the hole slab. The black dots are the reciprocal lattice points, the outlined hexagon is the first Brillouin zone, and the grey triangle within it is the reduced Brillouin zone. K' is the boundary of the first Brillouin zone in the projected lattice.88
- Figure 5.9: Field cross-sections for the hole-radius $0.45a$ guided mode at K' ($0.3026 c/a$), showing the z component of the magnetic field. The contours of the dielectric function are shown in black. (a) Horizontal (xy) cross-section in the plane bisecting the slab ($z = 0$). (b) Vertical (yz) cross-section in the plane perpendicular to the waveguide and bisecting a defect hole.89
- Figure 5.10: Projected band structure for the increased-index rod-slab waveguide from Fig. 5.4(b), showing the (odd-symmetry) guided bands for defect rod radius $0.25a$ (compared to the bulk radius of $0.20a$).91
- Figure 5.11: Horizontal ($z = 0$) E_z field cross-sections for the rod-radius $0.25a$ guided modes at X . (a) p_x guided mode ($0.3887 c/a$). (b) p_y guided mode ($0.4205 c/a$).91
- Figure 5.12: Projected band structure for the increased-index rod-slab waveguide from Fig. 5.4(b), showing the (odd-symmetry) guided bands for defect rod radius $0.275a$.92
- Figure 5.13: Projected band structure for the increased-index hole-slab waveguide from Fig. 5.4(b), showing the (even-symmetry) guided bands for defect hole radii $0.15a$ and $0.25a$ (compared to the bulk radius of $0.30a$). For each radius, there are two or three guided bands: p_x (solid line), p_y (dashed line), and d_{xy} (dotted line, triangles). Guided bands for a defect hole radius of 0 , in which a row of holes is entirely removed, are

shown in bold.	93
Figure 5.14: Horizontal ($z = 0$) H_z field cross-sections for the hole-radius $0.15a$ guided modes at K' . (a) p_x guided mode ($0.2874 c/a$). (b) p_y guided mode ($0.3148 c/a$). (c) d_{xy} guided mode ($0.3377 c/a$).	94
Figure 5.15: Projected band structure for the strip waveguide in a rod slab, from Fig. 5.4(c), showing the (odd-symmetry) guided bands for waveguide width $0.25a$. There are two pairs of guided bands: gap-guided (solid lines) and index-guided (dashed lines). The gap-guided bands are split at X into a 0.07% gap, while the index-guided bands are split at X into a 0.8% gap.	95
Figure 5.16: Horizontal ($z = 0$) E_z field cross-sections for the $0.25a$ strip-waveguide gap-guided modes at X . (a) Lower gap-guided mode ($0.3626 c/a$). (b) Upper gap-guided mode ($0.3629 c/a$).	97
Figure 5.17: Vertical ($y = 0$) E_z field cross-sections of a unit cell, parallel to and bisecting the waveguide, for the strip-waveguide gap-guided and index-guided modes at X . (a) Upper gap-guided mode ($0.3629 c/a$). (b) Lower index-guided mode ($0.2632 c/a$).	97
Figure 6.1: Photonic-crystal slabs in 2d and 3d, with defects (resonant cavities) formed by increasing the radius of a central rod. (a) 2d slab: a sequence of dielectric rods with lattice constant a , radius $0.2a$, and ϵ of 11.56. (b) 3d slab: a square lattice of dielectric rods with lattice constant a , radius $0.2a$, height $2a$, and ϵ of 12.	104
Figure 6.2: Total Q (solid circles) and radiation Q_{\perp} (hollow circles) plotted vs. $\Delta\omega$ above the band edge for monopole defects in the 11×11 slab of rods from Fig. 6.1(b). Black points are for a single-rod defect (left inset, black), and red points are for a 5-rod super-defect (upper-right inset, red).	108
Figure 6.3: Monopole defect mode for $\epsilon=9$ super-defect ($Q = 11000$, $Q_{\perp} = 13000$): electric-field z -component in the plane bisecting the rods.	109
Figure 6.4: Q vs. ω for a dipole state in the 2d rod-sequence structure (inset) of Fig. 6.1(a). Both the total Q (black line, filled circles) and also the $Q_{\pm m}$ of the first few nonzero multipole moments (colored lines, hollow circles) are shown.	112
Figure 6.5: E_z radiation pattern for 2d dipole states of Fig. 6.4, using a color table that exaggerates small field magnitudes, with dielectric boundaries shown in black. (a) Point just before the peak ($Q=1773$, $\omega=0.328$). (b) Point at the peak ($Q=28700$, $\omega=0.309$), showing nodal lines from cancellation of the lowest multipole moment. (c) Point just beyond the peak ($Q=6624$, $\omega=0.300$).	113
Figure 6.6: (a) Q vs. ω for a quadrupole state in the 3d rod-slab structure of Fig. 6.1(b). E_z in the mid-plane at peak Q is shown in the inset. The central rod is $r=0.45a$ with $\epsilon=13$; ϵ of the four $r=0.25a$ neighbors was varied to control ω , and labels the points. The solid line is a Lorentzian curve fitted to the peak (with $R^2 = 0.9994$). (b) Electric-field energy density for this mode, plotted in a plane $2a$ above the rods. (i) Point just before the peak ($Q=426$, $\omega=0.346$). (ii) Point at the peak ($Q=1925$, $\omega=0.349$), showing nodal planes from cancellation of the lowest multipole moment. (iii) Point just beyond the peak ($Q=408$, $\omega=0.352$).	114
Figure 7.1: Computer rendering of a novel 3d photonic crystal, showing several horizontal periods	

and one vertical period, consisting of an fcc lattice of air holes (radius $0.293a$, height $0.93a$) in dielectric (21% fill). This structure has a 21% gap for a dielectric constant of 12.....118

Figure 7.2: Schematic of the crystal in Fig. 7.1. (a) The vertical structure, showing the different material layers as they might be deposited during fabrication. (The layers are given different colors for clarity, but would normally be the same material and thickness.) (b) Plan view of a horizontal cross-section intersecting the “A” cylinders, with the offset locations of cylinders in other layers also labelled.119

Figure 7.3: (a) Band diagram for the structure in Fig. 7.1, showing frequency in scale-invariant units vs. wavevector along important symmetry directions in the irreducible Brillouin zone. The inset shows the first Brillouin zone and its symmetry points. (b) Vertical transmission spectrum for slightly over one period (3 layers plus a capping hole slab, as shown in the inset) of this structure, showing the $\Gamma-L'$ gap.....120

Figure 8.1: Eigenvalue convergence as a function of grid resolution (grid points per lattice constant a) for three different methods of determining an effective dielectric tensor at each point: no averaging, simply taking the dielectric constant at each grid point; averaging, the smoothed effective dielectric tensor of Eq. (8.12); and backwards averaging, the same smoothed dielectric but with the averaging methods of the two polarizations reversed.).....135

Figure 8.2: Eigensolver convergence for two variants of conjugate gradient, Fletcher-Reeves and Polak-Ribiere, along with preconditioned steepest-descent for comparison. 143

Figure 8.3: The effect of two preconditioning schemes from section 2.4, diagonal and transverse-projection (non-diagonal), on the conjugate-gradient method.145

Figure 8.4: Comparison of the Davidson method with the block conjugate-gradient algorithm of Sec. 8.3.1. We reset the Davidson subspace to the best current eigenvectors every 2, 3, 4, or 5 iterations, with a corresponding increase in memory usage and computational costs.....147

Figure 8.5: Conjugate-gradient convergence of the lowest TM eigenvalue for the “interior” eigensolver of Eq. (8.27), solving for the monopole defect state formed by one vacancy in a 2D square lattice of dielectric rods in air, using three different supercell sizes (3×3 , 5×5 , and 7×7).148

Figure 8.6: Scaling of the number of conjugate-gradient iterations required for convergence (to a fractional tolerance of 10^{-7}) as a function of the spatial resolution (in grid points per lattice constant, with a corresponding planewave cutoff), or the number p of bands computed.....151

List of Tables

Table 7.1: Optimal parameters and gaps for various dielectric contrasts 123

Chapter 1

Photonic Crystals and Other Obscure Topics

1.1 In the Beginning, There Was Maxwell

The thing that hath been, it is that which shall be; and that which is done is that which shall be done: and there is no new thing under the sun.

Ecclesiastes 1:9.

In 1864, James Clerk Maxwell presented his treatise, “A dynamic theory of the electromagnetic field,” and revealed the equations underlying all phenomena in electricity and magnetism [1]. This remarkable work laid bare the deep connections between the electricity of Franklin, the magnetic lodestones of the ancient mariners, and the light that had filled the first eyes gazing into the heavens. And yet Maxwell’s equations, encompassing as they are, were only the beginning.

Of course, it is not enough to understand electricity and magnetism—one must also comprehend the stuff upon and within which those forces operate. Maxwell himself knew the behavior of certain kinds of materials, such as dielectrics (of which glass is a famous example), and many other important systems would emerge over the years: from the semiconductors in our computers to the black holes that warp the universe itself along with its light. (On the other hand, one material of Maxwell’s analyses turned out *not to exist*: the “luminiferous ether” was vanquished by relativity’s assertion that light propagates without a fixed medium.) Indeed, Maxwell’s electromagnetism has figured prominently in virtually every subsequent advance of physics, from quantum mechanics to relativity to the “standard model” of particle theory. Forgive us our myopia, however, if we allow ourselves to forget those extraordinary discoveries for a moment—what, then, remains to be found in Maxwell’s equations themselves?

An equation is only useful to the extent that you can solve it, and even in simple circumstances the solution of Maxwell's equations is not easy (as any student in an electromagnetism course will unhappily inform you). Nevertheless, the phenomena that can emerge in such basic problems were largely solved and understood by clever Maxwell and his contemporaries. Since then, it often seemed that the only remaining task was the "engineering" challenge of solving and utilizing those equations in ever more realistic and messier systems. Even the glassy fiber-optic cables that today span our oceans for telecommunications would surprise Maxwell only in their uses and in their perfection of manufacture, not in their underlying principles. The story of *condensed-matter physics*, however, has been that even simple, well-known equations can have unexpected consequences, especially in complicated circumstances.

1.2 Periodic Surprises

Milo tried very hard to understand all the things he'd been told, and all the things he'd seen, and, as he spoke, one curious thing still bothered him. "Why is it," he said quietly, "that quite often even the things which are correct just don't seem to be right?"

Norton Juster, *The Phantom Tollbooth*

Imagine shining a light through a large block of glass with a single bubble of air in it. When the light strikes the bubble, some of it will reflect and some of it will continue forward at a slightly different angle (be *refracted*). As you look through the glass, this scattered light allows you to see the bubble, perhaps with an attractive sparkling caused by all of the reflections and refractions. Picture now a second bubble in the glass, just like the first but at a different place. As before, the light will reflect and refract, this time from both bubbles, sparkling in a more intricate pattern than before. All of this is exactly predicted by Maxwell's equations, although solving for the precise details would probably require

the help of a computer. Next, suppose that we fill the glass with *millions* of bubbles, all identical and arranged in a perfectly periodic lattice extending in all directions. Surely, this is not much different from before, although the sparkles will form a far more complicated mess? (“Ugh,” says the engineer.) In fact, if you are *very* careful, you may see something quite different: *nothing at all*. What would Maxwell say?

James would be gratified to learn that he has not been contradicted—everything is predicted by his equations, as always. The engineer, working out the integrals carefully on her computer, would discover to her amazement that all of those fiendish reflections and refractions *cancel* one another before reaching our eyes, and the light passes through the crystal unimpeded. This is no accident, but is a general consequence of the periodicity of the structure—actually, the explanation is closely related to one of the great mysteries of physics from a century ago, in which the cancellation was of electrons rather than of light.

When you flip a wall switch to illuminate the dirty socks on your floor, an electric current—a flow of tiny, charged particles called *electrons*—carries your good intentions to the bulb. Unfortunately, this current requires constant power to maintain, as your monthly electric bill attests; otherwise, the electrons will halt their flow after caroming off an atom or two in the wire (generating heat and light as a side effect). Scientists of the 19th century, however, couldn’t leave well-enough alone: they set out to calculate and measure the rate at which electrons bounced off atoms. What they found was terribly strange. You see, the forces between charged particles are enormous—powerful enough to keep an electron hurtling around a Hydrogen nucleus at a speed of over 1000 miles per second in a circle four billionths of an inch across—so, an electron flying through the metal wire, surrounded by charged nuclei every few billionths of an inch (not to mention all the other electrons), shouldn’t get very far. And yet it was found that, in a pure chunk of a good conductor such as copper wire, an electron apparently meanders blithely past 100 or more nuclei before

bumping into one! The solution to this puzzle came in two pieces (some assembly required). First, in the 20th century, it was reluctantly realized that electrons are actually *waves*: rather than hard little bullets whizzing through space, they are smeared out in a ghostly probability envelope covering an entire volume at once—and, unlike bullets, they can interfere and even cancel if two waves are superimposed. Second, since a conductor like copper forms a periodic lattice of nuclei, a *crystal*, we can apply the lovely* theorem that we alluded to earlier with the bubbles.

This theorem, called alternatively “Bloch’s” [2] or “Floquet’s” [3] theorem depending upon whether you drink beer or wine, says that, in a *periodic* medium, waves can be found that propagate without scattering, and is equally true whether the waves consist of electrons or of light beams (or sound waves, or...). This is why the electrons can drift so far in the metal—they only bounce off of impurities or imperfections—and why a light beam could travel through our bubbly glass without reflecting towards our eyes. And that’s not all: the machinery of the Bloch-Floquet theorem leads us to many other curious phenomena...

1.3 Photonic Crystals

How do you solve a problem like $[\nabla \times \frac{1}{\epsilon} \nabla \times \hat{H} = \left(\frac{\omega}{c}\right)^2 \hat{H}]$?

How do you hold a moonbeam in your hand?

Oscar Hammerstein II, *The Sound of Music*

By combining Maxwell’s equations with the theorems of solid-state physics, a surprising and simple result has emerged from what seemed like a horrendous problem, that of light bouncing among an infinity of periodic scatterers. But there’s still more to be learned from the humble electron, a behavior of light that Maxwell never guessed although the equa-

* Physicists have strange attractions. (Mathematicians, contrariwise, have “strange attractors.”)

tions were right in front of his eyes: just as there are electrical insulators, which keep the currents in the wires where they belong, one can also build an optical insulator, a *photonic crystal*.

Photonic crystals again arise from the cooperation of *periodic* scatterers—thus, they are called “crystals” because of their periodicity and “photonic” because they act on light.* They can occur when the period (the separation of the bubbles or scatterers) is on the order of the wavelength of the light.† When that’s true, and if your bubbles are the right shape, it is possible that all of those reflections and refractions will cancel not only the light scattered sideways, but the light moving *forward* as well. Then, because the light has to go somewhere (energy is conserved), it has no choice but to go back—it is forbidden from entering the photonic crystal. This happens no matter what direction the light is coming from, in a certain range of wavelengths called the *photonic band gap*.‡ (A similar effect, an *electronic* band gap, is what gives electrical semiconductors their lucrative properties.)

Once you have such a medium, impervious to light, you can manipulate photons in many interesting ways. By carving a tunnel through the material, you have an optical “wire” from which no light can deviate. Even more dramatically, by making a cavity in the center of the crystal, you have an optical “cage” in which a beam of light could be caught and held—although you couldn’t easily admire it, because the very fact that it cannot escape would render it invisible. These kinds of abilities to trap and guide light have many potential applications in optical communications and computing, where one would like to make tiny optical “circuits” to help manage the ever-increasing traffic through the world’s

* Perversely, at about the same time that electrons were reborn as waves, Einstein discovered that the light waves of Maxwell’s equations are “particles” called *photons*.

† Light, like all waves, has a characteristic *wavelength*: the distance from one wave crest to the next, where each wavelength corresponds to a different color. For example, all of the colors that you can see fall in the range of about 0.4 to 0.6 microns—one micron is 1/1000 millimeters—but the wave crests are too small and move too fast for you to observe directly.

‡ Some people call these systems “photonic band-gap materials” instead of “photonic crystals.”

optical communications networks. Other devices, too, are made possible by this increased control over light: from more-efficient lasers and LED light sources, to opening new regimes for operating optical fibers, to cellular phones that don't pump half of their signal power into your head (as harmless heat, but wasteful).

What's so new about photonic crystals? Aren't they only mirrors, just like the one in which we admire ourselves every morning? Metal mirrors, however, besides operating on entirely different principles, have a major disadvantage: some of the light that strikes them is absorbed rather than reflected, and this absorption becomes intolerably large as you move towards the short wavelengths (infrared and visible) that are most important for lasers and optical communications.* On the other hand, photonic crystals rely only on weak interactions of the light with the material, and can have low absorption at nearly any wavelength.

Actually, a form of photonic crystal can be found in nature, producing the iridescent colors of abalone shells, butterfly wings, and some crystalline minerals. These beautiful materials archetypically comprise alternating microscopic layers of two substances (or two arrangements of the same substance)—they are only periodic in *one* direction, and thus form a *one-dimensional* (1d) photonic crystal, depicted in Fig. 1.1. When light shines on the layers, it is reflected by the photonic band gap.† Because the natural optical contrast between the layers is weak, however, the band gap is narrow—only a *specific color* is reflected, resulting in the brilliant hues we so admire. Moreover, as the angle of the light changes, the periodicity in that direction alters and the band gap shifts wavelengths—that is, different colors are reflected at different angles, shimmering as we tilt our heads.

* Metal mirrors work just fine for *microwave* wavelengths, and Bell Labs tried for many years to use metal tubes as microwave communications “pipes.” How these tubes eventually lost the battle against today's glass fibers makes an interesting story. (For one thing, you couldn't bend the tube.)

† The classical way to understand the 1d band gap is that, if the periodicity is half a wavelength, the cumulative reflections are *in phase*: the reflected wave crests line up, and so are additive.

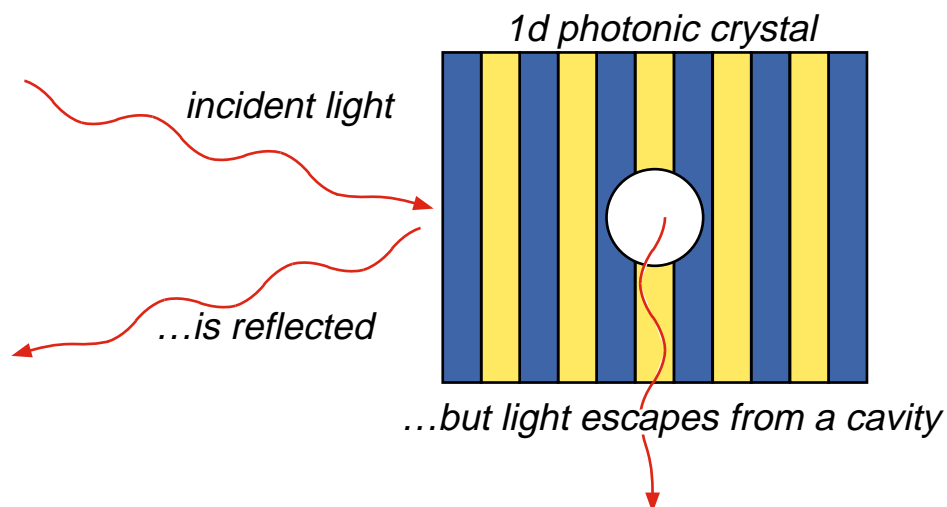


Figure 1.1: Schematic of a *one-dimensional* photonic crystal: alternating dielectric layers (represented by blue and yellow), with a band gap in one direction. Incident light is reflected for colors of light in the band gap, but it cannot be used to trap light in 3d—the light inevitably escapes “sideways.”

Just such a crystalline mineral attracted the interest of Lord Rayleigh in 1887, and he was (it seems) the first to formulate a precise theory of these phenomena [4]–[5]. In fact, Rayleigh stumbled onto a form of the Bloch-Floquet theorem, deriving quite a general solution for waves in 1d periodic media;* he even correctly showed that *any* such medium exhibits a band gap (although he didn’t call it that). Unfortunately, he failed to pursue this result to its full generality—instead, he later abandoned it for a more direct (but less illuminating) method: performing a heroic calculation to sum the waves reflected and refracted from each interface [6]. No one suspected that the same magical cancellation could happen for light shining from *any* direction in a crystal periodic along all *three* axes. No one, that is, until 100 years later, in 1987, when the first papers were published marrying solid-state physics with Maxwell’s equations and suggesting that a 3d crystal could produce an omnidirectional band gap [98]–[99]. “Suggested,” because an actual *example*

* Rayleigh was guided in his analysis of periodic media by earlier studies of the lunar orbit, in which the moon is periodically tugged by the sun. If the generality of that work had been recognized, we might now call it the “Hill” theorem (since it predates Floquet by six years) [7].

(even in theory) was not found until three years later—unlike in 1d, any old periodic structure will not do, and much trial-and-error was required. The first concrete design involved levitating dielectric balls [101], so the engineers asked to build the thing were understandably upset—especially since the balls had to be sub-microns in size for it to work at infrared wavelengths. Nevertheless, after several years, more realistic structures were proposed and even fabricated, using technologies such as those by which tiny circuits are etched onto computer chips.

The challenge of realizing the benefits of photonic crystals in practical systems continues to be a major subject for research. The only such devices currently in production are still based on one-dimensional band gaps like the crystals of Lord Rayleigh (often called *Bragg mirrors*^{*}), and range from dielectric mirrors and reflective coatings to DFB lasers.

Before we continue, we must take a moment to apologize: if it has seemed to you that our explanations of photonic crystals have often failed to explain, then we are in complete agreement. A more substantial understanding can only come after one has had the pleasure of swimming through the actual mathematics of Maxwell's equations and Bloch-Floquet eigensystems, and that would require a work somewhat longer than this chapter. Fortunately, we are able to recommend to the interested reader a marvelous introductory text by an author unrivaled in his eloquence and clarity, and whose signature is coincidentally required on this thesis. That book, cited in Ref. [8], should be accessible to anyone with an undergraduate background in electromagnetism. (A briefer review can be found in the *Nature* article of Ref. [9].) Alternatively, Chapter 2 of this thesis provides a whirlwind tour for the impatient and the mathematically inclined.

* Named after Sir William Henry Bragg, who in 1915 received the Nobel prize in physics for his work on X-ray crystallography—probing periodic *atomic* crystals by bouncing X-rays off them.

1.4 Things to Come

*I'm very well acquainted too with matters mathematical,
I understand equations, both the simple and quadratical,
About binomial theorem I'm teeming with a lot o' news—
With many cheerful facts about the square of the hypotenuse.*

Sir William Schwenk Gilbert, *The Pirates of Penzance*

Although we cannot take credit for the discoveries described in the previous sections, we will endeavor in the subsequent chapters to contribute some marginal advance to the understanding of photonic crystals and related systems. Chapter 2, however, is simply a mathematical introduction to photonic crystals, while Chapter 9 is our conclusion. The remaining chapters are derived from a series of articles published in various scientific journals, meaning that they are either pleasantly self-contained or irretrievably abstruse, depending upon one's perspective. In any case, we shall summarize them here.

1.4.1 Intersecting Optical Pipes

To begin with, we consider an interesting effect that one can achieve with photonic crystals, and more generally with the mathematical and conceptual techniques that they embody. In particular, we show how two optical “wires” can be crossed without interfering, which might be surprising if you think in terms of another wave: sound.

Imagine that you face an intersection of two corridors, deep within the bowels of MIT. On the opposite side of the intersection is one of your fellow graduate students, but you can't reach her because it's noon* and the janitors have roped off the crossing to wax the floor. Instead, you call across to brag that you used the Boston White Pages to pad out your thesis: “Our batty advisor won't even notice!” Meanwhile, speak of the devil, *he* is having his own shouted conversation with a colleague across the same intersection, but in the per-

* In Tech time—2AM, Boston time.

pendicular direction: “No need to feel guilty—he’s young, and has plenty of time to find another trade if we reject his dissertation.” Naturally, neither of you wants the other to overhear his words—is this possible? (No fair talking through tubes, either; both voices must pass through the *same space*.) It turns out that, in certain cases, the answer is *yes*.

You can see where this is heading: an almost identical situation shows up with light, and poses a substantial problem in designing optical circuits. In order to weave together a group of components, one needs the ability to cross waveguides (optical “corridors”). Doing so in the naïve way, however, would allow the crossed channels to overhear one another—this is called *crosstalk*, and will turn your device into a lemon. (With electronic circuits, this difficulty is circumvented by passing the wires over or under one another, but that solution is less practical in optics for a variety of reasons.)

In Chapter 3, we show how, by exploiting general principles of symmetry and resonance, it is possible to design an optical crossing where the two waveguides do not “see” one another—light travels across in one direction or the other, through the same space, without leaking into the perpendicular pipe. Although the inspiration and the ideal realization of this concept resides in full photonic crystals, we show how it can be applied to conventional and hybrid optical systems as well. The application to eavesdropping advisors will, perhaps, be the subject of a future publication.

1.4.2 Photonic-Crystal Slabs

Photonic crystals, in their full three-dimensional glory, are hard to make. Imagine baking ten-layer cake, where each layer is a checkerboard pattern of chocolate and vanilla, and the patterns of adjacent layers must be offset by exactly half a checker. Oh, and the cake is only 1/1000 of an inch high. Wouldn’t it be much easier to bake only a single layer? You can roll out the dough as flat as you need, and even use a cookie-cutter to make the pattern. The optical equivalent of this recipe shortcut is a *photonic-crystal slab*: a

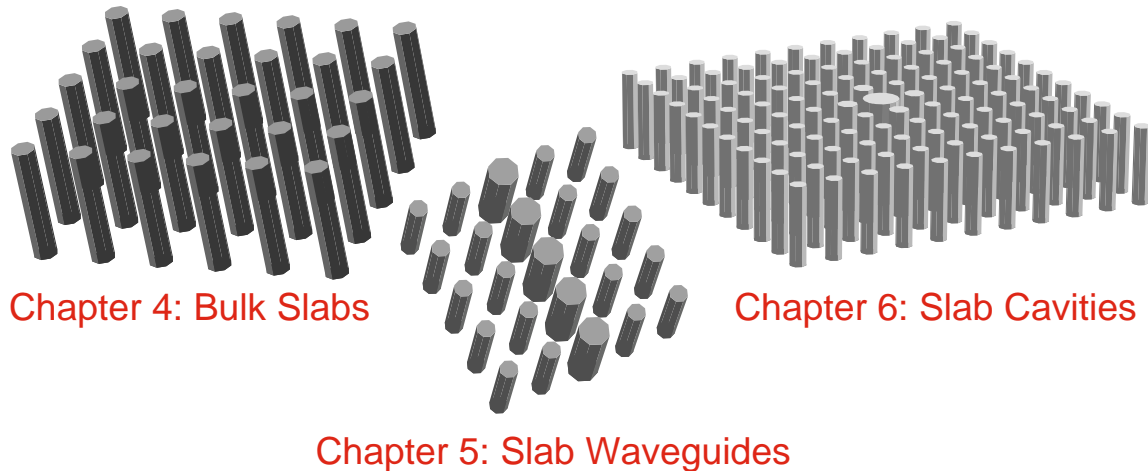


Figure 1.2: Schematic of the photonic-crystal slab chapters, covering first the bulk systems, then waveguides formed by linear defects, then cavities formed by point defects.

dielectric (*e.g.*, Silicon) with only a two-dimensional pattern and a finite thickness, which can be baked via the well-known lithographic techniques used to manufacture computer chips. Slabs are very attractive to the engineers, but they involve a compromise—what confines the light in the vertical direction? Actually, there *is* a confinement mechanism, albeit imperfect: *index guiding*, sometimes misleadingly called *total internal reflection*.

Index guiding is the principle behind today’s optical fibers; simply put, it is the tendency of light to stay in stronger dielectrics (which have a higher *index of refraction*) compared to weaker dielectrics such as air. Just as a bump in the road may throw your car into the ditch, however, any irregularity in the medium can cause light to scatter out, and this ease of loss makes life more difficult with a slab than in a full 3d crystal (where the road is surrounded by impenetrable walls). Fortunately, thanks again to the Bloch-Floquet theorem, at least the *periodic* variation of the slab itself does not cause scattering; if only regularly-spaced bumps did not disturb your driving!

In the next three chapters, as depicted in Fig. 1.2, we explore and explain the unique phenomena that can occur in such hybrid systems, probing both the limitations and the possibilities of photonic-crystal slabs. Chapter 4 lays the foundations by characterizing the

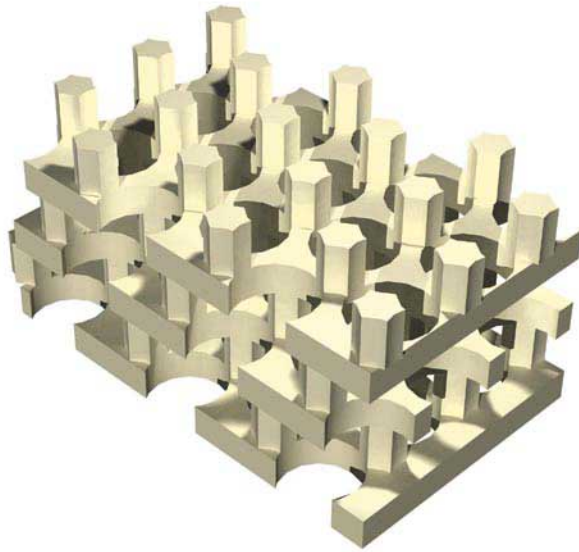


Figure 1.3: A portion of a novel 3d photonic crystal that we propose in Chapter 7. This structure has a large, complete band gap, can be constructed layer-by-layer using traditional lithography, and its behaviour is understandable in terms of much-simpler two-dimensional photonic-crystal structures.

bulk slabs themselves, without defects. Chapter 5 breaks the slab's periodicity in only one direction by a linear defect, showing how lossless waveguides can be formed. Finally, Chapter 6 breaks the periodicity entirely and attempts to trap light in a cavity. This is not possible, however, due to the lack of a full band gap, and so we propose and investigate various mechanisms to keep the light confined for as long as we can.

1.4.3 A New Photonic Crystal

Despite their difficulty of manufacture, tremendous progress has been made in the fabrication of full three-dimensional photonic crystals. Thus far, however, the menu of practical structures has been limited to a mere handful of possibilities, and so the prospect of new designs with unique and desirable properties holds great attraction in the field. In Chapter 7, we propose a novel photonic crystal that can be (and, in fact, has been) realized even at microscopic lengthscales. Although it may look complex, such a periodic system can easily be analyzed in its entirety with the help of our friends Bloch and Floquet.

Our structure, depicted in Fig. 1.3, has a number of appealing qualities whose combination is unique. First, its photonic band gap covers a wide range and persists even if only weak dielectric materials are employed. Second, it consists of a sequence of patterned layers that can be constructed one at a time via the “cookie-cutter” of modern micro-lithography. Third, each layer has high symmetry (it looks the same from six directions in the plane). Physicists love symmetry for its own sake, but here it has substantial practical advantages as well—it means that complex optical circuits can be constructed by modifying only a *single* layer of the structure. Finally, each layer resembles a common *two-dimensional* photonic-crystal structure, and this allows us to understand its behavior in terms of those vastly simpler systems (for which a substantial body of literature, intuition, and computation is already available).

1.4.4 So, You Want To Study Photonic Crystals?

Finally, we consider a matter less of physics than of mathematics—given the equations of Maxwell and Bloch-Floquet, it is still a substantial challenge to compute their solutions. Several methods are already known, of course, but in Chapter 8 we attempt to bring together the best developments in numerical analysis, computer science, and physics in order to take a certain class of techniques further towards its ultimate conclusion, and to provide a comprehensive review and comparison of the ideas of ourselves and others. In this way, we hope to furnish a valuable reference for the reader who wishes to engage in these kinds of computations. We offer more than that, however...

In the literature, all too often one sees a proposal for a new computational algorithm, usually with a few sample results, that nevertheless leaves the reader with no way to evaluate or employ the method for herself without investing several months or even years in programming her own implementation on a computer. Moreover, despite the fact that an entire field of scientists and engineers runs the same type of computation on similar types

of problems, to a large extent each group of scientists develops their own software. This duplication of effort is, it seems to us, extraordinarily wasteful and unnecessary. With that in mind, and in the scientific spirit of contributing freely to human knowledge, we have made the results of our work available as free software on the Web—see Ref. [122], <http://ab-initio.mit.edu/mpb/>. Using it, we invite you analyze for yourself the band gaps and the propagation of light in arbitrary three-dimensional photonic crystals.

*And like a silver clarion rung
The accents of that unknown tongue,
Excelsior!*

Henry Wadsworth Longfellow, “Excelsior”

Chapter 2

Band Gaps & Maxwell's Equations: A Whirlwind Tour

2.1 Maxwell's Equations

In this chapter, we very briefly review the properties of Maxwell's equations and the Bloch-Floquet theorem, and discuss the origin of the band gap in photonic crystals. We assume herein a fair amount of mathematical sophistication; see Ref. [8] for a gentler introduction. Our starting point, naturally, is Maxwell's equations in cgs units (written in a slightly more modern form than the 20 equations in 20 variables used by Maxwell [1]):

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{H}}{\partial t}, \quad (2.1)$$

$$\vec{\nabla} \times \vec{H} = \frac{4\pi}{c} \vec{J} + \frac{1}{c} \frac{\partial \epsilon \vec{E}}{\partial t}, \quad (2.2)$$

$$\vec{\nabla} \cdot \epsilon \vec{E} = 4\pi \rho, \quad (2.3)$$

$$\text{and } \vec{\nabla} \cdot \vec{H} = 0, \quad (2.4)$$

where \vec{E} and \vec{H} are the electric and magnetic fields, \vec{J} is the free current density, ρ is the free charge density, and ϵ is the dielectric function; all of these are potentially functions of both the position \vec{x} and the time t . c , of course, is the speed of light in vacuum.

Recall that a *dielectric* is a material in which tiny dipoles, the bound charges, line up to oppose any applied electric field—this process is encapsulated entirely by the dielectric function ϵ . In particular, we consider here the case of *linear* dielectrics, for which ϵ is independent of \vec{E} —this is reasonable, as nonlinear effects in optical materials are typically so weak that you cannot observe them without enormous fields or great distances of propagation. We will also assume that ϵ is independent of frequency (and therefore independent of time as well), which is reasonable for the relatively narrow frequency ranges

that will be of interest in photonic crystals. Finally, we assume that we can neglect any absorption of light by the material; in this case, ϵ is a real number.

Furthermore, in optics we are normally interested only in cases where no free charges or currents are included, so \vec{J} and ρ are zero. If we then combine Eq. (2.1) and Eq. (2.2), we can eliminate \vec{E} in favor of \vec{H} to write:

$$\vec{\nabla} \times \frac{1}{\epsilon} \vec{\nabla} \times \vec{H} = \frac{\partial^2 \vec{H}}{\partial ct^2}. \quad (2.5)$$

This equation can be compared to Schrödinger's equation from quantum mechanics, with a Hermitian (self-adjoint) “Hamiltonian” operator $\vec{\nabla} \times \frac{1}{\epsilon} \vec{\nabla} \times$ on the left (instead of Schrödinger's $-\nabla^2 + V$) and a time derivative $\frac{\partial^2}{\partial ct^2}$ on the right (instead of $i\frac{\partial}{\partial t}$). Here, $\frac{1}{\epsilon}$ plays a role analogous to that of the potential function V . Unlike in quantum mechanics, we also have the *transversality* constraint of Eq. (2.4), but this merely means that we operate in a subspace of the set of all possible eigensolutions; it has no effect on most of the following discussion.

Next, like in quantum mechanics, we look for time-harmonic states whose time dependence is $e^{-i\omega t}$ for some (angular) frequency ω —all possible solutions can be expressed via this form, since the equation is linear—and thus Eq. (2.5) becomes:

$$\vec{\nabla} \times \frac{1}{\epsilon} \vec{\nabla} \times \vec{H} = \left(\frac{\omega}{c}\right)^2 \vec{H}. \quad (2.6)$$

This is a Hermitian eigenproblem over an infinite domain, and generally produces a continuous spectrum of eigenfrequencies ω . Things become much more interesting, however, when we consider in the next section the case of a *periodic* dielectric structure ϵ .

When we say that the eigen-operator is “Hermitian,” we are referring to the fact, easily proved from the identities of vector calculus, that:

$$\int \vec{A}^* \cdot \vec{\nabla} \times \frac{1}{\epsilon} \vec{\nabla} \times \vec{B} = \int \left(\vec{\nabla} \times \frac{1}{\epsilon} \vec{\nabla} \times \vec{A} \right)^* \cdot \vec{B}, \quad (2.7)$$

where \vec{A} and \vec{B} are arbitrary fields with boundary conditions such that the boundary terms in the above integrals disappear (for example, if they are periodic or have compact support). In other words, the operator acts the same when applied to the left or to the right under inner products (integrals). It's also worth noting that this operator is positive semi-definite. The combination of these two properties implies that the eigenfrequencies ω in Eq. (2.6) are real numbers.

We should note one important property of Maxwell's equations in general, and of Eq. (2.6) in particular: they are *scale-independent*. If you multiply all the dimensions of your system by 10, say, the solutions are exactly the same except that the frequencies are divided by 10. This means that we can solve the equations *once* and then apply the same results to problems at all lengthscales and frequencies. Because of this scale invariance, it is convenient to use dimensionless units for distance and time: we pick some natural lengthscale a in the system (usually the periodicity), and then express all distances as a multiple of a and all angular frequencies ω as a multiple of $2\pi c/a$. (This is equivalent to writing the frequencies as a/λ , where λ is the vacuum wavelength.)

2.2 The Bloch-Floquet Theorem and Electromagnetism

The Bloch-Floquet theorem [152] tells us that, for a Hermitian eigenproblem (or actually, a larger class of differential equations) whose operators are *periodic* functions of position, the solutions can always be chosen of the form $e^{i\vec{k}\cdot\vec{x}} \cdot (\text{periodic function})$.^{*} A periodic function $f(\vec{x})$ is one such that $f(\vec{x} + \vec{R}_i) = f(\vec{x})$ for any *primitive lattice vector* \vec{R}_i of the crystal—there can be up to three such lattice vectors, if the function is periodic in all three dimensions. A corollary of this theorem is that the wavevector \vec{k} is a conserved quantity, and hence propagating waves with a fixed \vec{k} cannot scatter.

^{*} This theorem is easily proven by noting that the translation operator of the crystal commutes with the Hamiltonian operator, and then recalling that simultaneous eigenfunctions can be chosen for commuting operators.

When applied to our situation, the Bloch-Floquet theorem means that the solution of Eq. (2.6) for a periodic ϵ can be chosen of the form:

$$\vec{H} = e^{i(\vec{k} \cdot \vec{x} - \omega t)} \vec{H}_{\vec{k}}, \quad (2.8)$$

where $\vec{H}_{\vec{k}}$ is a periodic function of position. If we substitute Eq. (2.8) into Eq. (2.5), we find that the function $\vec{H}_{\vec{k}}$ satisfies the “reduced” Hermitian eigenproblem:

$$(\vec{\nabla} + i\vec{k}) \times \frac{1}{\epsilon} (\vec{\nabla} + i\vec{k}) \times \vec{H}_{\vec{k}} = \left(\frac{\omega}{c}\right)^2 \vec{H}_{\vec{k}}. \quad (2.9)$$

Because $\vec{H}_{\vec{k}}$ is periodic, we need only consider this eigenproblem over a *finite* domain: the unit cell of the periodicity. There is a general theorem that eigenproblems with a finite domain have a *discrete* set of eigenvalues. Thus, the eigenfrequencies ω are a countable sequence of continuous functions: $\omega_n(\vec{k})$ (for $n = 1, 2, 3, \dots$). When they are plotted as a function of the wavevector \vec{k} , these frequency “bands” form the *band structure* of the crystal.

Another important consequence of the Bloch-Floquet theorem is that the solutions are periodic as a function of the wavevector \vec{k} ; this is easy to see if we examine a one-dimensional example. Suppose that we have a one-dimensional structure with period a ($\epsilon(x+a) = \epsilon(x)$) and consider the eigenstates for the wavevector $k + 2\pi/a$. The function $e^{i2\pi x/a}$ is periodic (with period a), however, so this part of Eq. (2.8) can be absorbed into the function $\vec{H}_{\vec{k}}$! Since the resulting expression then solves the eigenproblem for a wavevector k , and those solutions are unique, we can therefore conclude that the eigen-solutions at $k + \frac{2\pi}{a}$ are the same as those at k . Because k is only unique to within multiples of $\frac{2\pi}{a}$, then, it is sufficient to *only* solve the eigenproblem for k in a finite range, $-\frac{\pi}{a} \dots \frac{\pi}{a}$, called the *first Brillouin zone*. Moreover, we usually have additional symmetries that further reduce the range of unique wavevectors—for example, time-reversal symmetry tells us that k and $-k$ produce the same results, so in this case we can look at just $k = 0 \dots \frac{\pi}{a}$.

This symmetry-reduced region is called the *irreducible Brillouin zone*, and in three dimensions is given by well-known polyhedra [13] (whose vertices have canonical, obscure labels like Γ , X , K , *etc.*).

Let us make two additional comments before we continue. First, in the special case where the crystal has *translational symmetry* ($a \rightarrow 0$), the irreducible Brillouin zone extends to $k = \infty$, and is called a *dispersion relation*; this is commonly used to characterize waveguides. Second, when one is interested in just the extrema of the frequency bands, it is usually sufficient to consider only wavevectors along the edges of the irreducible Brillouin zone (or often merely the vertices)—this allows the band structure to be plotted in an ordinary two-dimensional diagram, many examples of which are found in this thesis.

2.3 The Origin of the Photonic Band Gap

Now, we will attempt to make some general arguments regarding the possibility of a photonic band gap: a range of ω into which there intrude no propagating states $\omega_n(\vec{k})$. In order to better understand the gap, we will exploit two general properties of Hermitian eigenvalue problems. First of all, the eigenfunctions must be orthogonal; that is, their inner product (integral over the unit cell) is zero:

$$\int \vec{H}_{\vec{k}}^{(m)*} \cdot \vec{H}_{\vec{k}}^{(n)} = 0 \quad (2.10)$$

for any two eigenstates ($m \neq n$) at a given \vec{k} point. Second, the lowest band solves a *variational problem*:

$$\omega_1(\vec{k})^2 = \min_{\vec{H}_{\vec{k}}} \frac{\int |(\nabla + i\vec{k}) \times \vec{H}_{\vec{k}}|^2 / \epsilon}{\int |\vec{H}_{\vec{k}}|^2} c^2. \quad (2.11)$$

(The numerator is just the expectation value of the eigen-operator.) The second band satisfies the same variational problem, with the added constraint that it be orthogonal to the first band according to Eq. (2.10), and so on. Note that the curl of the magnetic eigenfield,

in the numerator, is simply proportional to the electric field from Eq. (2.2).

In order to minimize the variational problem of Eq. (2.11), the field for the lowest band should try to satisfy two conditions. First, its curl should not be too big—that is, it should vary slowly, and not contain any fast oscillations. This is analogous, in quantum mechanics, to having a low kinetic energy. Second, the curl—the electric field—should be concentrated in regions of *high* dielectric constant ϵ ; this is analogous to having a low potential energy. So far, so good, but where’s the band gap? In order to see the gap, we naturally must look at the *second* band. Like the first band, the second band also wants to oscillate slowly and be concentrated in the high dielectric. However, it must also be orthogonal to the first band from Eq. (2.10). In order to accomplish this, the second band must be concentrated in the lower dielectric and/or have a sign oscillation inside the high dielectric to make the integral zero—both of these possibilities will incur a large increase in frequency in Eq. (2.11), a band gap!

Actually, we’re not quite there yet. True, there will generally be a large ω difference between the first and second bands, a “gap” at each \vec{k} point. In order to have an *omnidirectional* band gap, however, the gaps at *all* \vec{k} points must overlap in some frequency range. For this reason, a fairly large contrast in ϵ between the high and low dielectric regions is usually necessary to make the gaps in different directions large enough to intersect.

The situation is even worse than that, however, because the fields have a direction as well as a sign and a magnitude. In order to be orthogonal to the first band according to Eq. (2.10), the second band need only be perpendicular to the first, without sacrificing its concentration in the high dielectric! To combat this phenomenon, we need to make use of the boundary conditions on the fields, from Maxwell’s equations, that apply at dielectric interfaces. In particular, the surface-parallel component of the curl of \vec{H} is allowed to be discontinuous across a dielectric boundary, while the surface-normal component must be

continuous. Now, suppose that our high dielectric consists of thin veins surrounded by lower dielectric. In that case, the numerator of Eq. (2.11) can be “weighted” *discontinuously* in these veins for electric fields running parallel to them, but *not* for electric fields perpendicular to the veins. This allows us to discriminate between bands with fields in different directions, and makes a complete band gap possible despite the variety of field polarizations. (Even so, complete gaps typically arise between the second and third bands, not between the first and second, precisely because of polarization-related degeneracies.)

When these conditions are satisfied, and if you are very lucky, your periodic crystal will have a complete photonic band gap in three dimension. We’ve waved our hands quite a bit in this section, partially in the interest of brevity, but partially because there are no rigorous criteria for the existence of a complete three-dimensional gap. In contrast, for one dimensional systems, life is easy—*any* nonzero periodic dielectric variation will result in a complete band gap (although it may be small)! This has a simple explanation via perturbation theory that, unfortunately, exceeds the scope of this chapter. A very similar argument, however, implies that your best chance for a complete gap in three dimensions resides in crystals with nearly spherical Brillouin zones. The crystal lattice with the B. Z. closest to a sphere is the fcc (face-centered cubic) lattice,^{*} and almost all known photonic crystals in three dimensions fall into this category (including the crystal we introduce in Chapter 7). Furthermore, the same derivation also explains why the gap arises only for wavelengths of about twice the crystal periodicity. We hope, nevertheless, to have armed the reader with sufficient knowledge to brave the remaining chapters.

* There is class of non-periodic structures called *quasi-crystals* that, in some sense, have more-spherical Brillouin zones, at the cost of additional complexity and some other disadvantages. There is evidence that these structures can have substantial band gaps even with low dielectric contrast.

Chapter 3

Elimination of Crosstalk in Waveguide Intersections*

3.1 Introduction

The ability to intersect waveguides is crucial in constructing integrated optical circuits, due the desire for complex systems involving multiple waveguides. We present a novel theoretical framework for achieving low crosstalk and high throughput in perpendicular intersections, based on symmetry considerations that can be systematically applied to any optical system. The transmission spectra are predictable without detailed calculations and are robust under perturbations that do not break the symmetry. The focus of this design is the elimination of crosstalk; additional considerations are required to maximize throughput when radiation losses are present. We illustrate the theory with two example systems—a two-dimensional photonic crystal and a conventional dielectric-waveguide intersection—and show how an efficient crossing can be produced by assembling well-understood elements that are independently tuned. In contrast, previous studies of waveguide intersections [10]–[11] have lacked general principles that could be applied *a priori* to diverse systems. Moreover, they have typically been concerned with shallow-angle crossings for wavelengths many times smaller than the waveguide width. Although perpendicular crossings in such systems exhibit negligible crosstalk, the crosstalk is close to 10% when the waveguide width is on the order of half a wavelength (as seen below). Our design applies even when the waveguide width is small, permitting single-mode waveguides with optimal miniaturization. Simulations of our structures show peak throughputs of nearly unity while crosstalk falls as low as 10^{-9} .

* This chapter is based on: S. G. Johnson, C. Manolatou, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, “Elimination of cross talk in waveguide intersections,” *Optics Lett.* **23**, 1855–1857 (1998).

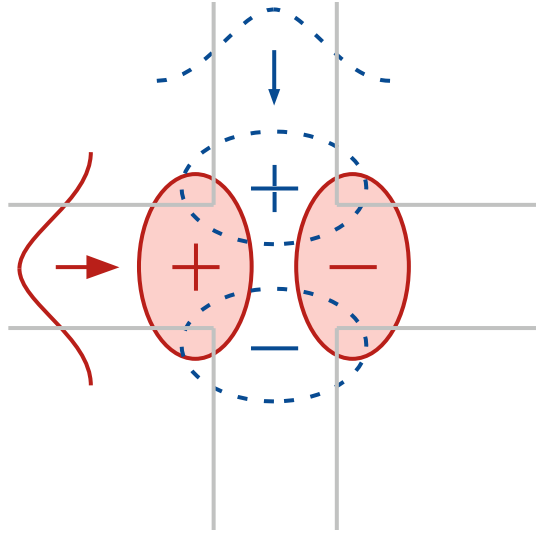


Figure 3.1: Abstract diagram of symmetry requirements for waveguide crossing, showing waveguide mode profiles and resonant-cavity mode contours. By symmetry, the solid-line (red) modes cannot couple with the dashed-line (blue) modes and vice-versa.

3.2 Eliminating Crosstalk by Symmetry

The fundamental idea is to consider coupling of the four branches, or ports, of the intersection in terms of a resonant cavity at the center. If the resonant modes excited from the input port can be prevented by symmetry from decaying into the transverse ports, then crosstalk is eliminated and the system reduces to the well-known phenomenon of resonant tunneling through a cavity. This situation can be achieved by means of the following conditions (see also Fig. 3.1):

- Each waveguide must have a mirror symmetry plane through its axis and perpendicular to the other waveguide, and have a single guided mode in the frequency range of interest. This mode will be either even or odd with respect to the mirror plane.
- The center of the intersection must be occupied by a resonant cavity that respects the mirror planes of both waveguides.

- Two resonant modes must exist in the cavity, each of which is even with respect to one waveguide’s mirror plane and odd with respect to the other.

These should be the only resonant modes in the frequency range of interest.

If these requirements are satisfied, then each resonant state will couple to modes in just one waveguide and be orthogonal to modes in the other waveguide, as shown in Fig. 3.1. (For simplicity, we depict a lowest-order even waveguide mode.) Therefore, under the approximation that the ports only couple to one another through the resonant cavity, crosstalk will be prohibited. The transmission to the output port is described by resonant tunneling, and one can use coupled-mode theory [12] to show that the throughput spectrum will be a Lorentzian peaked at unity on resonance. The width of the Lorentzian is given by the inverse of the cavity’s quality factor Q , which is proportional to the lifetime of the resonance mode.

Perfect throughput will not be attained, however, due to a combination of two effects: direct coupling and radiation loss. First, the input port can directly couple with the transverse ports, resulting in crosstalk—this can be decreased arbitrarily, but at the expense of the throughput bandwidth. Direct coupling is reduced by making the cavity larger, since this increases the distance between the localized waveguide modes. The drawback to increased cavity size is that it can either create extraneous resonance modes or, as in the case of the photonic crystal cavities below, increase the Q of the resonance. Second, many systems suffer from losses due to radiation or other mechanisms. If these losses do not couple strongly to the transverse waveguide, then their primary effect will be to decrease throughput and increase reflection [12].

The required parity of the cavity modes is easy to achieve in practice, due to the fact that a perpendicular intersection of identical waveguides typically has C_{4v} symmetry (the

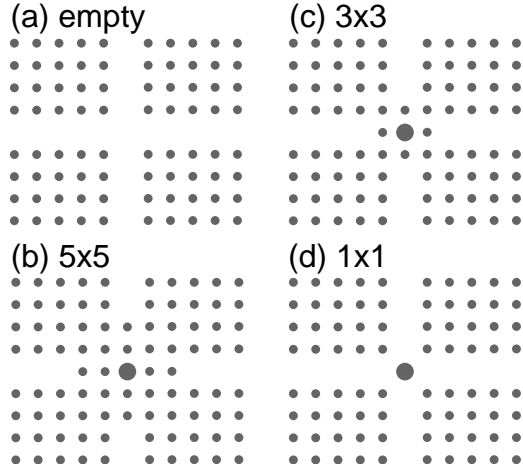


Figure 3.2: Waveguide intersections in a two-dimensional photonic crystal consisting of a square lattice of dielectric rods in air.

symmetry group of the square). In this case, any non-accidental degenerate modes (such as many higher-order modes) will necessarily be a pair having the requisite symmetry from Fig. 3.1, as that is the only two-dimensional irreducible representation of C_{4v} [13].

3.3 Intersecting Photonic-Crystal Waveguides

A realization of the intersection design can be made in a photonic crystal [14], for which the band gap in the bulk material prevents any radiation losses. For ease of computation, we use a two-dimensional crystal, consisting of a square lattice of dielectric rods in air. The rods have lattice constant a , radius $0.2a$, and dielectric constant 11.56. This structure prohibits propagation of TM light (in-plane magnetic field) in the frequency range 0.286 to 0.421 c/a [14]. Removing a row or column of rods produces a waveguide [14]. A resonant cavity is created by increasing the radius of a single rod to $0.33a$, leading to doubly-degenerate modes of the requisite symmetry with a frequency of 0.361 c/a (*i.e.* a wavelength in air of $2.77a$) [15].

These elements are then combined into an intersection, and we simulate the collective behavior of several structural variations (Fig. 3.2) using a finite-difference time-domain simulation [16]. In each case, we send a wide-spectrum TM gaussian pulse to the input

port (left). The fractional power transmission is then evaluated as a function of frequency for the output port (right) and one of the transverse ports (top), yielding the throughput and crosstalk, respectively. The resulting transmission spectra are shown in Fig. 3.3. Up to five percent of each throughput and crosstalk value can be due to reflections from the boundary of the simulation, which is why the peak throughput in cases (c) and (d) appears to exceed unity.

As predicted, the throughput for cases (b) and (c) is a Lorentzian centered on the resonance frequency (which varies due to the finite size of the cavity). At their resonance frequencies, (b) and (c) achieve crosstalks of 5×10^{-10} and 4×10^{-4} , respectively. In case (d), there are no photonic crystal layers separating the central rod from the waveguide, and confinement is due only to index contrast. The cavity confinement is so weak that coupled-mode theory is no longer accurate. Even so, case (d) reaches nearly 100% throughput with a crosstalk at the same frequency of 1.8%. In contrast, both the throughput and the crosstalk for the empty intersection lie in the 20–40% range.

Both the crosstalk and the resonance Q derive from the exponential decay of light through the bulk photonic crystal. Because the crosstalk must tunnel through twice as many crystal layers as the resonance decay, the crosstalk should be approximately Q^{-2} . Looking at Fig. 3.3(b), this estimate for the crosstalk is seen to be valid to within about an order of magnitude, corresponding to the Q's of roughly 1000, 100, and 10 for cases (b), (c), and (d). Here, we do not consider the additional resonance minima apparent in the crosstalk at some frequencies, which are independent of the throughput resonance and are due to a destructive-interference phenomenon that is still under investigation.

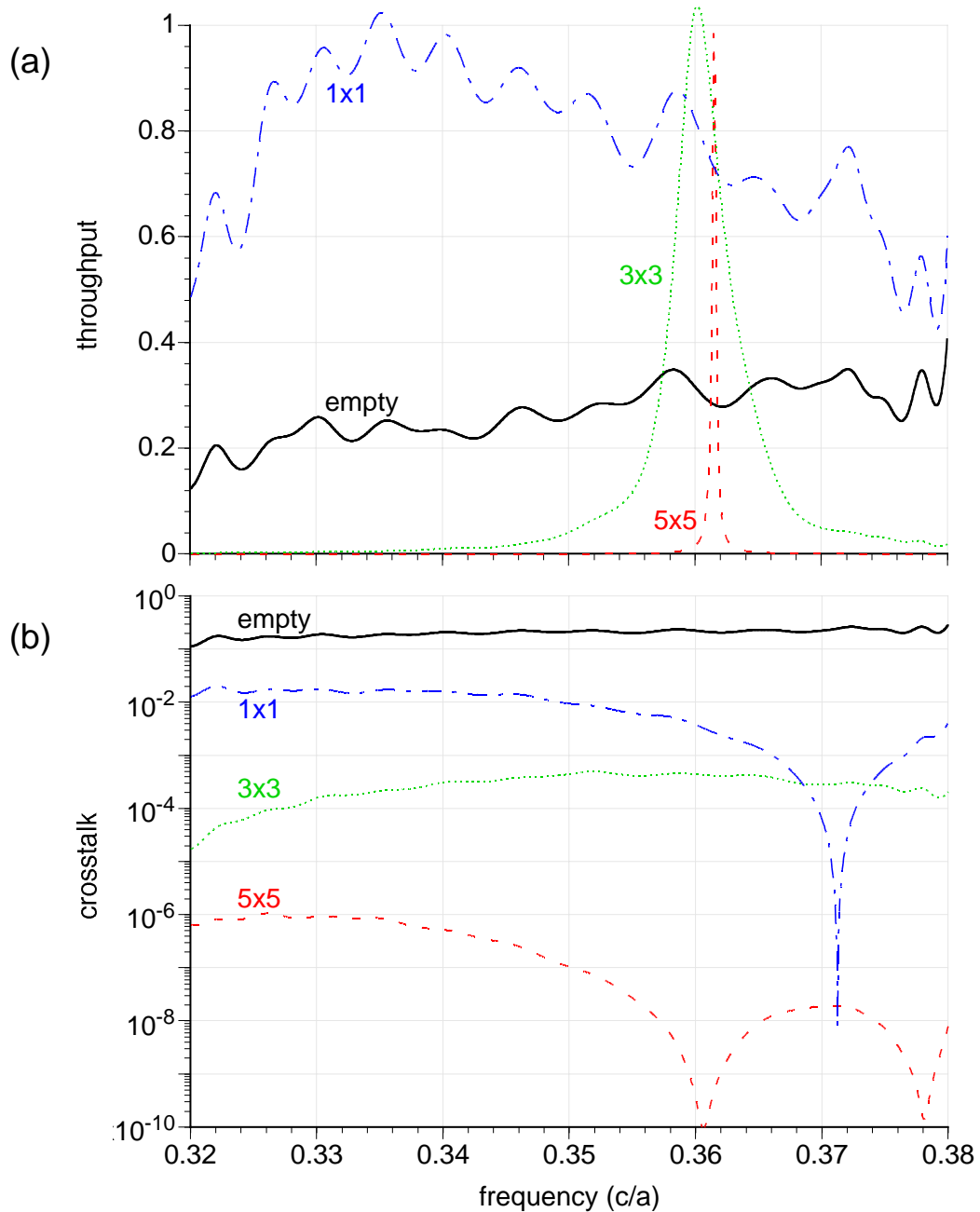


Figure 3.3: TM transmission spectra for the four intersections shown in Fig. 3.2: (a) throughput from the input port to the output port and (b) crosstalk from the input port to one of the transverse ports.

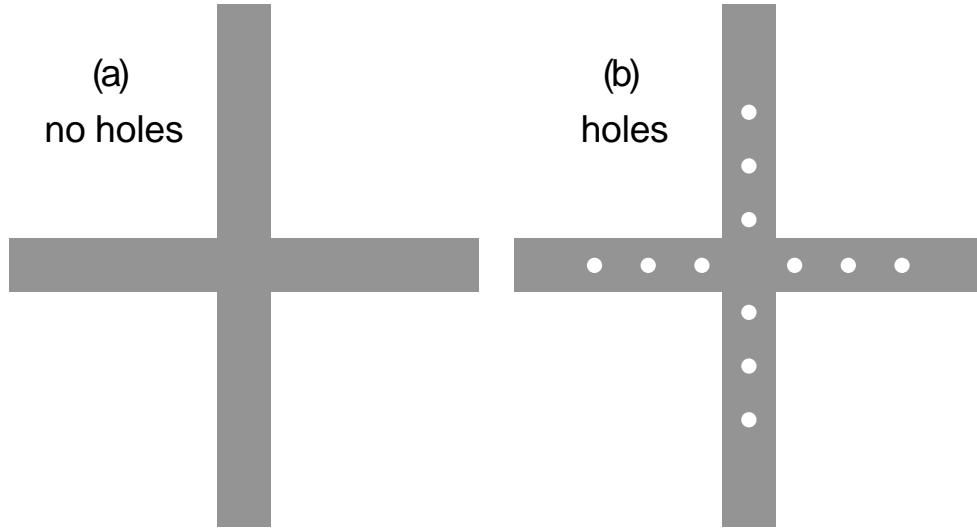


Figure 3.4: Intersection of two-dimensional, high index-contrast waveguides in air, both (a) without holes and (b) with air holes added to create a resonant cavity using a photonic band gap.

3.4 Intersecting Conventional Waveguides

The same principles of intersection design can be applied to other systems, such as conventional dielectric waveguide crossings. In particular, we considered an intersection of high index-contrast waveguides in air, of width a and dielectric constant 11.56, as shown in Fig. 3.4(a). In Fig. 3.4(b), we have modified the crossing to include a resonant cavity by introducing a periodic sequence of air holes on each branch of the intersection, creating a band gap in the TE guided modes of the waveguides [17]. The holes have radius $0.2a$ and period a . A defect center-to-center separation of $2.2a$ produces a pair of doubly-degenerate, p-like modes with the desired symmetry at a frequency of $0.22 c/a$ (*i.e.* a wavelength in the dielectric of $1.3a$).

As before, we use a two-dimensional simulation to evaluate the throughput and crosstalk of these two structures, this time with TE-polarized light. The results are shown in Fig. 3.5. Not only does the resonant cavity increase the throughput from 75% to over 90%, but the crosstalk is decreased by two orders of magnitude, from 7% to 0.08%. The resonance Q is close to 40 and, as before, the crosstalk can be estimated by Q^{-2} . The main

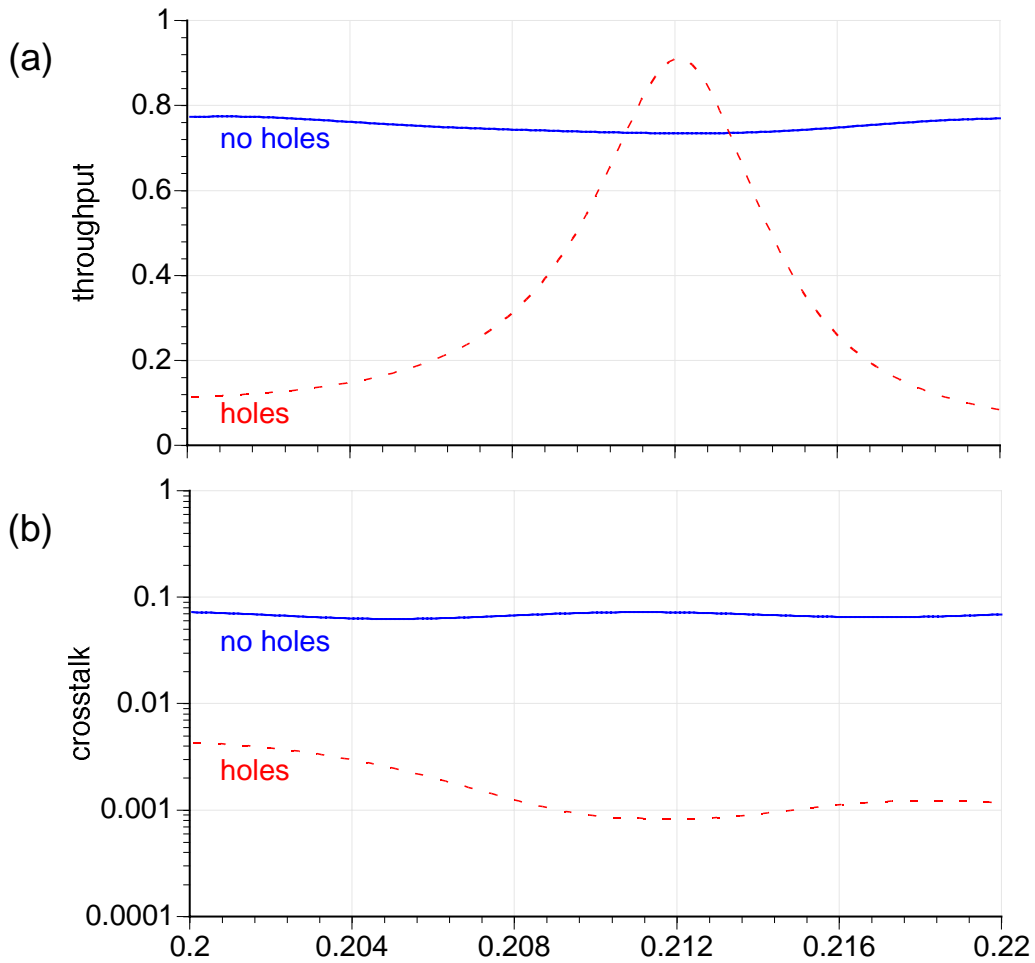


Figure 3.5: TE transmission spectra for the two intersections shown in Fig. 3.4.

difference between the behavior of this crossing and the one from Fig. 3.2 is that there are now radiation losses from scattering off the intersection, resulting in lower throughput.

3.5 Summary

We have shown how, by the application of general principles and symmetry considerations, it is possible to systematically construct efficient waveguide intersections in diverse structures. Although our examples are limited to two-dimensional systems, the same method is applicable in three dimensions. Photonic crystal cavities have been used because of the flexibility and tunability they provide, but the theory is independent of the confinement mechanism. Furthermore, several additional refinements are possible. Many

degrees of freedom remain in high index-contrast crossings (as in Fig. 3.4), such as the shape of the intersection and the size and number of holes, that could be used to further reduce radiation losses and increase throughput. Another important possibility is to broaden and flatten the resonance peak, which could be accomplished by adding additional resonators in the waveguides adjacent to the intersection [18].

Chapter 4

Guided Modes in Photonic-Crystal Slabs*

4.1 Introduction

The discovery of photonic crystals, periodic dielectric materials with a photonic band gap, has opened up new methods for controlling light, leading to proposals for many novel devices [9]. Straightforward application of these results to three dimensions requires a structure with a three-dimensional band gap. Fabricating such structures, however, has been a challenge because they tend to have complex three-dimensional connectivity and strict alignment requirements [19]–[22]. Such designs have been the subject of many recent developments [23]–[27]. An alternative system, the photonic-crystal slab, has been proposed that promises easier fabrication using existing techniques [28]–[35]. This is a dielectric structure that has only two-dimensional periodicity and uses index guiding to confine light in the third dimension. Photonic-crystal slabs retain or approximate many of the desirable properties of true photonic crystals, but at the same time are much more easily realized at submicron lengthscales. We present in this chapter a novel band-structure analysis of photonic-crystal slabs, providing a systematic understanding of many important features of these systems.

* This chapter is based on: S. G. Johnson, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and L. A. Kolodziejwski, “Guided modes in photonic crystal slabs,” *Phys. Rev. B* **60**, 5751–5758 (1999).

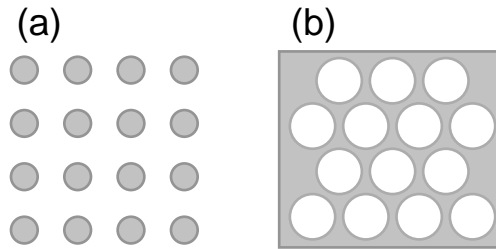


Figure 4.1: Two-dimensional photonic crystals. (a) Square lattice of dielectric rods in air, with lattice constant a and radius $0.2a$. (b) Triangular lattice of air holes in dielectric, with lattice constant a and radius $0.45a$. In both cases, the dielectric constant of the high-index material is 12.

In many ways, photonic-crystal slabs are analogous to two-dimensional photonic crystals, such as those depicted in Fig. 4.1, and this analogy aids greatly in the visualization and analysis of slab systems. Two-dimensional calculations, however, cannot be applied directly to three-dimensional slab structures. In particular, the band structure computed for a two-dimensional structure, as shown in Fig. 4.2, applies in three dimensions only to a structure that is infinitely “extruded” in the third dimension. Moreover, these two-dimensional bands correspond only to states that have no wavevector component in the vertical direction (perpendicular to the plane of periodicity). The inclusion of vertical wavevectors produces a continuum of states depicted by the shaded region in Fig. 4.2, destroying the band gap of the two-dimensional structure. The restriction of a slab to finite height recreates the band gap in the guided modes of a slab, but also forces a new analysis of the system that is fundamentally three-dimensional and distinct from the two-dimensional calculations. New issues such as slab thickness, index contrast with the substrate, and mirror symmetry assume a prominent role in determining the properties of photonic-crystal slabs.

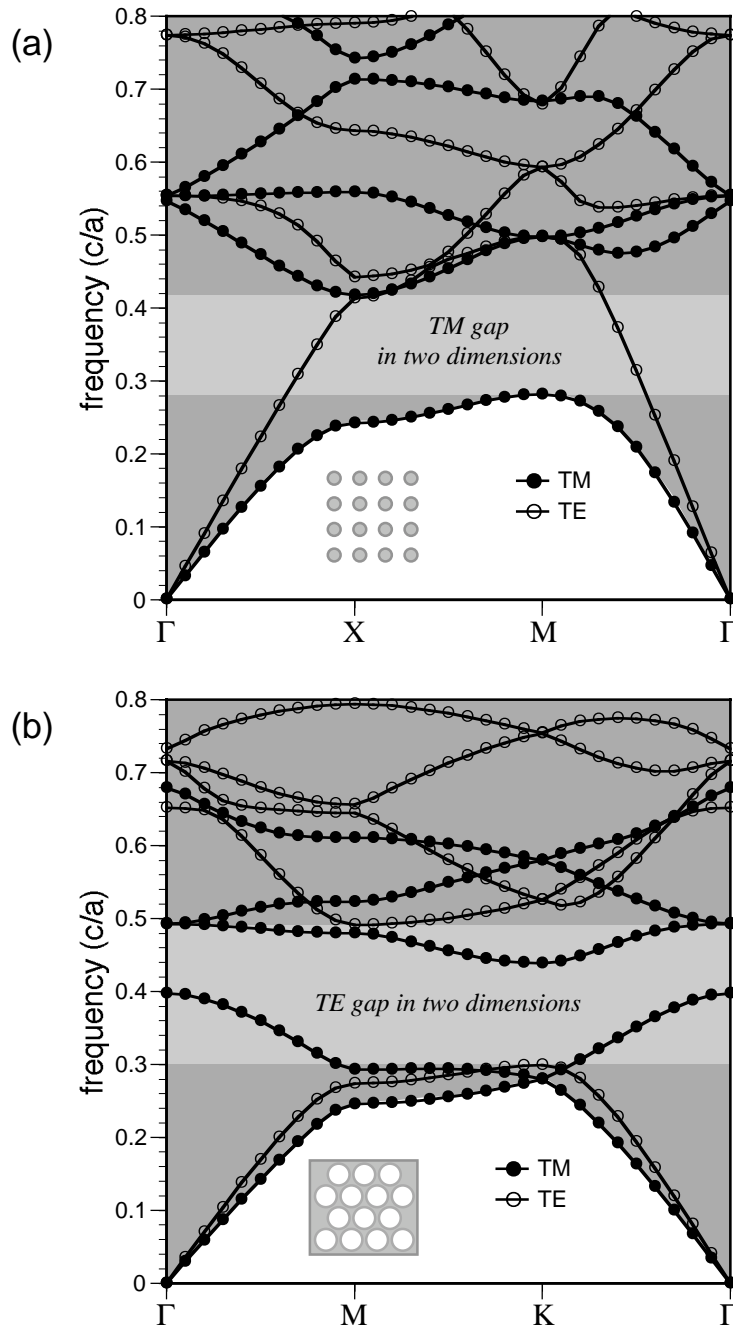


Figure 4.2: Band diagrams for the photonic crystals from (a) Fig. 4.1(a) and (b) Fig. 4.1(b). The shaded region indicates the frequencies of states introduced when vertical propagation (*i.e.*, perpendicular to the plane of periodicity) is permitted.

The chapter is structured as follows. We first outline the numerical methods that were used in our calculations. We then introduce slab band structures for two characteristic systems, the slab analogues of Fig. 4.1, suspended in air. The selection of an appropriate slab thickness is discussed, and an estimate for the optimal slab thickness is derived using a variational approach. Finally, the effects of different backgrounds, the regions above and below the slab, are examined in the context of the two example structures. Band structures are given for solid, periodic, and symmetry-breaking backgrounds.

4.2 Computational Method

The computation of a slab band structure is performed in two stages. First, the states of a slab are computed in a periodic cell. Then, the light cone is calculated and overlaid as an opaque region on the band diagram. The resulting non-obscured bands are the guided modes of the system, as described in the following section.

The eigenstates of the slab are computed using preconditioned conjugate-gradient minimization of the Rayleigh quotient in a plane-wave basis [36]. Such a computation requires a periodic cell. The slab is already two-dimensionally periodic, and we impose a three-dimensional periodicity by assuming a periodic sequence of slabs separated by a sufficient amount of background region. Because the guided modes are localized within the slab, the addition of periodic slabs at large intervals does not affect their frequencies noticeably. The non-guided modes are affected, but since they fall inside the light cone their exact frequencies are inconsequential.

The light cone is a continuous region indicating all possible frequencies of the bulk background. It is sufficient to compute only the lower boundary of the light cone, since all higher frequencies are automatically included. In the case of a uniform background region, the boundary is simply the wavevector divided by the index. For a periodic back-

ground, the boundary is the lowest band of the corresponding two-dimensional system, as is discussed in a later section, and is computed using the two-dimensional version of the method in [36]. In all cases, the light cone is depicted with a uniform shading that does not reflect the varying density of states in this region.

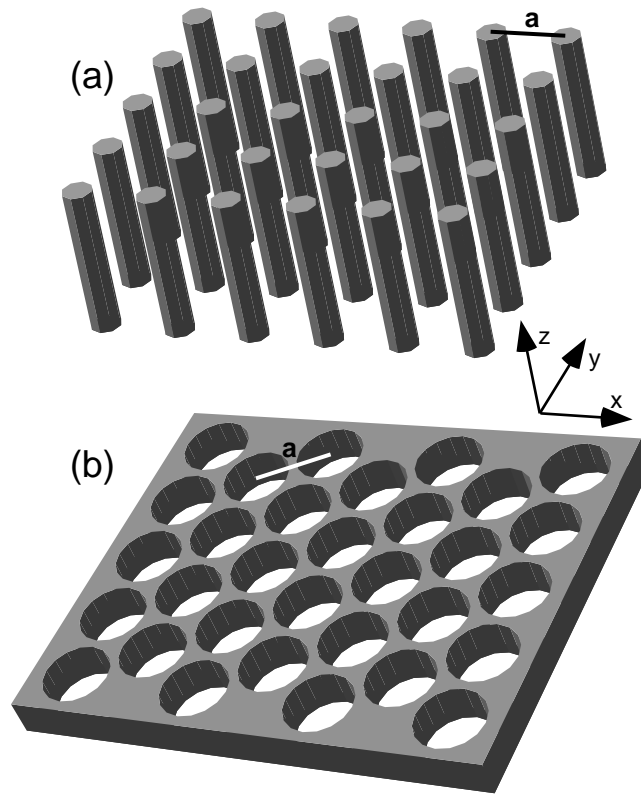


Figure 4.3: Photonic-crystal slabs analogous to the two-dimensional structures from Fig. 4.1. (a) Square lattice of rods in air with height $2.0a$. (b) Triangular lattice of holes in a dielectric slab with thickness $0.6a$. (Other parameters are as in Fig. 4.1.) Both slabs are suspended in air.

4.3 Photonic-Crystal Slab Band Structures

Shown in Fig. 4.3 are the photonic-crystal slab analogues of the two-dimensional structures from Fig. 4.1. These two systems, a square lattice of dielectric rods in air and a triangular lattice of air holes in a solid dielectric slab, embody the two basic topologies of two-dimensional crystals, and are characteristic of many possible slab structures. (An air-

bridge structure similar to Fig. 4.3(b) was fabricated in [35].) As seen from Fig. 4.2, in two dimensions the rod structure has a band gap in the TM modes (magnetic field in plane) and the hole structure has a band gap in the TE modes (electric field in plane). (The hole structure also has a small band gap in the TM modes, but we focus here on the more robust TE gap, which is larger and separates the lowest two TE bands.)

The corresponding projected band structures for the three-dimensional slab structures are shown in Fig. 4.4. These graphs, whose computation was discussed in the previous section, illustrate many features that are common to all photonic crystal band diagrams.

Perhaps the most important feature of the projected band diagram, the element that distinguishes slabs from ordinary photonic crystals, is the light cone, a continuum of states indicated by a shaded region in the plot. The light cone consists of states, or radiation modes, that are extended infinitely in the region outside the slab; we refer to this region as the “background.” Guided modes, which are states localized to the plane of the slab, can only exist in the regions of the band diagram that are outside the light cone. The primary interest in the radiation modes lies in how they interact with and constrain these guided modes.

Any state that lies below the light cone in the band diagram cannot couple with modes in the bulk background. Thus, the discrete bands below the light cone are guided—the states are infinitely extended within the plane of the slab, but decay exponentially into the background region. This confinement is analogous to total internal reflection, and is due to the guided modes seeing a higher effective index in the slab than in the background regions (air, in this case). When a guided band reaches the edge of the light cone, it becomes a resonant state: it extends, albeit with low amplitude, infinitely far into the background, and cannot be used to permanently confine light within the slab. We restrict our

discussion, and our use of the term “guided modes,” to truly localized states, which grow arbitrarily small as the distance from the slab becomes large

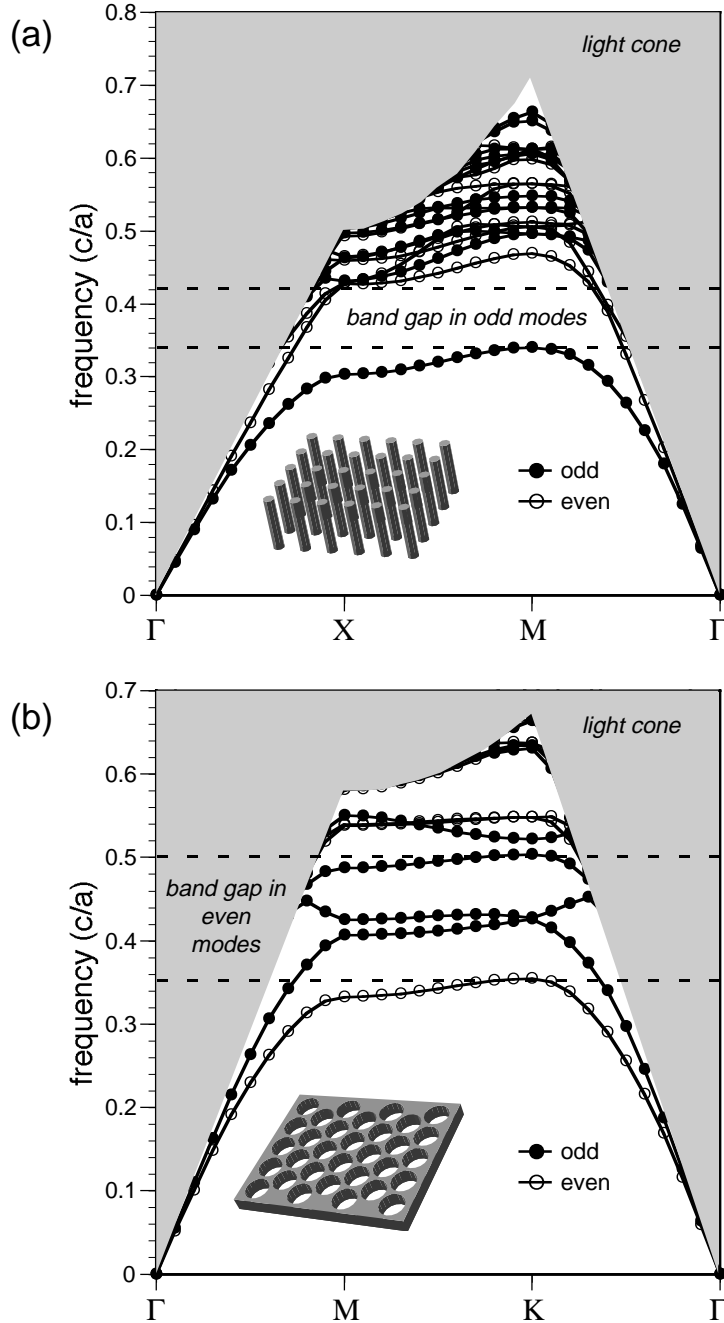


Figure 4.4: Projected band diagrams corresponding to the two slabs in Fig. 4.3. Whether states are even or odd with respect to the horizontal mirror plane of the slab is indicated by open or filled circles, respectively.

The reason we refer to these systems as “photonic crystal” slabs is that, like their two-dimensional brethren, they have a band gap—but not of the traditional sort. A “band gap” in this case is a range of frequencies in which no *guided* modes exist. It is not a true band gap because there are still radiation modes at those frequencies. Still, the lack of guided modes in the band gap gives rise to many of the same phenomena that occur in two-dimensional crystals, such as the ability to confine light in the plane to waveguides or resonators [32]. (The presence of the radiation modes in the gap has the consequence that resonant cavity modes will eventually decay into the background.) The fact that light in the band gap of the slab is forbidden from propagating in the plane of the slab, and can only radiate into the background, was used in [30] to design an efficient LED.

As in two dimensions, one is able to decompose the guided modes into two non-interacting classes. The lack of translational symmetry in the vertical direction, however, means that the states are not purely TM or TE polarized. Instead, due to the presence of a horizontal symmetry plane bisecting the slab, the guided bands can be classified according to whether they are even or odd with respect to reflections through this plane, and are indicated on the band diagram by hollow or filled circles, respectively. As shall be seen below, these even and odd states have strong similarities with TE and TM states, respectively, in two dimensions. (In fact, in the mirror plane itself, the even and odd states are purely TE, and TM, respectively.) It is not surprising, then, that the slab of rods has a gap in its odd modes (Fig. 4.4(a)), and the slab of holes has a gap in its even modes (Fig. 4.4(b)).

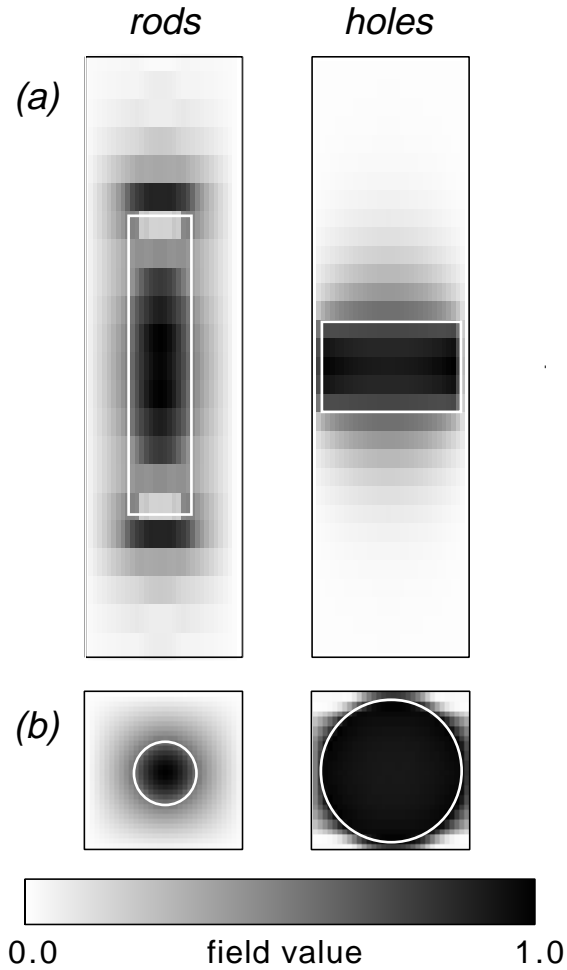


Figure 4.5: Lowest-frequency states of the rod and hole structures from Fig. 4.3 at M and K , respectively, depicted in the unit cell. The shading indicates the z components of the electric and magnetic fields, respectively, which happen to be positive throughout the unit cell for these states. Outlines of the rods/holes are shown in white. (a) Vertical cross-section. (b) Horizontal cross-section. In both cases the cross-sections bisect the holes or rods.

In Fig. 4.5, z -components of the electric and magnetic fields are shown for the lowest-order odd and even guided modes from the rod and hole slabs at the M and K symmetry points, respectively. It is apparent from these figures that these states are strongly guided within the thickness of the slab. Moreover, within the slab they are TM- and TE-like, and closely resemble the corresponding states in the two-dimensional system.

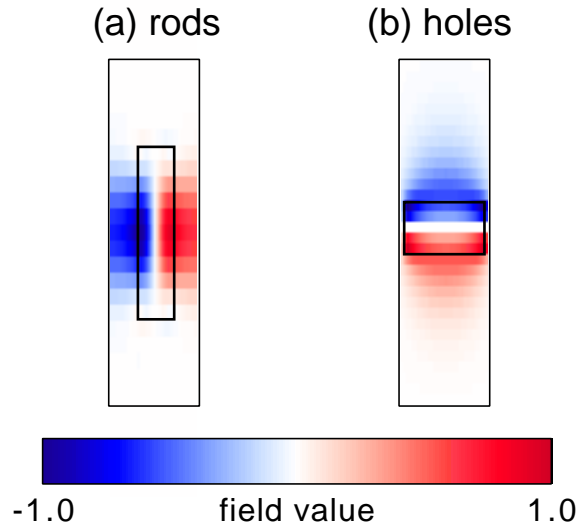


Figure 4.6: (a) Horizontal cross-section of E_z for the first odd (TM-like) excited state at M in the rod structure from Fig. 4.4(a). (b) Vertical cross-section of the H_z for the first TE-like (but odd) excited state at K in the hole structure shown in Fig. 4.4(b).

Higher-order states in the slabs are formed in two ways. First, they can gain additional nodes or other structure in the horizontal plane, corresponding to higher-order states in the two-dimensional system. Second, they can be formed by adding vertical structure such as horizontal nodal planes. In the case of the rods, the lowest higher-order odd state is of the first type, and is depicted in Fig. 4.6(a). The second odd mode at the M point in the holes slab, as shown in Fig. 4.6(b), is of the second type—it corresponds to the lowest-order even mode modified by the addition of a single horizontal nodal plane. Note that adding a nodal plane bisecting the slab transforms a state from odd to even and vice-versa, so even and odd higher-order modes aren't necessarily TE-like and TM-like, respectively.

Higher-order modes of the second type do not correspond to excitations in the two-dimensional system, and are responsible for destroying the gap if the slab becomes too thick, as is discussed in the following section.

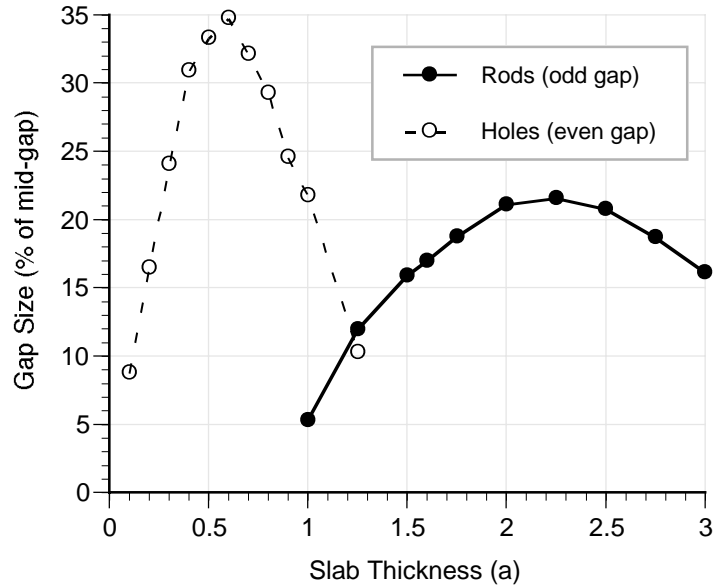


Figure 4.7: Gap size (as a percentage of mid-gap frequency) versus slab thickness for the two slabs from Fig. 4.3.

4.4 Effects of Slab Thickness

The slab thickness plays an important role in determining whether a photonic-crystal slab has a band gap in its guided modes. Shown in Fig. 4.7 is a graph of the band gap size as a function of slab thickness for the rod and hole slabs from Fig. 4.3. The existence of an optimal thickness for each slab is easily understood. If the slab is too thick, then higher-order modes can be created with little energy cost by adding horizontal nodal planes. Such modes will lie only slightly above the lowest-order mode, preventing any gap. If the slab is too thin, then the slab will provide only a weak perturbation on the background dielectric constant. Guided modes will still exist, but they will hug the edge of the light cone and be only weakly guided; any gap will be miniscule. Below, we propose a method for estimating slab thicknesses that produces large gaps, and explain the sharp distinction seen in Fig. 4.7 between the optimal thicknesses for the rod and hole slabs.

We postulate that the optimal thickness will be on the order of half the two-dimensional gap-bottom wavelength. The justification for this is based on the fact that the wave-

length is the lengthscale for variations in field amplitude at a given frequency. (We use the gap-bottom frequency instead of, say, mid-gap, because the state at the bottom of the gap is the basis for both the state at the lower edge of the slab gap and the excited states at upper edge.) If the slab thickness is a wavelength or more, then there will be little energy barrier to creating a higher-order state via a nodal plane. If the slab thickness is less than half a wavelength, on the other hand, then the mode cannot be strongly confined within the slab.

Only the frequency, rather than the wavelength, of the gap bottom is known, however. An effective dielectric constant must be determined in order to compute the corresponding wavelength. We accomplish this by constructing an estimate for the slab gap-bottom state using the two-dimensional wavefunction, and evaluating the dependence of its frequency expectation value on the vertical wavelength. The wavefunction can be approximated by:

$$|\tilde{H}\rangle = \int \vec{H}_0(x, y) e^{ikz} a(k) dk. \quad (4.1)$$

Here, \vec{H}_0 is the eigenfunction of the two-dimensional system at the lower edge of the gap. The $a(k)$ amplitudes are assumed to be chosen so that they produce a state strongly localized within the slab. A Fourier basis is used for the z -dependence so that the vertical wavevector may be an explicit parameter. We have chosen this wavefunction based on our experience (*e.g.* in Fig. 4.5) that the lowest-order guided band in the slab system is similar in appearance to the corresponding two-dimensional state with the addition of vertical confinement. (We note that the field in (1) is not divergenceless, however.)

We now evaluate the energy (*i.e.*, frequency) of this state. Since the state is assumed to be localized within the slab, the frequency should not be much affected if we evaluate it in an infinitely thick slab, which can be done exactly. (In the case shown in Fig. 4.5, more than 90% of the computed wavefunction's energy was found to lie within the height of the slab.) In this approximation, the frequency is found to be one of:

$$\omega_{TM}^2 = \frac{\langle \tilde{H} | \vec{\nabla} \times \frac{1}{\epsilon(x, y)} \vec{\nabla} \times | \tilde{H} \rangle}{\langle \tilde{H} | \tilde{H} \rangle} = \omega_0^2 + \frac{c^2 \langle k^2 \rangle}{\bar{\epsilon}_{TM}} \quad (4.2)$$

$$\omega_{TE}^2 = \frac{\langle \tilde{E} | \vec{\nabla} \times \vec{\nabla} \times | \tilde{E} \rangle}{\langle \tilde{E} | \epsilon(x, y) | \tilde{E} \rangle} = \omega_0^2 + \frac{c^2 \langle k^2 \rangle}{\bar{\epsilon}_{TE}}, \quad (4.3)$$

where E represents the electric field, and ω_0 is the two-dimensional frequency. Equations (2) and (3) are equivalent only for divergenceless (transverse) fields. In the case of (1), which is not transverse, we use (2) for TM fields and (3) for TE fields, for simplicity of the resulting expression. In both cases, the vertical wavelength is related to frequency simply by an effective dielectric constant that depends upon the polarization of the corresponding two-dimensional state, and is given by:

$$\bar{\epsilon}_{TE} = \frac{\langle E_0 | \epsilon | E_0 \rangle}{\langle E_0 | E_0 \rangle} \equiv \langle \epsilon \rangle_E \sim \epsilon_{high} \quad (4.4)$$

$$\bar{\epsilon}_{TM}^{-1} = \frac{\langle H_0 | \epsilon^{-1} | H_0 \rangle}{\langle H_0 | H_0 \rangle} \equiv \langle \epsilon^{-1} \rangle_H \sim \epsilon_{low}^{-1}. \quad (4.5)$$

In the TE case, the dielectric constant is weighted by the electric field, which is concentrated in high dielectric for the lower bands [9], and thus one should get an $\bar{\epsilon}$ near to the high dielectric constant. For the example of air holes in dielectric (Fig. 4.1(b)), the top-most state of the bottom TE band in the two-dimensional system gives an $\bar{\epsilon}$ of 5.06 from Eq. (4.4). This value, while greater than the uniform-weight mean ϵ of 3.92, is still far from 12 (the high dielectric), as a result of the structure's dielectric veins being so thin. (This structure has both a TE and a TM gap, but we focus on the TE gap corresponding to the even modes.) In Eq. (4.5), the TM case, we average the inverse dielectric constant, a mean that favors lower values. Moreover, the mean is weighted by the magnetic field, which tends to be less in the high dielectric than the electric field. (The magnetic field must loop around the electric field in the high dielectric, since the relationship of the magnetic and electric fields is the same as that of a current loop and its magnetic field.) There-

fore, the TM average dielectric should be closer to the low dielectric constant. In the case of the dielectric rods in air (Fig. 4.1(a)), Eq. (4.5) for the highest-frequency mode of the bottom TM band in two-dimensions produces an $\bar{\epsilon}$ of 1.25 (versus the uniform-weight inverse mean of 1.13).

From the above calculations, then, an estimate for the optimal slab thickness is given by:

$$\text{slab thickness } h \sim \frac{1}{2\omega_{\text{gap-bottom}}\sqrt{\bar{\epsilon}}}. \quad (4.6)$$

Here, the frequency is given in units of c/a and the thickness is in units of a . The effective dielectric constant $\bar{\epsilon}$ can be estimated from the averages in (4) or (5), according to whether a TE or a TM gap is of interest, and can use as a weighting factor either the state from a two-dimensional calculation or simply a uniform value.

When applied to the rod and hole systems with the effective dielectric constants computed above, Eq. (4.6) predicts optimal thicknesses of $1.6a$ and $0.7a$, respectively, compared to the computed gap maxima at $2.3a$ and $0.6a$ from Fig. 4.7. The estimates come close enough to the optimal values that large gaps are produced (whereas there would be no gap at all if the two estimates were applied to each others' systems). Moreover, this approximation explains how the sharp difference in optimal thickness between the rod and the hole systems derives from the polarizations of the modes exhibiting the gap.

It should be noted that the size of the gap is not necessarily the only consideration in selecting the slab thickness. For example, when localizing states in a resonant cavity, longer decay times can sometimes be achieved by using slightly thicker slabs—this causes the states to be more localized within the slab, and also pulls the frequencies down to where the density of radiation modes is lower (since the density of states in vacuum goes as ω^2). It is often better to tune parameters, such as the slab thickness, based upon the

actual phenomenon that is being optimized, rather than indirectly via the size of the band gap. Nevertheless, the approximation in (6) gives a reasonable starting point for subsequent fine-tuning.

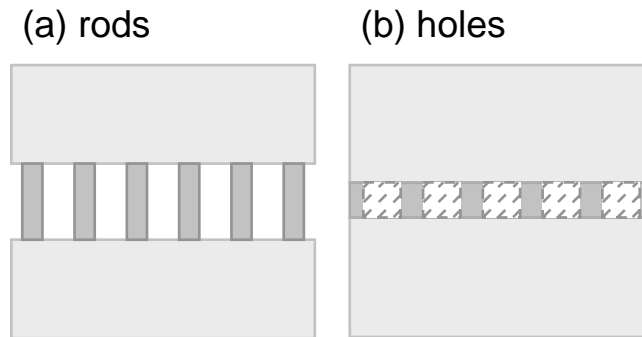


Figure 4.8: Side-view of the slabs from Fig. 4.3 with a uniform, low-index background above and below the slabs. (a) is the rod slab and (b) is the hole slab. The background has a dielectric constant of 2.0 and all other parameters are as in Fig. 4.3.

4.5 Slabs with Solid Backgrounds (Sandwiches)

So far, we have focused on the idealized system of slabs suspended in air. We now turn to the case where the background regions above and below the slab are occupied by a uniform dielectric material forming a substrate or “sandwich” as shown in Fig. 4.8. (We focus first on having the same dielectric above and below the slab in order to maintain mirror symmetry; the effects of symmetry breaking are considered in a later section.) The dielectric constant of the substrate is 2, but the holes and the space between the rods continue to be occupied by air. The resulting band structures are given in Fig. 4.9. The background index remains lower than the effective index in the slab, and hence index guiding is still able to produce the guided bands (and band gap) evident in Fig. 4.9.

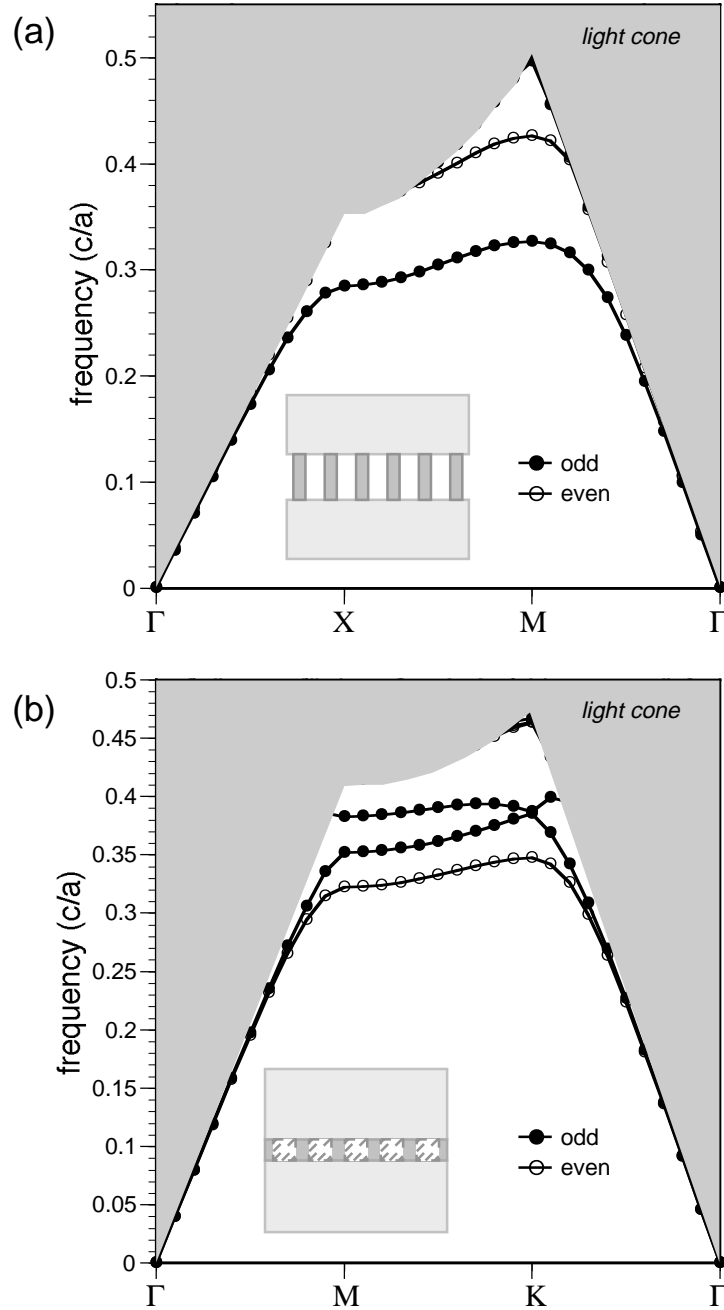


Figure 4.9: Projected band diagrams for the structures shown in Fig. 4.8.

While such a substrate will have a finite thickness in reality, the localization of the guided modes within the slab means that the substrate can be considered infinite as long as it is sufficiently thick. (In this case, a substrate thickness of several wavelengths is suffi-

cient for the guided mode amplitude to be negligible beyond the substrate.) Thus, the light cone states are those of an infinite uniform dielectric, whose frequencies are reduced from those in vacuum by a factor of the index. The increased index above and below the slab also has the effect of pulling down the frequencies of the guided modes, allowing them to remain under the now-lowered light cone. In addition, the guided modes are somewhat less localized—for example, 89% of the energy of the lowest band for the hole structure at K is within the height of the slab for the solid background, versus 96% for an air background.

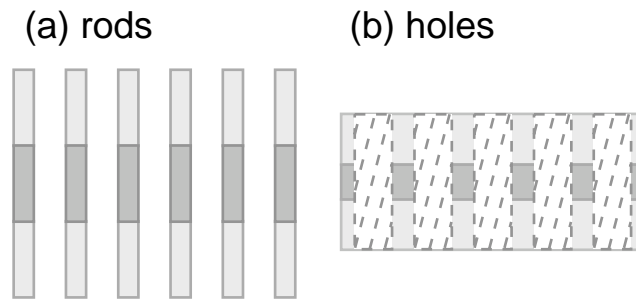


Figure 4.10: Side-view of the slabs shown in Fig. 4.3 with a periodic background formed by “extruding” the structures with a low-index material (dielectric constant 2.0). All other parameters are as in Fig. 4.3.

4.6 Slabs with Periodic Backgrounds

The effective index above and below the slab can be reduced from that of a solid background by using an “extruded,” low-index version of the slab, as shown in Fig. 4.10. For example, in the case of the hole slab, the holes extend through the low-dielectric substrate as well as through the slab. The resulting band structures are shown in Fig. 4.11, and again demonstrate guided modes and a band gap. Such a structure has advantages both in ease of fabrication (both slab and substrate can be etched at the same time) and in confinement of resonant cavity states (since localized states couple less strongly with a lower-index background). The periodicity of the background does not, however, produce useful photonic-

crystal slab effects by itself. Even if there were some sort of band gap in the background states, it would lie at higher frequencies than the band gap of the guided modes and would therefore not provide any additional confinement capabilities.

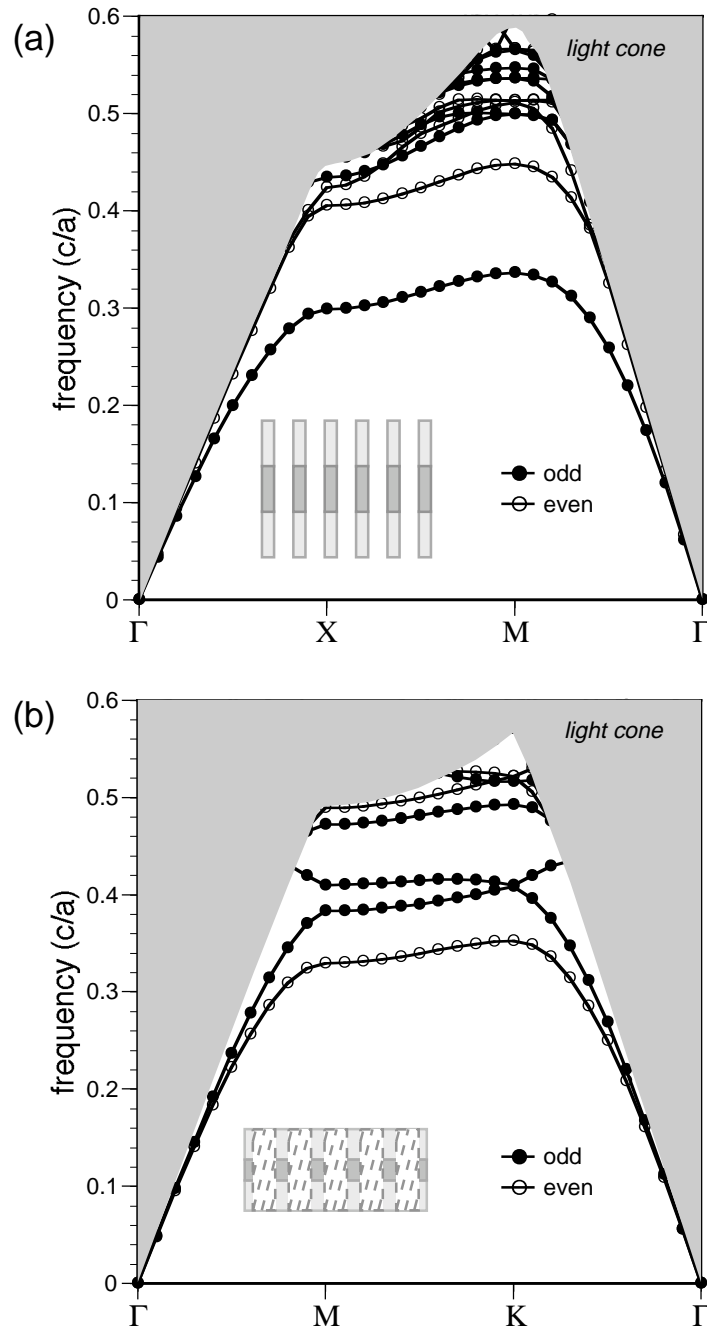


Figure 4.11: Projected band diagrams for the structures shown in Fig. 4.10.

The light cone in this system consists of all the states existing in the bulk background, an infinitely-extruded two-dimensional photonic crystal. This is similar to the shaded regions from Fig. 4.2, except that the material has a smaller index. That is, the lower edge of the light cone is simply the lowest band of the two-dimensional system. The fact that the lowest band of the two-dimensional structure forms a lower bound for the frequencies of the extruded structure in three dimensions may not be immediately apparent, and so we consider it below.

Because the dielectric function of the bulk background, $\varepsilon(x, y)$, has translational symmetry in the z direction, any background state's magnetic field can be written in the form:

$$|H\rangle = \vec{H}(x, y)e^{i(kz - \omega t)}. \quad (4.7)$$

Such a field is “manifestly” propagating in the positive- z direction (upwards) for positive k and downwards for negative k . Therefore, the group velocity $\frac{d\omega}{dk}$, which can be expressed as the upwards flux divided by the energy density (using the derivative of the eigenvalue equation), will ordinarily have the same sign as k . It then follows that the minimum frequency occurs at zero k , at which point the solution $|H\rangle$ is simply the two-dimensional eigenstate. This argument holds at each point in the Brillouin zone, and so the light cone is bounded by the lowest two-dimensional band.

Although the above is true for the present system, there are actually rare cases in which the group velocity can be opposite to the phase velocity in a structure with translational symmetry! These “backward waves” have been found in coaxial cables (dielectric or metallic) and also in anisotropic materials [37]–[38]. When this occurs, the minimum-frequency states are at a (small) non-zero k —this is the frequency that determines the light line. To compute the exact light line in general, then, one must solve for the lowest “2d” band at small positive k and check to make sure that it is not at a lower frequency.

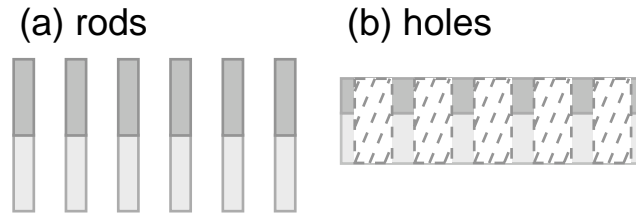


Figure 4.12: Side-view of the slabs shown in Fig. 4.10 (extruded backgrounds) with the upper background replaced by air. All other parameters are as in Fig. 4.10. This background breaks the mirror symmetry of the two photonic-crystal slabs.

4.7 Slabs with Symmetry-Breaking Backgrounds

It is also possible to have a background that is not symmetric, for example a substrate below the slab and air above the slab. (Rod and hole slabs with solid substrates have been studied experimentally, *e.g.*, in [39]–[41].) The light cone in this case is the union of the light cones for the backgrounds above and below the slab. Such a structure is depicted in Fig. 4.12, in which the periodic backgrounds from the previous section here lie only below the slabs. The resulting band diagram is given in Fig. 4.13.

The most important consequence of a symmetry-breaking background is that the guided modes can no longer be classified as even or odd. Thus, there is no longer any band gap in the guided modes, and the photonic crystal properties of the slab are ostensibly lost. If the guided modes are sufficiently localized within the slab, however, so that the background is only a small perturbation, the wavefunctions may still be approximated as even or odd and some effects of the band gap will persist.

In order to maintain the distinction between even and odd guided modes, it is only necessary to preserve mirror symmetry where the guided modes have non-negligible amplitude. Thus, a solid substrate can be used below the slab (and not above) with no effects on the band gap as long as the substrate is separated from the slab by a buffer region that is sufficiently thick. Shown in Fig. 4.14 are examples of how a solid substrate might be used

below the structures from the previous two sections without affecting the band structures significantly.

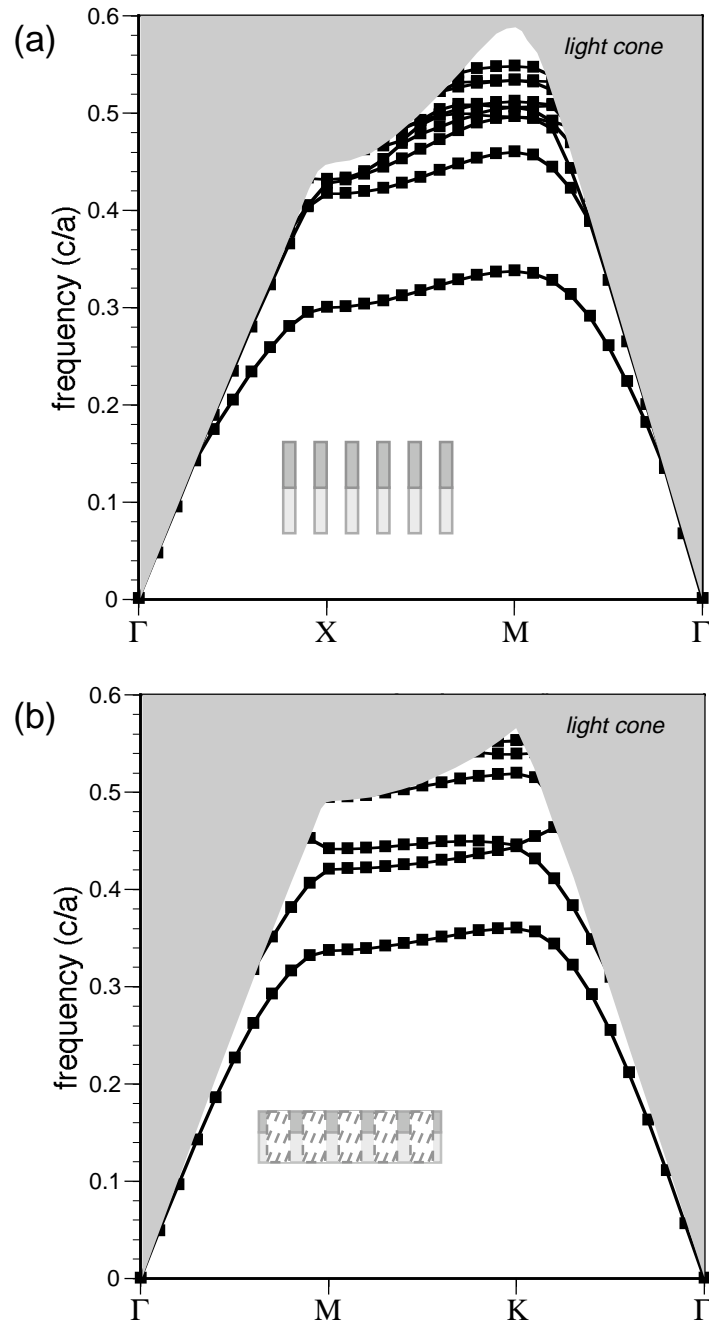


Figure 4.13: Projected band diagrams for the structures shown in Fig. 4.12. The bands can no longer be segregated into even and odd modes, and there is no longer a band gap in the guided modes.

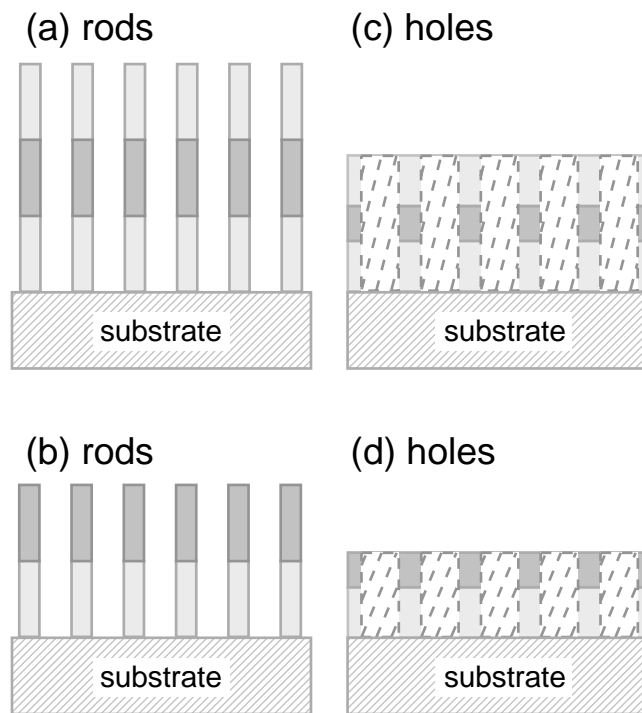


Figure 4.14: A solid substrate may lie below a slab without affecting the band structure by being sufficiently far from the slab. (a), (b), (c), and (d) depict such solid substrates below the symmetric and antisymmetric rod and hole structures from Figs. 10(a), 12(a), 10(b), and 12(b), respectively.

4.8 Summary

Photonic-crystal slabs are an important system in the practical application of photonic crystals, and the band structure formalism provides a powerful tool in their analysis. Unlike two- and three-dimensional photonic crystals, the band structures for slabs are projected and have the unique feature of a light cone enveloping the states. The presence of the light cone means that a complete band gap is impossible—still, waveguides and resonant cavities are possible [32]. Furthermore, the presence of a band gap in the guided modes is useful in its own right. For example, Fan *et al.* have proposed using slab band gaps to prevent in-plane losses from LED emission [30]. Perhaps the greatest promise of photonic-crystal slabs, however, is that they will allow exciting results from two-dimensional crystals such as waveguide bends [42] and channel-drop filters [43] to be imple-

mented easily on optical and infrared lengthscales.

Chapter 5

Linear Waveguides in Photonic-Crystal Slabs*

5.1 Introduction

Photonic crystals, which prohibit the propagation of light for frequencies within a band gap, have enabled exciting new ways to control light and construct integrated optical devices [44]. An important element of optical circuits is a linear waveguide to carry light to and from components, and photonic crystals also provide unique advantages for waveguides. Photonic-crystal waveguides, guided by the band gap of the bulk crystal [44]–[47], can exhibit near-zero reflection and loss through sharp bends [48] and when coupled with resonant cavities [49]–[50], due to the gap’s prevention of radiation losses. Previous theoretical studies of photonic-crystal waveguides, however, have been restricted to purely two-dimensional systems and have not examined the effects of vertical confinement. In this chapter, we present a systematic study of waveguide modes in a three-dimensional system, photonic-crystal slabs. Photonic-crystal slabs are two-dimensionally periodic dielectric structures of finite height that have a band gap for propagation in the plane and use index-confinement in the third dimension; they have been proposed as a more-easily fabricated alternative to true three-dimensionally periodic photonic crystals [45],[51]–[62]. Although their structure and properties strongly resemble those of two-dimensional crystals, slab systems require a fundamentally different, three-dimensional analysis [30],[59],[62]. We will demonstrate how such analyses apply to waveguides and explain the novel considerations that arise for line defects in photonic-crystal slabs. We shall show that, merely to produce guided modes (as well as to be single-mode and in-gap), the parameters of the defect must be carefully chosen.

* This chapter is based on: S. G. Johnson, P. R. Villeneuve, S. Fan, and J. D. Joannopoulos, “Linear waveguides in photonic-crystal slabs,” *Phys. Rev. B* **62**, 8212–8222 (2000).

Waveguides must satisfy three criteria in order to achieve optimal performance in many integrated-optical-circuit applications; *e.g.*, for maximum transmission through sharp bends (on the scale of the wavelength) and resonant cavities. First, of course, the waveguide must support true guided modes (as opposed to resonances, as defined later). This also implies that the waveguide structure must be periodic along the direction of propagation, in order to have a well-defined Bloch wavenumber k and to thereby propagate without reflections. Second, the waveguide should be single-mode in the frequency range of interest—this condition is necessary in [48] and the increased reflections in a multi-mode waveguide are a straightforward consequence of coupled-mode theory [63]. (Extraneous modes may have little effect if coupling to them through the bend or cavity can be made negligible, but this is not generally the case for wavelength-scale structures.) Third, the guided mode should lie within the band gap of a photonic crystal in order to prohibit radiation losses (which both decrease transmission and increase reflection [63]); otherwise, some losses are inevitable at a bend or resonator, which breaks translational symmetry.

Conventional linear dielectric waveguides operate by index-confinement (total internal reflection, in the short-wavelength limit). Such waveguides can be made to satisfy the first two criteria from above (*i.e.*, be guiding and single-mode), and thereby have relatively high transmission through sharp bends and resonators [50],[64],[65]. However, the absence of a photonic band gap means that transmission is always limited by radiation losses. In a purely two-dimensional photonic-crystal linear waveguide, a linear (one-dimensionally periodic) defect is introduced into the crystal, creating a localized band that falls within and is guided by the photonic band gap [44]–[47]. Light is therefore prohibited from escaping the waveguide, and all three criteria can be satisfied to achieve perfect transmission through bends and resonant cavities [48]–[50]. These results have been

experimentally confirmed using structures that are uniform in the third dimension for many wavelengths, and thus approximate a two-dimensional system [66]; the large aspect ratios of such geometries and their lack of vertical confinement or a true band gap in the guided modes limit their practical applications, however. There are several other possibilities for realizing photonic-crystal waveguides in three dimensions. The most ideal structure, and the only way to totally eliminate radiation losses, would be a three-dimensionally periodic structure with a complete band gap, in which the photonic crystal completely surrounds the waveguide. Such structures, though, are challenging to construct at submicron lengthscales.[67]–[70] In this chapter, we instead focus on photonic-crystal slabs. Waveguides in photonic-crystal slabs are analogous to those in two-dimensional crystals, except that they use index confinement in the third dimension. Similar waveguides, which use photonic-crystal confinement in two dimensions and index confinement in the third, have been subject to experimental studies [71]–[72], as well as theoretical work [73] subsequent to the submission of the article upon which this chapter is based. There is also another type of photonic-crystal waveguide that we do not consider in this chapter, the

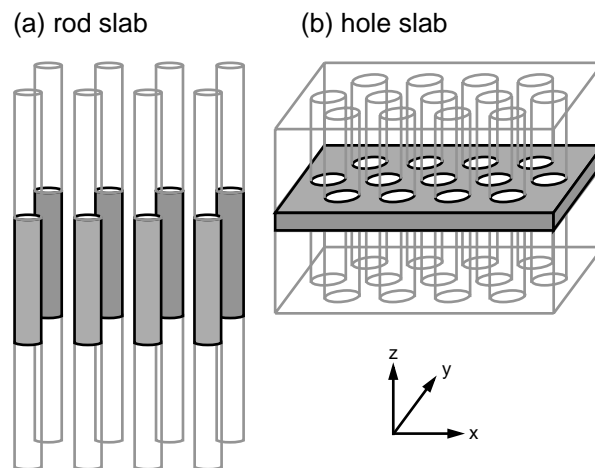


Figure 5.1: Photonic-crystal slabs. (a) Square lattice of dielectric rods in air with ϵ of 12.0, lattice constant a , radius $0.2a$, and height $2.0a$, with low-index (ϵ of 2.0) rods extending infinitely above and below. (b) Triangular lattice of air holes extending through both a high-index (ϵ of 12.0) finite-height dielectric slab and low-index (ϵ of 2.0) semi-infinite dielectric regions above and below. The holes have lattice constant a and radius $0.3a$, while the high-index slab is of thickness $0.5a$.

photonic-crystal fibre [74]–[75]. This is a two-dimensionally periodic system of (in principle) infinite thickness, in which the waveguide mode propagates perpendicular to the plane of periodicity; in contrast, photonic-crystal slab waveguides are of finite thickness and run parallel to the plane of periodicity. A waveguide operating on principles similar to the fibres’ was also considered in [76].

In this chapter, we analyze linear-defect waveguides in two characteristic photonic-crystal slabs, a square lattice of dielectric rods in air and a triangular lattice of air holes in dielectric, illustrated in Fig. 5.1. Above and below the slabs are semi-infinite “extruded” substrates with the same cross-sections as the slabs, but having a lower dielectric constant. (We use a substrate in order to model a more practical system than a suspended slab, and having the substrate on both sides of the slab preserves mirror symmetry.) The band diagrams of these two structures are given in Fig. 5.2 and Fig. 5.3, respectively; their calculation and interpretation is described in [62]. Each band diagram includes a shaded region (defining the light cone) and discrete (guided) bands lying below the light cone, which are

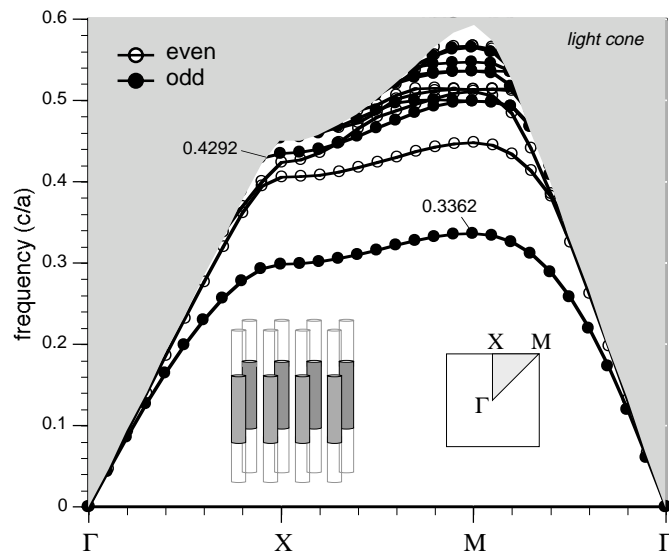


Figure 5.2: Band structure for the rod slab from Fig. 5.1(a). Hollow and filled circles represent modes which are even and odd with respect to the horizontal ($z = 0$) mirror plane bisecting the slab. There is a band gap in the odd (TM-like) modes in the frequency range $0.3362\text{--}0.4292 c/a$.

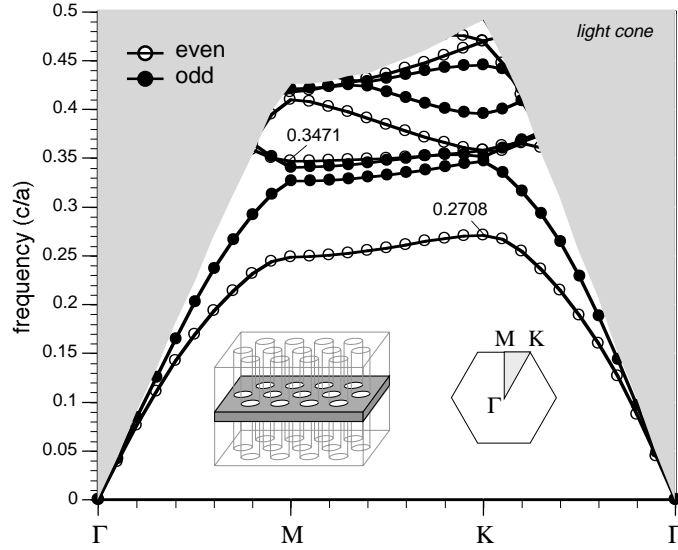


Figure 5.3: Band structure for the hole slab from Fig. 5.1(b). Hollow and filled circles denote even and odd modes as in Fig. 5.2. There is a band gap in the even (TE-like) modes in the frequency range 0.2708–0.3471 c/a .

confined vertically in the vicinity of the slab by index confinement. (Thus, the slab forms a planar waveguide, but in this chapter we use the term “waveguide” only to describe linear waveguides.) The presence of the light cone, with the corresponding constraint that all guided modes be index-confined in the vertical direction, gives rise to the most significant differences between photonic-crystal slabs and their two-dimensional cousins. The guided modes are either even or odd with respect to the horizontal mirror symmetry plane of the slab; this is analogous to TE and TM, respectively, in two dimensions, although here the modes are not completely polarized except in the symmetry plane. The rod and hole slabs have band gaps in their odd and even modes, respectively, though these are not complete gaps due to the light cone. Nevertheless, in the gap no *guided* modes exist for the corresponding symmetry. The slab band gap has been shown to support resonant cavities in point defects [59]–[60], and we will show in this chapter that it can also confine waveguide modes in linear defects. The differing thicknesses of the slabs derive from the polarizations of the modes exhibiting the band gap, and were chosen to achieve large gaps.

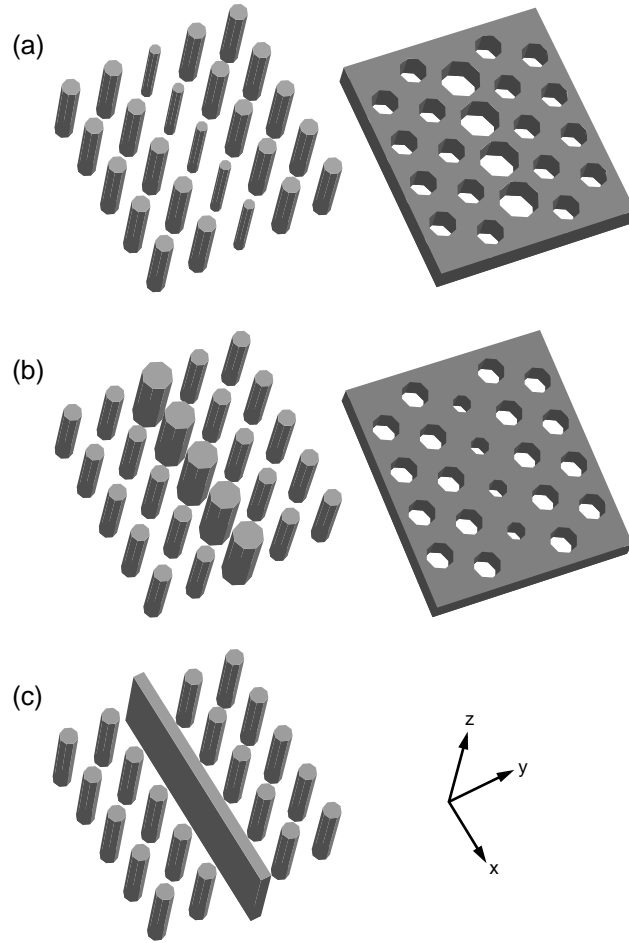


Figure 5.4: Linear defects, which give rise to waveguide modes, in the rod and hole photonic-crystal slabs from Fig. 5.1. The low-index material is not shown, but “extrudes” above and below the structures as in Fig. 5.1. (a) Reduced-index waveguides, created by decreasing/increasing the radii of a line of nearest-neighbor rods/holes, respectively. (b) Increased-index waveguides, created by increasing/decreasing the radii of a line of nearest-neighbor rods/holes, respectively. (c) Dielectric-strip waveguide surrounded by the rod slab, from which a line of nearest-neighbor rods has been removed and replaced with the waveguide.

In order to create a linear waveguide, we break the periodicity of the slabs in one direction by creating linear defects of various kinds, as shown in Fig. 5.4. The possible perturbations of the slab fall into two main categories: reduced-index waveguides, as in Fig. 5.4(a), for which the amount of high-index material is decreased; and increased-index waveguides, as in Fig. 5.4(b), for which the amount of high-index material is increased. We also consider a third case, shown in Fig. 5.4(c), in which a column of rods is removed

and replaced by a conventional dielectric “strip” waveguide. In all cases, we refer to the direction along the waveguide as x , the perpendicular in-plane direction as y , and the vertical direction (out of plane) as z .

The remainder of the chapter begins with a discussion of the computational methods that we employed, followed by detailed analyses of reduced-index, increased-index, and strip waveguides in photonic-crystal slabs. We then show that waveguides along nearest-neighbor directions are the most feasible. Finally, by estimating a lower bound on the fraction of the electric-field energy in the dielectric, we argue that it is not possible for photonic-crystal slabs to guide modes that reside primarily in air (unlike two- or three-dimensional crystals with complete gaps). (Guided modes can still reside, however, in a reduced-index region.)

5.2 Computational Method

The band structures of Fig. 5.2 and Fig. 5.3 were computed as described in [62], using preconditioned conjugate-gradient minimization of the Rayleigh quotient [77] in a three-dimensional plane-wave basis [78], with an imposed periodicity at a large interval in the vertical direction. This is a general, non-separable, vectorial solution of the full Maxwell’s equations in a complete three dimensional basis [79]–[80]—the only approximations are the discretization of the system (*i.e.*, the planewave cutoff) and the imposition of a vertical supercell (which has a negligible effect on the localized, exponentially-decaying guided modes). The light cone boundary derives from the lowest two-dimensional band of the low-index extruded region.

Calculation of projected band structures for line defects (linear waveguides) is similar to the above, except that an additional supercell of seven or eight periods is used in the direction perpendicular to the waveguide. (Waveguide modes are sufficiently localized so

that the adjacent waveguides introduced by the supercell have a negligible effect on the mode frequencies.) The bands from the unperturbed slab are projected onto the Brillouin zone of the line defect, and form the boundaries of the slab band continuum. Both the slab bands and the light cone are depicted with a uniform shading in the projected band structure, despite the varying density of states in these regions. Only modes that fall outside both the light cone and the slab bands are truly guided in the line defect, decaying exponentially away from the waveguide. When the guided bands cross into the continuum regions, they become resonances that extend infinitely away from the waveguide, albeit with low amplitude. We do not consider such resonance modes in this chapter.

5.3 Reduced-Index Waveguides

Photonic-crystal waveguides formed by removing high-index material [44],[47],[48], effectively reducing the index of the waveguide compared to its surroundings, are dramatic examples of the photonic band gap, as they have no analogues in conventional index-guiding. Such waveguides are also possible in photonic-crystal slabs, such as in Fig. 5.4(a) where we have reduced the radius of a line of rods or increased the radius of a line of holes. The effective index is lower in the waveguide than in the surrounding slab, so that modes can only be constrained horizontally by the band gap. On the other hand, the index is higher in the waveguide than the regions above and below the slab, so that modes can be guided vertically by index-confinement. In this section, we will first consider the reduced-radius rod-slab waveguide and then the increased-radius hole-slab waveguide.

The dispersion relations for various reduced-radii rod-slab waveguides are shown in Fig. 5.5. There is only one guided band for each radius, so the waveguide is single-mode. The line defect is periodic along its axis, and so the dispersion relation is plotted on the $\Gamma-X$ reduced Brillouin zone of this lattice. The light cone and the bulk slab bands from Fig. 5.2 are projected onto their $\Gamma-X$ wavevector component, resulting in the continuous

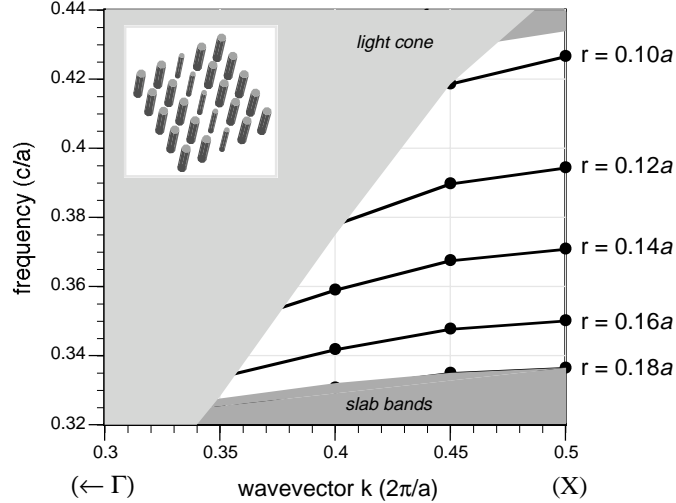


Figure 5.5: Projected band structure for the reduced-index rod-slab waveguide from Fig. 5.4(a), showing the (odd-symmetry) guided bands for various defect rod radii (compared to the bulk radius of $0.20a$).

shaded regions of Fig. 5.5. This band diagram continues symmetrically beyond X , so all of the bands must have zero group velocity (slope) at X in order to be analytic functions of the wavevector. This condition of zero slope at the Brillouin-zone edge also applies to all of the other waveguide dispersion relations that we consider in this chapter. (The zero slope is not apparent in Fig. 5.5 because the bands were only computed at a few points.)

These discrete bands, all of odd (TM-like) symmetry, decay exponentially away from the defect because they fall in neither the light cone (vertically-radiating modes) nor the continuum of odd slab-bands (which propagate freely within the slab), and so cannot couple with those extended modes. This strong confinement is illustrated in Fig. 5.6, which depicts horizontal and vertical cross sections of the electric-field z component at X for the radius $0.14a$ line defect. We call this an s mode because the field distribution around a rod resembles an s electron state, in contrast to other waveguide modes in subsequent sections.

By the time the defect radius decreases to $0.08a$ (versus $0.20a$ in the bulk), the guided mode has entirely disappeared into the light cone. This stands in stark contrast to a two-dimensional lattice of rods, in which a line defect with the rods entirely removed still sup-

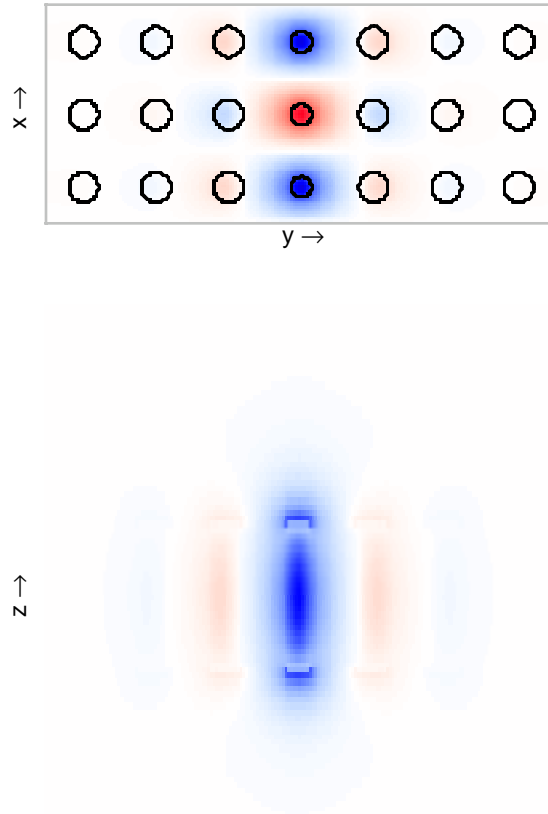


Figure 5.6: Field cross-sections for the rod-radius $0.14a$ guided mode at X ($0.3707 c/a$), showing the z component of the electric field. The contours of the dielectric function are shown in black. (a) Horizontal (xy) cross-section in the plane bisecting the slab ($z = 0$). (b) Vertical (yz) cross-section in the plane perpendicular to the waveguide and bisecting a row of rods.

ports a guided mode [44],[47],[48]. This difference is due solely to the slab’s constraint of vertical index confinement—that is, the guided bands must lie outside the light cone. The two-dimensional dispersion relation of the removed-rod waveguide disappears into the upper gap edge at a wavevector of roughly $0.4 (2\pi)/a$ [44], so it is not surprising that it falls into the light cone of the slab in this case. Even more generally, however, we do not expect a photonic-crystal slab waveguide to support guided modes that exist primarily in

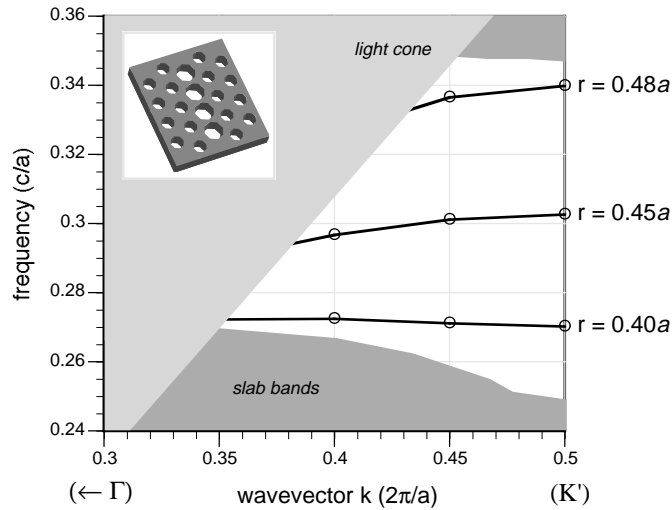


Figure 5.7: Projected band structure for the reduced-index hole-slab waveguide from Fig. 5.4(a), showing the (even-symmetry) guided bands for various defect hole radii (compared to the bulk radius of $0.30a$).

air—a large fraction of the electric field must lie within the dielectric. The basic reason for this is that the vertical confinement of the mode entails an “energy cost,” and having much of the field within the dielectric is the only way to bring the mode back down below the light cone. We quantify this argument in section VII. In the case of the Fig. 5.5 waveguides, the fraction of the electric-field energy inside the high dielectric at X ranges from 55% for radius $0.10a$ to 68% for radius $0.18a$. (Nevertheless, the mode is still within the reduced-index region and is not index-guided in the plane.)

A reduced-index waveguide in the hole slab is formed by increasing the radius of a line of nearest-neighbor holes. The resulting band diagrams for various radii are shown in Fig. 5.7. Here, only modes with even (TE-like) symmetry are depicted, since that is the symmetry of the bands exhibiting a gap. As with the rods, there is only a narrow range of radii that supports guided modes (all of which are single-mode). Although the radius $0.40a$ band is guided, it is barely in the band gap, which begins at the maximum frequency of the lower slab bands. The projected band structure for line defects retains a wavevector quantum number, and this means that the range of frequencies available for guided modes may

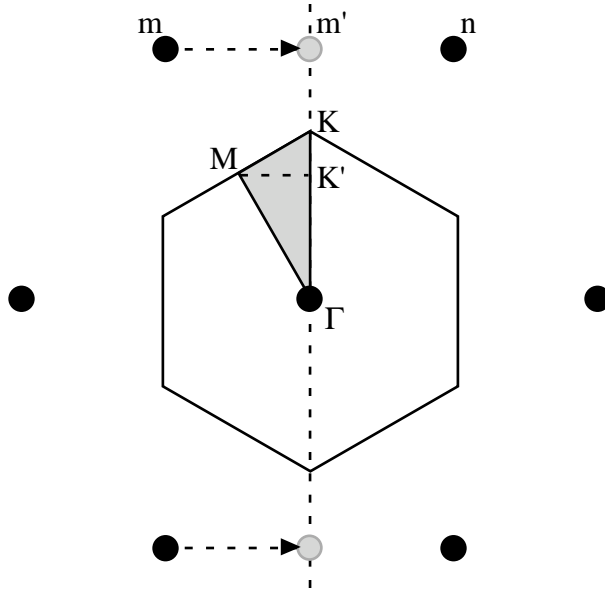


Figure 5.8: Diagram of how the reciprocal lattice is projected onto Γ - K (vertical dashed line) for a nearest-neighbor line defect in the hole slab. The black dots are the reciprocal lattice points, the outlined hexagon is the first Brillouin zone, and the grey triangle within it is the reduced Brillouin zone. K' is the boundary of the first Brillouin zone in the projected lattice.

be greater than the range of the band gap, as is the case here at the K' point. Since this quantum number disappears at bends and other places that break translational symmetry, however, it is important to consider only guided modes within the true band gap.

The nearest-neighbor direction corresponds to the Γ - K direction in the slab Brillouin zone, but K is not the edge of the line-defect Brillouin zone. Figure 5.8 depicts how the reciprocal lattice of the slab is projected to form the reciprocal lattice of the line defect. The boundary of the projected Brillouin zone lies halfway between Γ and m' (projected from m)—this is the point K' , projected from M (which lies halfway between Γ and m). Use of the correct projection is important for the computation of the light cone and the slab bands continuum for Fig. 5.7.

As before, the waveguide band is s -like and strongly localized both vertically and horizontally, as shown in Fig. 5.9. In this case, we depict the z component of the magnetic field because the mode is TE-like. Since the electric field loops around the magnetic field

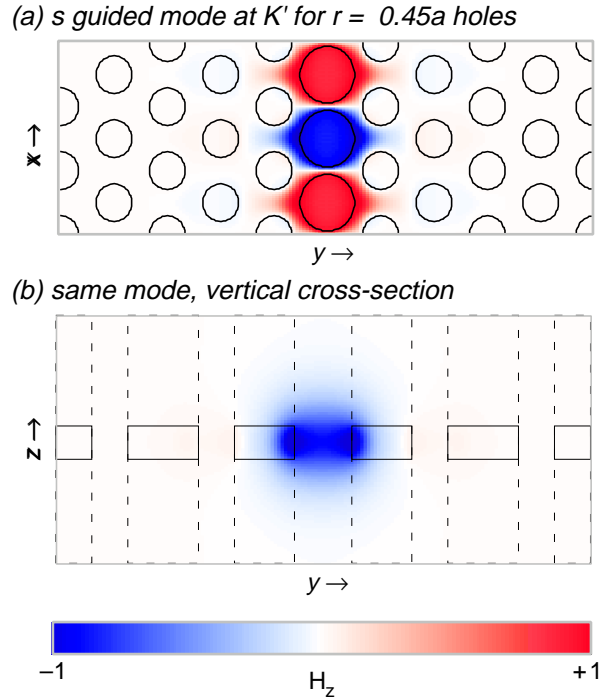


Figure 5.9: Field cross-sections for the hole-radius $0.45a$ guided mode at K' ($0.3026 c/a$), showing the z component of the magnetic field. The contours of the dielectric function are shown in black. (a) Horizontal (xy) cross-section in the plane bisecting the slab ($z = 0$). (b) Vertical (yz) cross-section in the plane perpendicular to the waveguide and bisecting a defect hole.

like fields around a current, it is concentrated in the dielectric just as it was for the rods—the fraction of the electric-field energy inside the high dielectric at K' ranges from 63% for radius $0.48a$ to 86% for radius $0.40a$. The critical parameter in this case is the thickness of the narrow veins between the holes. It is only as that thickness approaches zero that the guided modes rise up far into the gap, while for small changes in defect hole radius there are no modes at all in the gap. In contrast, when the rod-slab radius is decreased, the electric field is forced directly into the air and the frequency quickly increases even for small perturbations.

There are many other ways to create reduced-index waveguides. For example, one could remove a row of rods and then shift the lattices towards one another on either side of the waveguide. This might, however, be inconvenient for applications in high-density inte-

grated optical devices, since lattice dislocations from different components may conflict. When a row of rods is removed, one can run into additional problems due to the electric field being concentrated in the adjacent rows of remaining rods. This leads to near-degeneracies between states that are even and odd with respect to the $y = 0$ mirror plane, creating a multi-mode waveguide. Yet another way to create a reduced-index waveguide is to actually decrease the dielectric constant within a linear (one-dimensionally periodic) region, without necessarily changing the geometric structure.

5.4 Increased-Index Waveguides

A waveguide can also be created by a line defect in which high-dielectric material is added, leading to an increased effective index. As shown in Fig. 5.4(b), we examine two types of increased-index waveguides: first, a rod slab with a line of increased-radius rods; and second, a hole slab with a line of decreased-radius holes. In general, increased-index defects can introduce two sorts of guided modes. First, there are modes that lie below both the light cone and the slab band continuum—these modes are conventionally index-guided horizontally as well as vertically. They do not satisfy our third waveguide criterion, however: they do not lie in the band gap, being below the lowest slab bands. Thus, we do not consider such guided bands here. The second type of guided mode, which we shall study, is a state that is pulled down into the gap from the upper slab bands. Such a mode is confined horizontally purely by the band gap and index-confined vertically, just like the reduced-index waveguides in the previous section. Because these increased-index modes derive from the upper slab bands, which have a higher density of states and often come in degenerate pairs (at symmetry points), there is a greater tendency here towards multi-mode waveguides. As we shall see, however, the single-mode criterion can be satisfied if the effective index is not increased too greatly.

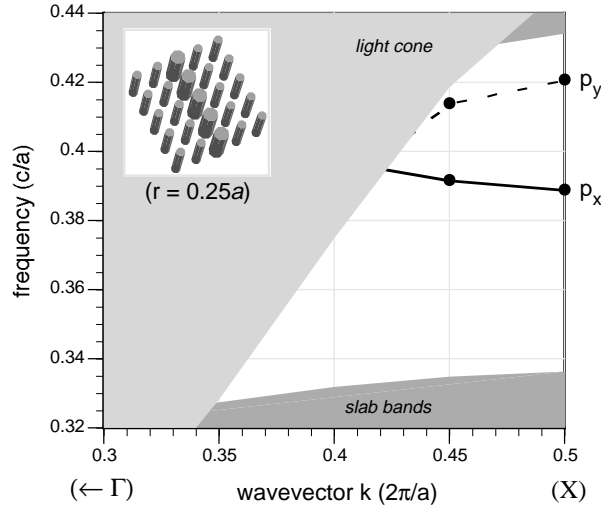


Figure 5.10: Projected band structure for the increased-index rod-slab waveguide from Fig. 5.4(b), showing the (odd-symmetry) guided bands for defect rod radius $0.25a$ (compared to the bulk radius of $0.20a$).

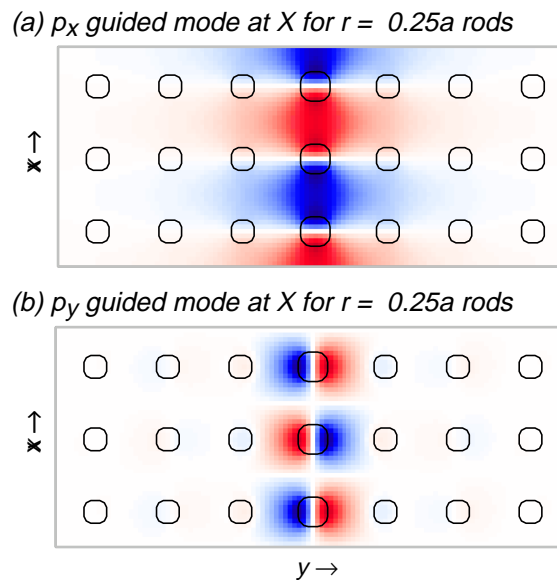


Figure 5.11: Horizontal ($z = 0$) E_z field cross-sections for the rod-radius $0.25a$ guided modes at X . (a) p_x guided mode ($0.3887 c/a$). (b) p_y guided mode ($0.4205 c/a$).

Increasing the radii of a line of nearest-neighbor rods results in the band diagram of Fig. 5.10, showing the odd bands for rod radii of $0.25a$. This rod radius gives rise to a pair of non-degenerate guided modes, labeled p_x and p_y , but because these modes do not overlap in frequency the waveguide remains single-mode. (Note that here, one can identify modes with the same wavevector but opposite group velocities.) Fig. 5.11 shows cross-

sections of the electric field for the two states at defect radius $0.25a$. As can be seen in Fig. 5.11, the field around a rod is similar to a p electron state, oriented in the x and y directions for the p_x and p_y states, respectively.

When the defect rod radius is increased to $0.275a$, three additional states are pulled down into the gap, as shown in Fig. 5.12. In this case, the p_x and p_y states cross (which is possible because they have different symmetry with respect to the $y = 0$ mirror plane), meaning that they are not single-mode over their whole guided range. The states labeled $p_x^{(2)}$ and $p_y^{(2)}$ have similar horizontal cross-sections to p_x and p_y , but are second-order excitations in their vertical cross section (*i.e.*, their electric fields have two vertical nodes). The state labeled $s_H^{(1)}$ is a first-order vertical excitation of the lowest TE-like band (*i.e.*, a single vertical node in the magnetic field), which is s -like in its H_z cross-section. (The lowest TE-like band is of even symmetry, but adding a vertical node transforms it to odd symmetry and allows it to be guided by the gap.) The $p_x^{(2)}$, $p_y^{(2)}$, and $s_H^{(1)}$ states have no analogues in two dimensions.

Figure 5.13 shows the dispersion relation of even modes for the decreased-radius hole waveguides with radii of 0 , $0.15a$ and $0.25a$. The waveguides are single-mode at each fre-

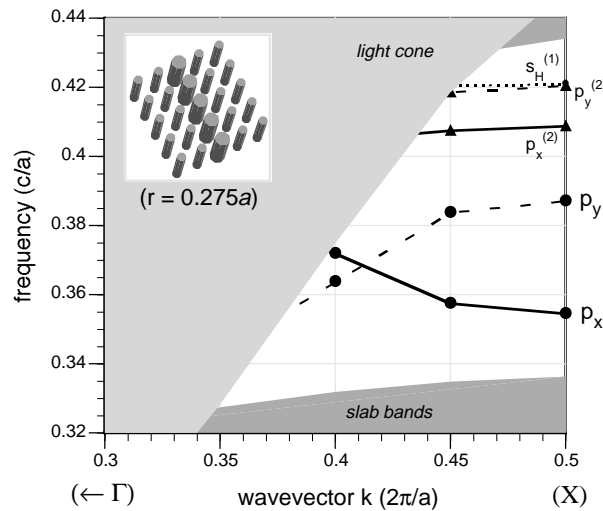


Figure 5.12: Projected band structure for the increased-index rod-slab waveguide from Fig. 5.4(b), showing the (odd-symmetry) guided bands for defect rod radius $0.275a$.

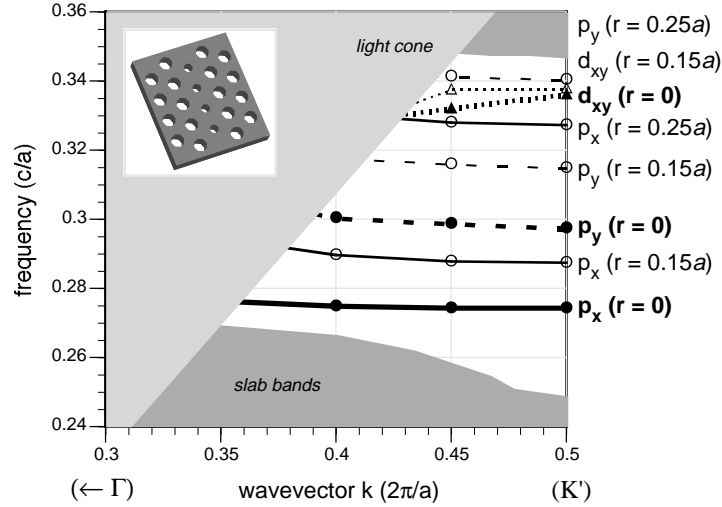


Figure 5.13: Projected band structure for the increased-index hole-slab waveguide from Fig. 5.4(b), showing the (even-symmetry) guided bands for defect hole radii $0.15a$ and $0.25a$ (compared to the bulk radius of $0.30a$). For each radius, there are two or three guided bands: p_x (solid line), p_y (dashed line), and d_{xy} (dotted line, triangles). Guided bands for a defect hole radius of 0 , in which a row of holes is entirely removed, are shown in bold.

quency, and there are p_x and p_y modes for each radius as for the rod slabs. In addition, the radius 0 and $0.15a$ waveguide supports a higher-order d_{xy} mode. All three modes for radius $0.15a$ are depicted in Fig. 5.14.

In both the rod and the hole slabs, increasing the amount of dielectric brings more and more states into the gap, which makes it harder to achieve single-mode waveguides. One can ameliorate that situation by removing a row of holes and then shifting the lattices on either side of the waveguide towards one another (reducing the effective index in the waveguide). Such a waveguide was considered for the hole structure in two dimensions by [46], and similar geometries should be feasible in photonic-crystal slabs.

5.5 Strip Waveguides in Photonic-Crystal Slabs

In this section, we study another possible photonic-crystal slab waveguide, formed by replacing a row of rods in the slab of Fig. 5.1(a) with a conventional dielectric strip waveguide of the same thickness as the slab, as shown in Fig. 5.4(c). As before, there is a

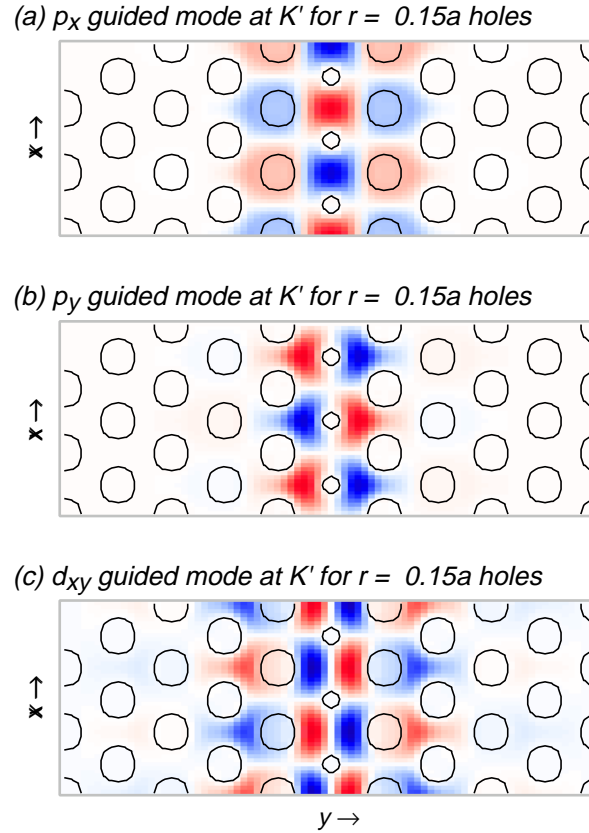


Figure 5.14: Horizontal ($z = 0$) H_z field cross-sections for the hole-radius $0.15a$ guided modes at K' . (a) p_x guided mode ($0.2874 c/a$). (b) p_y guided mode ($0.3148 c/a$). (c) d_{xy} guided mode ($0.3377 c/a$).

low-index substrate with the same cross-section as the slab and waveguide both above and below this structure.

In Fig. 5.15 we show the odd-symmetry band diagram for such a waveguide with a width of $0.25a$. In this case, the frequency axis extends down below the slab band continuum so that both the index-guided modes (dashed lines, below the slab bands) and the gap-guided modes (solid lines) are visible. The two gap-guided modes at X are shown in Fig. 5.16—they differ by a 90° phase shift—and the horizontal cross-sections of the index-guided modes are nearly identical to those of the gap-guided modes. The vertical cross-sections of the gap- and index-guided modes differ sharply, however, as shown in

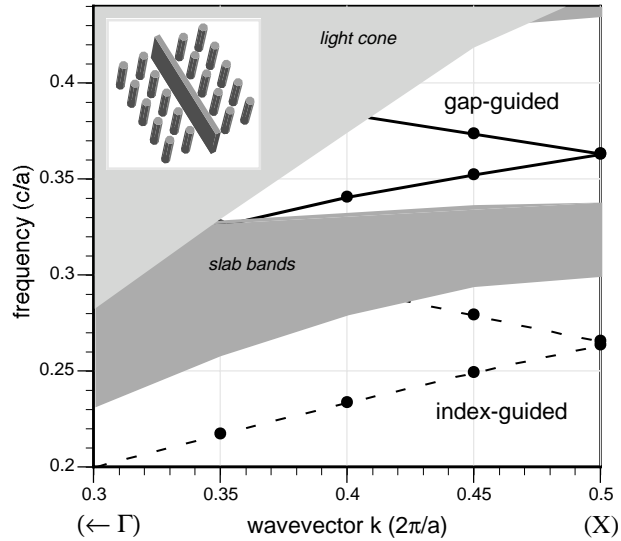


Figure 5.15: Projected band structure for the strip waveguide in a rod slab, from Fig. 5.4(c), showing the (odd-symmetry) guided bands for waveguide width $0.25a$. There are two pairs of guided bands: gap-guided (solid lines) and index-guided (dashed lines). The gap-guided bands are split at X into a 0.07% gap, while the index-guided bands are split at X into a 0.8% gap.

Fig. 5.17—the index-guided mode is the fundamental mode and the gap-guided mode is a second-order excitation.

In the absence of the rod slab, both of the modes shown would lie below the light cone and be purely index-guided. Furthermore, since the strip has translational symmetry, the Brillouin zone that we apply in Fig. 5.15 would be artificial—the guided bands would simply fold back and forth at the zone boundaries and be degenerate at X . Because this folding is synthetic, the guided modes do not couple with the light cone even when they fold on top of it. Adding the rod slab to this system has three effects: first, the frequencies of the strip are perturbed; second, the degeneracy at X is broken by the true periodicity of the system; and third, an upper cutoff for the guided modes is introduced by the light cone (since the band folding is no longer artificial). (Also, if it were not for the band gap of the rods, the gap-guided modes would only be resonances—they would couple with the slab and leak slowly away.) Because the guided modes are strongly localized around the dielec-

tric strip, however, the effect of the rods is small. The bands shown in Fig. 5.15 have an root-mean-square deviation of less than 1% from the bands of an isolated strip, and the splitting at X is also less than 1%. Because of the true periodicity of the system, the folded band structure must be analytic and the bands thus have zero group velocity at X —due to the weakness of the perturbation, though, the slope drops towards zero only as the wavevector comes within 1% of X . The greater the periodic perturbation, the greater the splitting and the larger the region of low group velocity, leading eventually to the highly-nonlinear dispersions of the previous sections.

Despite the similarity in frequency with the modes of an isolated strip, the waveguide here has a significant advantage—it supports (single-mode) bands in the photonic band gap, which will inhibit losses in the cases of sharp bends and cavity interactions. The index-guided lower modes in this structure, on the other hand, have no advantage over a conventional waveguide. These modes might otherwise be preferred since they are the fundamental modes of the waveguide; if this is an important consideration, they can be pushed up into the band gap by reducing the waveguide width to $0.10a$ (at the expense of increased height-to-width aspect ratio).

5.6 Waveguides in Other Directions

In the preceding sections, we have only considered waveguides in the nearest-neighbor directions of the slabs. Another possibility might be waveguides in the other symmetry directions (*i.e.*, along next-nearest neighbors), as this is known to give rise to guided modes in two dimensions [47]. Because of the presence of the light cone, however, such waveguides are not possible in the photonic-crystal slab structures of Fig. 5.1.

The problem is that, for the next-nearest-neighbor directions, the edge of the projected Brillouin zone comes at too small a wavevector, and the resultant folding of the light cone

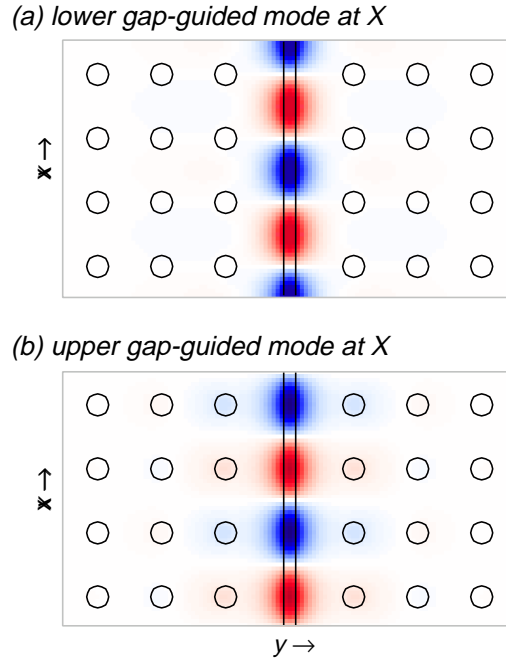


Figure 5.16: Horizontal ($z = 0$) E_z field cross-sections for the $0.25a$ strip-waveguide gap-guided modes at X . (a) Lower gap-guided mode ($0.3626 c/a$). (b) Upper gap-guided mode ($0.3629 c/a$).

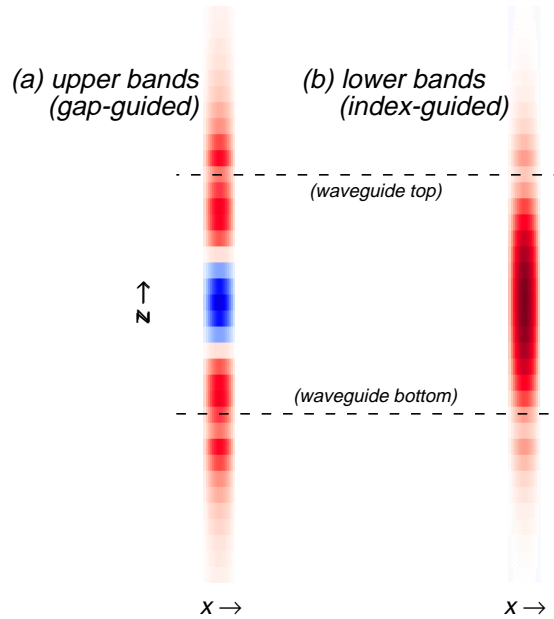


Figure 5.17: Vertical ($y = 0$) E_z field cross-sections of a unit cell, parallel to and bisecting the waveguide, for the strip-waveguide gap-guided and index-guided modes at X . (a) Upper gap-guided mode ($0.3629 c/a$). (b) Lower index-guided mode ($0.2632 c/a$).

eliminates the band gap. For example, the next-nearest-neighbor direction in the hole slab corresponds to the M direction. When the reciprocal lattice is projected onto $\Gamma-M$, however, one of the reciprocal lattice points (n in Fig. 5.8) falls onto M . Therefore, the Brillouin zone edge is at $M/2$, at which point the light-cone frequency is only $0.223 c/a$. Since this upper cutoff for guided modes is lower than the gap bottom, there is no band gap and photonic-crystal waveguides are not possible. Similarly, in the rod slab (square lattice), the Brillouin-zone boundary for the next-nearest-neighbor direction is also at $M/2$, corresponding to a light-cone frequency of $0.332 c/a$ —again below the gap bottom.

If the rod slab were suspended in air (*i.e.*, no substrate), the light cone at $M/2$ would be at a frequency of $0.354 c/a$, while the band-gap bottom rises to only $0.3397 c/a$, leaving a slight (4%) band gap. So, in general, next-nearest-neighbor waveguides in photonic-crystal slabs are not impossible, but the available band gap is greatly reduced from the nearest-neighbor case.

5.7 Estimating the Field Energy in the Dielectric

In section III, we suggested that much of the field energy must lie within the dielectric material, not in the air, in order for a waveguide mode to fall underneath the light cone and be guided. In this section, we justify that claim; moreover, we find an analytic relationship between a mode's localization and the fraction of it inside the high-index material, explaining why some waveguide modes from two-dimensional systems cannot exist in a slab. First, we derive a rigorous lower bound on the fraction of the electric field energy inside the dielectric, in terms of the spatial extent of the mode and the degree to which it is transverse. Then, we approximate the transversality and spatial confinement of the mode to compute a numerical value for this lower bound, and compare with the results of our exact calculations.

The frequency of a guided mode can be expressed as an expectation value of the Maxwell operator for the electric field, evaluated over the unit cell:

$$\omega^2 = \frac{\langle E | \nabla \times \nabla \times | E \rangle}{\langle E | \varepsilon(\hat{x}) | E \rangle}, \quad (5.1)$$

where we take c to be unity. Now, we suppose that the dielectric function is piecewise constant, and write the field $|E\rangle$ as a sum of fields $|E\rangle_i$, such that $|E\rangle_i$ equals $|E\rangle$ where the dielectric is ε_i and is zero elsewhere. Then, we can rewrite Eq. (5.1) as:

$$\omega^2 = \frac{\langle E | \nabla \times \nabla \times | E \rangle}{\langle E | E \rangle} \sum_i \frac{1}{\varepsilon_i} \left(\frac{\langle E | \varepsilon_i | E \rangle_i}{\langle E | \varepsilon(\hat{x}) | E \rangle} \right) = \frac{\langle E | \nabla \times \nabla \times | E \rangle}{\langle E | E \rangle} \sum_i \frac{f_i}{\varepsilon_i}. \quad (5.2)$$

Here, f_i is the fraction of the electric field energy that is in the dielectric ε_i , and must sum to unity. In order to be guided, the frequency in Eq. (5.2) must be less than the light cone frequency ω_{lc} at that point in the Brillouin zone. We will write the electric field in a Fourier basis, in which case the curl operations become simply the total Fourier wavevector k squared. Because the electric field is not divergenceless, however, the curl also projects out the longitudinal field component $|E_L\rangle$, leaving only the transverse component $|E_T\rangle$. ($|E_T\rangle$ is divergenceless but not curl-free, while $|E_L\rangle$ is curl-free but not divergenceless. $|E_L\rangle$ is the static electric field of the bound charges.) In this case, the combination of Eq. (5.2) with the guiding constraint becomes:

$$\langle k^2 \rangle_{E_T} \frac{\langle E_T | E_T \rangle}{\langle E | E \rangle} \sum_i \frac{f_i}{\varepsilon_i} < \omega_{lc}^2 \leq k_B^2. \quad (5.3)$$

Here, k_B is the Bloch wavevector in the Brillouin zone of the line defect, and the second relation derives from the fact that the light cone must lie beneath the light cone of vacuum. $\langle k^2 \rangle_{E_T}$ is the mean-square wavevector of the transverse field, which is inversely related to the mode's spatial confinement. We will denote the transverse fraction of the field, $\langle E_T | E_T \rangle / \langle E | E \rangle$, by f_T . If we have a two-component system (dielectric ε and air), Eq. (5.3) leads to the following rigorous lower bound on the fraction of energy in the

dielectric:

$$f_{diel} > \frac{1 - k_B^2 / (f_T \langle k^2 \rangle_{E_T})}{1 - 1/\epsilon}. \quad (5.4)$$

The more localized the field is, the greater becomes the mean-square wavevector and the greater the minimum f_{diel} .

In order to evaluate Eq. (5.4) further, we need to estimate both f_T and a lower bound on the mean-square wavevector. We treat the wavevector by breaking it into its Cartesian components and examining each component separately (dropping the E_T subscript):

$$\langle k^2 \rangle = \langle k_x^2 \rangle + \langle k_y^2 \rangle + \langle k_z^2 \rangle. \quad (5.5)$$

Since the field is of the Bloch form (for a system periodic in x) and k_B is restricted to the first Brillouin zone, it can be shown $\langle k_x^2 \rangle$ is bounded below by k_B^2 . Furthermore, we will simply drop the $\langle k_y^2 \rangle$ term, which corresponds to horizontal confinement of the waveguide mode—we consider the best case for creating a guided mode, when it is very weakly confined in the plane (*e.g.*, near the gap edge).

$$\langle k^2 \rangle \geq k_B^2 + \langle k_z^2 \rangle. \quad (5.6)$$

Up to this point, we have maintained a strict lower bound on f_{diel} . Now, however, we will make two approximations. The first is based upon the observed mode profile, such as in Fig. 5.6 and Fig. 5.9. We are only interested in modes that are localized in z , so we assume that the vertical mode profile is a Gaussian with a standard deviation less than the height h of the slab, and thus:

$$\langle k_z^2 \rangle \geq \frac{1}{h^2}. \quad (5.7)$$

Second, in the case of the rod slab, we will assume f_T to be nearly unity—that is, neglect the longitudinal component of the field. We justify this assumption by noting that the fields of the rod modes are very similar to those of the two-dimensional structure [62], in

which the TM electric field is divergenceless. The only longitudinal component comes from the small regions near the rod ends where the electric field has a component normal to the dielectric surface. (For example, the rod-slab defect-radius $0.10a$ mode is 86% transverse at X .)

Finally, combining equations (4), (6), and (7), we have the following approximate lower bound on the fraction of electromagnetic energy in the dielectric (for TM-like fields):

$$\text{minimum } f_{diel}^{\text{tm-like}} \approx \frac{1}{([hk_B]^2 + 1) \cdot (1 - 1/\epsilon)}. \quad (5.8)$$

In the case of the rod slab at the X point, this lower bound evaluates to 0.545, which is very close to the minimum of 55% that we observed for the smallest defect rod supporting guided modes (radius $0.10a$). If we instead take f_T to be the observed value of 0.86, then the minimum f_{diel} becomes 0.457.

In contrast, let us consider the purely two-dimensional system of rods. Here, there is no light cone, so the upper-bound on the frequency is a constant ω_{max} (the top of the gap) instead of k_B . Also, there is no vertical confinement and the TM electric field is fully transverse. Thus, Eq. (5.4) becomes the following strict bound:

$$f_{diel}^{\text{tm}} > \frac{1 - \omega_{max}^2 / (k_B^2 + \langle k_y^2 \rangle)}{1 - 1/\epsilon}. \quad (5.9)$$

Therefore, for small Bloch wavevectors k_B (or for large ω_{max}) in two dimensions, it is possible to both confine in the y direction and also have a low fraction of the field in the dielectric. This corresponds precisely to what is observed in [44], in which guiding in air is demonstrated, but not near the edge of the Brillouin zone (large k_B). (Similarly, it should be possible to guide primarily in air using three-dimensional crystals with a complete band gap.)

We should emphasize that the lower bounds derived above hold for the electric field energy, not the magnetic field energy. Indeed, for the reduced-radius rod waveguides, we observed the fraction of the magnetic field energy in the dielectric to drop as low as 15% (for radius $0.14a$ rods).

5.8 Summary

Photonic-crystal slabs, with their hybrid approach of in-plane band gap combined with vertical index-confinement, promise a viable route to achieving photonic-crystal waveguide effects in real systems. We have shown how single-mode, in-gap waveguides can be achieved in a variety of ways with photonic-crystal slabs. At the same time, the finite height of the slabs and restriction of index-confinement have given rise to fundamentally new concerns compared to two-dimensional systems. For example, slab waveguides do not support modes guided in air or along next-nearest-neighbor directions, unlike in two dimensions. Thus, photonic-crystal slab waveguides require a full three-dimensional analysis, for which the formalism of the projected band structure provides a powerful visual and analytical tool.

Chapter 6

High- Q Cavities in Photonic-Crystal Slabs*

6.1 Introduction

Radiation losses are a general problem in many optical devices. A complete photonic band gap (PBG) prohibits losses [81], but the difficulty of its fabrication has spurred interest in simpler 2d-periodic dielectric structures of finite height: photonic-crystal slabs [82]–[89]. Without a complete gap, however, it is impossible to prevent radiation losses whenever translational symmetry is entirely broken, for example by a resonant cavity [83] or a waveguide bend [48]. In some cases, radiation is desirable [82], but for planar optical devices vertical scattering will generally be a loss mechanism and must be minimized.

In this chapter, we describe two independent general principles, mode delocalization and multipole cancellations, by which low radiation losses may be achieved in photonic-crystal slabs and in many related systems. We demonstrate these mechanisms by showing how they may be used to design high- Q cavities in a model photonic-crystal slab structure, a two-dimensional lattice of finite-height dielectric rods suspended in air. We also consider a simpler 2d analogue, a 1d sequence of dielectric cylinders in air, which lends itself to easier semi-analytical study. Both model structures are depicted in Fig. 6.1.

A resonant cavity of Lorentzian lineshape is described by a quality factor, Q , which is the number of cycles for its energy to decay by $e^{-2\pi}$. The decay rate $1/Q$ is thus the bandwidth at half-maximum of any filter or other device based on the cavity. A typical cavity has several decay mechanisms, each of which is described by its own lifetime. First,

* This chapter is based on two articles that were recently accepted for publication: S. G. Johnson, S. Fan, A. Mekis, and J. D. Joannopoulos, “Multipole-cancellation mechanism for high- Q cavities in the absence of a complete photonic band gap,” *Appl. Phys. Lett.* (May 2001); and S. G. Johnson, A. Mekis, S. Fan, and J. D. Joannopoulos, “Molding the Flow of Light,” *Computing in Science and Engineering* (Summer 2001).

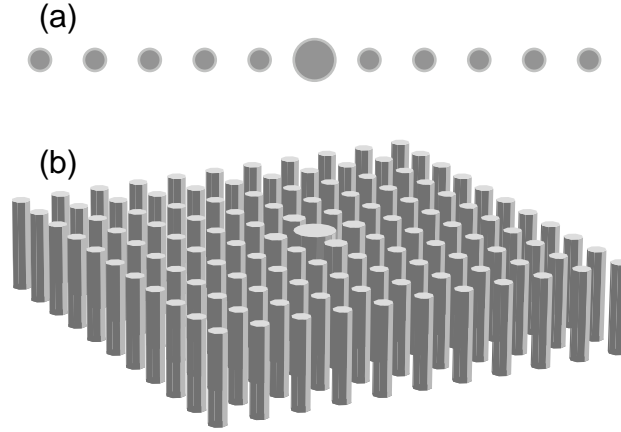


Figure 6.1: Photonic-crystal slabs in 2d and 3d, with defects (resonant cavities) formed by increasing the radius of a central rod. (a) 2d slab: a sequence of dielectric rods with lattice constant a , radius $0.2a$, and ϵ of 11.56. (b) 3d slab: a square lattice of dielectric rods with lattice constant a , radius $0.2a$, height $2a$, and ϵ of 12.

there are the radiation losses, characterized by Q_{\perp} (“perpendicular” to the slab)—this is the same as the total Q in an infinite, perfect crystal. Then, there are in-plane losses, Q_{\parallel} , due to the finite size of the crystal; these losses decrease exponentially as the crystal size increases. There are also losses due to absorption and disorder, which we do not consider in this chapter. Finally, there is the rate at which the light decays where you want it to go, *e.g.* into an adjacent waveguide for a filter, described by some Q_d (“device”). The total Q is then [92]:

$$Q = [Q_d^{-1} + Q_{\perp}^{-1} + Q_{\parallel}^{-1}]^{-1} \quad (6.1)$$

Since Q_{\parallel} can be made effectively infinite, Q_{\perp} determines an upper bound on the achievable total Q of the device. Moreover, the fraction of energy lost to radiation is given by [92],[83]:

$$\left(\frac{Q}{Q_{\perp}}\right)^2 = \left(\frac{Q_d}{Q_{\perp} + Q_d}\right)^2 \quad (6.2)$$

So, as long as Q_{\perp} is much larger than Q_d , the radiation losses will be very small. As a corollary, if Q_d is small, then Q_{\perp} need not be very high to achieve low losses. Since light decays very quickly into a waveguide from a bend, for example, small Q_{\perp} can be tolerated

there; whereas large Q_{\perp} is required for the case of high- Q devices/cavities we consider here. The purpose of this chapter is to maximize Q_{\perp} .

6.2 Computational Method

The mechanisms described herein are general for any type of photonic-crystal system, but for definiteness we will demonstrate this phenomenon in structures involving rods in air. The 3d system that we will examine is a square lattice of short dielectric rods in air (lattice constant a , radius $0.2a$, height $2a$, dielectric constant 12), shown in Fig. 6.1(b). This 3d structure is analyzed in [86] and exhibits perfectly-guided modes (extended in the slab and localized vertically). There is a bandgap in the guided modes of odd symmetry with respect to the horizontal mirror plane, corresponding to TM modes in 2d. We will also study a simple 2d analogue to the photonic-crystal slab, a one-dimensional sequence of dielectric rods (period a , radius $0.2a$, dielectric constant 11.56) [90], depicted in the inset of Fig. 6.1(a). Like the 3d slab, these 2d rods produce guided modes propagating in the lattice direction (x), with a bandgap from 0.264 – $0.448 c/a$, which are confined in the transverse (y) direction by index-guiding. (a is the lattice constant.) Here, we only consider TM-polarized light (electric field along z).

To analyze resonant cavities formed by point defects in the crystal, we use 2d and 3d finite-difference time-domain (FDTD) calculations with absorbing boundary conditions. The 2d computational cell contains 29 rods, with a $7a$ margin of air in the lateral directions. The 3d computational cell is an 11×11 lattice of rods, with the defect(s) at the center, surrounded by a margin of $4a$ above and below the rods and $2a$ on the sides. Around the sides of the crystal, we include a dielectric slab of the same height as the rods and extending to the boundaries of the cell, separated from the edge of the crystal by $0.5a$. The purpose of this slab is to leech off “whispering gallery modes” associated with the entire crystal, which otherwise have long lifetimes and make observation of the point-defect

mode more difficult; it did not appreciably affect the radiation Q_{\perp} of the cavities and only affected the in-plane Q_{\parallel} .

The FDTD grid resolution is 20 pixels/ a in 2d and 10 pixels/ a in 3d. At this resolution, the cylinders are “blocky” but are nevertheless periodic, and the basic PBG and guiding phenomena are not altered. (Some cases were also tried at 15 pixels/ a and consistent results were obtained.) There is a frequency shift, however, relative to the higher-resolution calculation in Fig. 5.1, and so we recomputed the odd-mode band gap using a single-rod FDTD calculation with Bloch-periodic boundary conditions at the sides (analyzing the spectrum excited by point sources as below), and found it to lie from 0.320–0.391 c/a .

We excite the defect modes of the cavity with one or more dipole sources, arranged in the same symmetry (if any) as the mode of interest. From the field amplitude as a function of time at a point in the cavity, the mode frequencies and decay constants (whence Q) are extracted by the filter-diagonalization method with a Fourier basis. In order to determine Q_{\perp} and Q_{\parallel} , the simulation is run again with Gaussian sources narrowly-peaked around the frequency of the desired mode. With only this mode excited, the vertical flux is measured through planes above and below the rods, and the horizontal flux through planes at the edges of the cell (extending to a height $0.5a$ above and below the rods). The lifetimes in each direction are then given by [85]:

$$Q_{\perp, \parallel} = \frac{\omega U}{P_{\perp, \parallel}}, \quad (6.3)$$

where ω is the angular frequency, U is the total electromagnetic energy, and P is the radiated power. U and P were time-averaged. The results were verified to be consistent with Eq. (6.1).

6.3 Mode Delocalization

In this section, we demonstrate that spatial delocalization of a cavity mode will generally

lead to increased Q_{\perp} . Since the mode becomes more delocalized as the edges of the band gap are approached, one then expects that Q_{\perp} will not peak at mid-gap but rather may increase towards one or both band edges.

The relationship between Q_{\perp} and localization can be seen in several ways. A qualitative argument is that the more spatially delocalized the mode is, the more localized it will be in wavevector space—so, a finite “wavevector” is approximately regained, and coupling to the light cone can be prevented by momentum conservation just as in the bulk slab or in linear waveguides. A more rigorous argument can be made using a Green’s function approach, also known as the volume-current method [94]. There, the field in a dielectric perturbation $\Delta\epsilon(\hat{x})$ is treated as a current $\vec{J} = -\frac{ik}{4\pi}\Delta\epsilon(\hat{x})\vec{E}(\hat{x})$ (assuming e^{-ickt} time dependence)—then, using a Green’s function \hat{G}_{ω} , the radiated field is expressed in terms of only the field at the perturbation:

$$\vec{E}(\hat{x}) = \int \hat{G}_{\omega}(\hat{x}, \hat{x}') 4\pi ik \vec{J}(\hat{x}') d^3\hat{x}'. \quad (6.4)$$

$\Delta\epsilon$ may be chosen with respect to *any* structure; this determines \hat{G}_{ω} . If we use the \hat{G}_{ω} of the unperturbed crystal, then $\Delta\epsilon$ only comes within the defect itself. This suggests two ways to decrease the radiated field: either delocalize the mode to decrease the field at the defect, or reduce the perturbation $\Delta\epsilon$ (which delocalizes the field, too).

To illustrate this effect, we create a monopole-like point defect in the photonic-crystal slab by changing the ϵ of a single rod, and the resulting Q as a function of ω is shown in Fig. 6.2. As expected, Q_{\perp} diverges towards the lower band edge (source of the defect mode), with the total Q dominated by the smallest of Q_{\perp} and Q_{\parallel} .

Another way to delocalize the mode, achieving greater delocalization for the same frequency, is to enlarge the size of the defect—in this case, we change ϵ in a 5-rod “super-

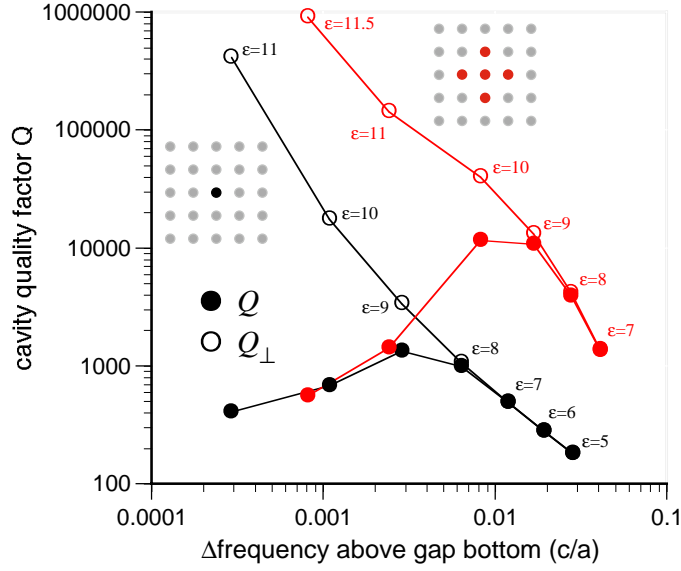


Figure 6.2: Total Q (solid circles) and radiation Q_{\perp} (hollow circles) plotted vs. $\Delta\omega$ above the band edge for monopole defects in the 11×11 slab of rods from Fig. 6.1(b). Black points are for a single-rod defect (left inset, black), and red points are for a 5-rod super-defect (upper-right inset, red).

defect.” The results, also shown in Fig. 6.2, display the same general behavior as a single defect but with dramatically increased Q . A total Q of over 10^4 is obtained for a defect ϵ of 9, 5% above the gap bottom, testifying to the localization of the state (shown in Fig. 6.3). To completely understand the super-defect, coherent scattering in Eq. (6.4) must be considered ; in effect, the super-defect forms a tiny photonic-crystal slab, whose guided modes only scatter from the boundaries. As the number of rods in the super-defect is increased, it will eventually become multi-mode, but in this case the cavity is single-mode.

A disadvantage of this method is that it increases the modal volume V , of concern since, *e.g.*, the enhancement factor η of spontaneous emission in a cavity goes like Q/V . We can estimate Q and η from Eq. (6.4) by assuming that \hat{G}_{ω} and \vec{E} are approximately constant inside the defect and neglecting coherence, in which case we find with the help of Eq. (6.3) that:

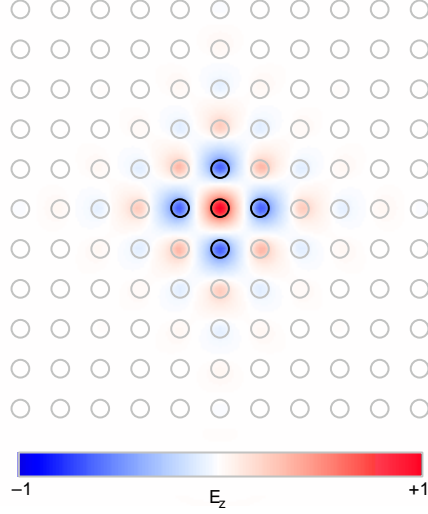


Figure 6.3: Monopole defect mode for $\epsilon=9$ super-defect ($Q = 11000$, $Q_{\perp} = 13000$): electric-field z-component in the plane bisecting the rods.

$$Q_{\perp} \sim \frac{\langle |\vec{E}|^2 \rangle}{\Delta\epsilon^2 \langle |\vec{E}|^2 \rangle_{defect}} \sim \frac{V}{\Delta\epsilon^2} \sim \frac{1}{\Delta\epsilon^2 \Delta\omega^{d/2}} \quad (6.5)$$

where $\Delta\omega$ is the frequency above the gap edge and d is the number of dimensions we are delocalizing (here, 2). (We use the fact that the decay length is proportional to $\Delta\omega^{-1/2}$.) It follows that η is proportional to $\Delta\epsilon^{-2}$, and thus the increase in Q outpaces the increase in V for small perturbations.

This inverse relationship between localization and Q can be, and has been, applied in other systems. The two-dimensional analogue was addressed in [91], and it was also applied to cavities in a photonic-crystal slab of holes rather than rods [83]. Alternatively, the state can be delocalized vertically using a low-contrast slab and operating above the light line. In that case, and a similar argument to the above has been made using the Green's function of the two-dimensional photonic crystal of infinite height (not the slab) [87]. (The relevant “band edge” there is of the mode propagating vertically along the 2d defect.) In such conventional structures as ring resonators, Q again grows with increasing size, in a phenomenon conceptually similar to that of the super-defect.

In principle, the same effect could occur at both band edges, but a number of things can prevent it from being observable, especially at the upper band edge. For example, there can be a finite threshold required in $\Delta\varepsilon$ before a mode is localized. Or, the upper boundary of the gap can be determined something other than the band edge of the perturbed mode (*e.g.*, an intersection with the light line).

6.4 Multipole Cancellation

There is also a conceptually different mechanism for minimizing radiation losses of a cavity, one which does not require delocalization of the state, and which peaks in the interior of the gap rather than at an edge. The idea is that if the field oscillates in sign, especially for a higher-order mode in which the differing signs have equal magnitude, it might be possible to force *cancellations* in the integral of Eq. (6.4). In particular, we study cancellations of terms in the multipole expansion of the field, using the vacuum Green's function. This \hat{G}_ω allows simple analytic study, possibly at the expense of worsened convergence—we shall see, however, that excellent convergence is still obtained with only a few multipole terms.

6.4.1 The Multipole Expansion for Radiation Fields

In the multipole method, one essentially expands Eq. (6.4) in terms of spherical harmonics Y_{lm} [95]. The radiated power is then an *incoherent* sum of the time-averaged power P_{lm} radiated by each multipole (proportional to the square of its multipole moment). These moments are a rapidly decreasing series, since high orders represent fast angular oscillations not present in a low-order cavity mode. Thus, if one can cancel the lowest multipole moment(s) without drastic changes in the localized field character, a large fraction of the radiated power will become zero, independent of the other moments. Such cancellations require only sign oscillations in the cavity field and degrees of freedom

to control their contributions to the moment integrals. In this way, one should be able to dramatically decrease losses without sacrificing localization.

The 3d multipole expansion is presented in [95], and we give here the similar but simpler 2d expansion for TM fields. One finds the far-field radiation to be:

$$E_z(r, \varphi) = \sum_{m=-\infty}^{\infty} a_m \frac{e^{im\varphi}}{\sqrt{2\pi}} H_m^{(1)}(kr), \quad (6.6)$$

where $H_m^{(1)}$ is the complex (outgoing) Bessel function. a_m is the multipole moment, determined purely from the near-field pattern:

$$a_m = \frac{i\pi k^2}{2} \int J_m(kr') \frac{e^{-im\varphi'}}{\sqrt{2\pi}} \Delta\varepsilon(\hat{x}') E_z(\hat{x}') d^2x', \quad (6.7)$$

where J_m is a Bessel function and $\Delta\varepsilon = \varepsilon - 1$. The time-averaged radiated power of each multipole is then $P_m = \frac{c}{4\pi^2 k} |a_m|^2$. Thus, any variation of $\Delta\varepsilon$ that can lead to Eq. (6.7) being zero for the lowest-order a_m will maximize Q , for which any number of root-finding procedures could be employed.

We can define a Q for *each* multipole moment by a ratio of the mode energy U to the radiation rate: $Q_m = ckU/P_m$, just as in Eq. (6.3). Because the P_m sum incoherently, the total Q is just the inverse sum of $1/Q_m$. Moreover, one can now compute Q from only the local cavity fields (raising the possibility of loss calculations using, *e.g.*, frequency-domain methods). Although the above relations are for complex fields, they can be used for definite-frequency real fields (*e.g.* simulation output) via: $\text{Im } \vec{E} = \vec{\nabla} \times \text{Re } \vec{H} / k\varepsilon$.

6.4.2 A Multipole Cancellation in 2d

In the 2d system (the 1d sequence of dielectric rods), we form a dipole-mode defect by increasing the radius of a single rod, pulling a single mode down into the gap [90]. The resulting Q vs. frequency for a range of radii is shown in Fig. 6.4, along with its Q_m decomposition. Q exhibits a sharp peak of almost 3×10^4 in the interior of the gap. In contrast, the delocalization mechanism for high- Q leads to a Q divergence towards a band

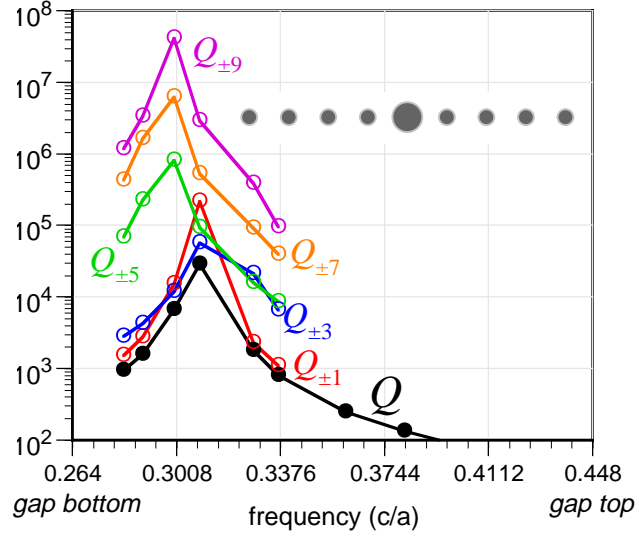


Figure 6.4: Q vs. ω for a dipole state in the 2d rod-sequence structure (inset) of Fig. 6.1(a). Both the total Q (black line, filled circles) and also the $Q_{\pm m}$ of the first few nonzero multipole moments (colored lines, hollow circles) are shown.

edge. To verify that the peak in Q comes from multipole cancellation, one need only look at the radiation pattern, shown in Fig. 6.5: at the peak Q , extra nodal lines appear, proving that the radiation pattern of Eq. (6.6) has transitioned to a higher order. Quantitatively, the computed Q_m are shown in Fig. 6.4, and the lowest multipole moment dominates everywhere *except* at the peak, where the next moment supercedes it. The expansion converges since the Q_m increase rapidly with m . (The close-set peaks in the higher moments may suggest that a more compact representation could be found, *e.g.*, using the crystal Green's function.) By symmetry, even multipole moments are zero and are not shown, and we have combined Q_m and the equal Q_{-m} into $Q_{\pm m} \equiv (Q_m^{-1} + Q_{-m}^{-1})^{-1}$. The total Q computed by combining these Q_m terms is within 4% of the Q measured from the field decay.

6.4.3 A Multipole Cancellation in 3d

We also present a three-dimensional state that demonstrates this effect in the photonic-crystal slab of rods. This (non-degenerate) mode has quadrupole symmetry, and is produced by increasing the radius of a rod to $0.45a$ and increasing its ϵ to 13. In order to achieve the multipole cancellation, an extra tuning parameter was required: we also

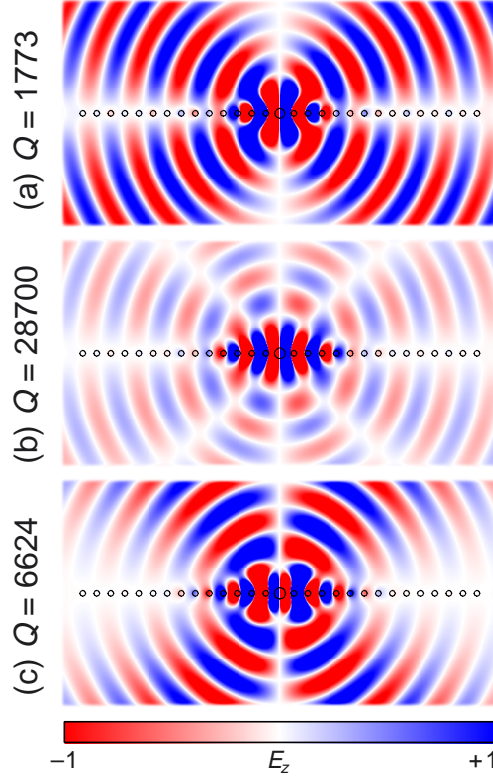


Figure 6.5: E_z radiation pattern for 2d dipole states of Fig. 6.4, using a color table that exaggerates small field magnitudes, with dielectric boundaries shown in black. (a) Point just before the peak ($Q=1773$, $\omega=0.328$). (b) Point at the peak ($Q=28700$, $\omega=0.309$), showing nodal lines from cancellation of the lowest multipole moment. (c) Point just beyond the peak ($Q=6624$, $\omega=0.300$).

increased the radii of the four adjacent rods to $0.25a$ and varied their dielectric constants to adjust the mode frequency, resulting in the Q shown in Fig. 6.6(a).^{*} Q and Q_{\perp} are essentially the same, due to the localization. Q peaks (at 1925) in the interior of the gap, so the underlying phenomenon must be different from the delocalization discussed previously. To verify that the peak is due to the postulated multipole cancellation, we plot in Fig. 6.6(b) the fields for ω below, at, and above the peak. Two clear nodal planes appear precisely at the peak, indicating the cancellation of the lowest-order multipole moment. (The near-field patterns are visually indistinguishable in the three cases.) In contrast, if the

^{*} The simultaneous variation of ϵ and radii was due to the limited computational resolution; in a real system, radii alone (or other geometric parameters) would be sufficient.

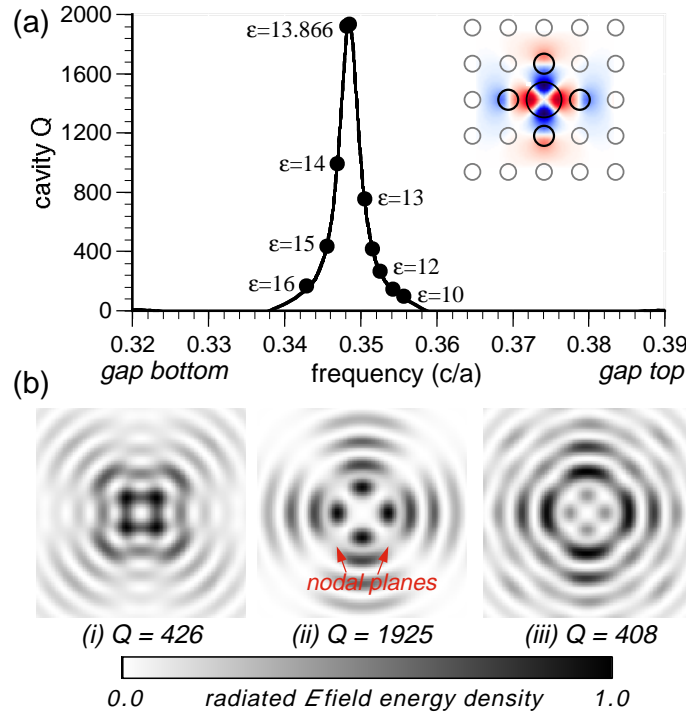


Figure 6.6: (a) Q vs. ω for a quadrupole state in the 3d rod-slab structure of Fig. 6.1(b). E_z in the mid-plane at peak Q is shown in the inset. The central rod is $r=0.45a$ with $\epsilon=13$; ϵ of the four $r=0.25a$ neighbors was varied to control ω , and labels the points. The solid line is a Lorentzian curve fitted to the peak (with $R^2 = 0.9994$). (b) Electric-field energy density for this mode, plotted in a plane $2a$ above the rods. (i) Point just before the peak ($Q=426$, $\omega=0.346$). (ii) Point at the peak ($Q=1925$, $\omega=0.349$), showing nodal planes from cancellation of the lowest multipole moment. (iii) Point just beyond the peak ($Q=408$, $\omega=0.352$).

dielectric constants or radii of all five defect rods are increased together, Q increases monotonically towards the lower band edge just as in Fig. 6.2, without nodal planes.

This particular cavity is a proof of concept, and we expect that significant refinements (*e.g.* increased Q) are possible. (In addition to the quadrupole mode at $0.349 c/a$, it also supports a higher-order monopole state at $0.358 c/a$ with a Q of 887.) A similar peaking of Q in the interior of the gap (with a peak value of more than 10^4) was reported in a photonic-crystal slab of holes, and we suspect that the explanation there must also be a far-field cancellation; this is under investigation.

6.5 Summary and Conclusions

Minimizing radiation losses in optical structures lacking a complete PBG is an important prerequisite for their use in efficient optical devices. We described and characterized two independent mechanisms for achieving high- Q resonant cavities, which are arguably the components most susceptible to radiative scattering. One mechanism, which had already been described in other contexts but not explicitly characterized as a function of frequency in three dimensions, is to delocalize the cavity mode. With only moderate delocalization, we demonstrate that $Q > 1000$ is achievable. In this case, Q characteristically diverges as a band edge is approached. A different and novel method is to engineer cancellations in the far-field radiation pattern, leading to a Q that peaks in the interior of the gap for strongly-localized states, again leading to $Q > 1000$.

Future work for the delocalization method will include a more complete analysis of why Q diverges towards some band edges and not others, and analytical predictions of the divergence scaling. The far-field cancellation mechanism also requires better understanding to predict under what circumstances and with what strength it will appear, and also to clarify its relationship to other interference phenomena such as coherent scattering. The Green's function formalism offers a promising approach for many of these issues. Additional problems to address include the effect of substrates on cavity Q . The main effects of a symmetric low-index substrate, for a high-contrast slab, are a slight shift in frequencies and an increase in the density of states in the light cone; thus, high- Q cavities are expected to persist. Preliminary results suggest that low-index symmetric substrates typically reduce Q by 30–50%. Asymmetric substrates, furthermore, break the distinction between even and odd modes that allows a band gap, and thus should reduce $Q_{||}$. Other questions include the losses from more complex devices that include waveguide-cavity junctions.

Chapter 7

Three-dimensionally periodic dielectric layered structure with omnidirectional photonic bandgap*

7.1 Introduction

Much research in recent years has been focused on photonic crystals—periodic dielectric (or metallic) structures with a photonic bandgap (PBG), a range of frequencies in which light is forbidden to propagate [98],[99]. Photonic crystals provide an unprecedented degree of control over light, introducing the possibility of many novel optical devices and effects [100]. The diamond, A7, and graphite geometries have been shown to be particularly promising for PBGs, and these have led to many interesting practical structures [101]–[110]. We present a design (depicted in Fig. 7.1) for a practical three-dimensional photonic crystal based on these geometries. Our design satisfies three desirable criteria: (i) it has a large, complete PBG; (ii) it is a stack of planar layers amenable to microfabrication; and (iii) each layer corresponds to one of the two basic, ubiquitous 2d photonic-crystal slab geometries: rods in air and air cylinders in dielectric. The third criterion distinguishes our structure from previous work, and the resulting high-symmetry cross sections should allow integrated optical networks to be formed by changing only a *single* layer. In this way, one can leverage the extensive analyses and results obtained for 2d slab structures [112], without incurring the intrinsic losses due to the lack of an omnidirec-

* This chapter is based on: S. G. Johnson and J. D. Joannopoulos, “Three-dimensionally periodic dielectric layered structure with omnidirectional photonic band gap,” *Appl. Phys. Lett.* **77**, 3490–3492 (2000).

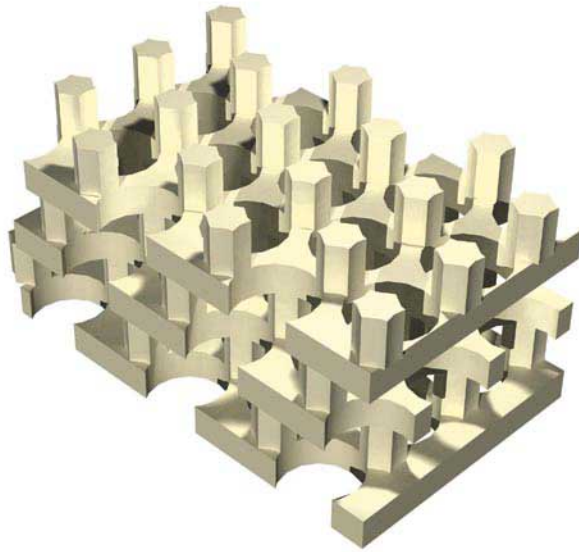


Figure 7.1: Computer rendering of a novel 3d photonic crystal, showing several horizontal periods and one vertical period, consisting of an fcc lattice of air holes (radius $0.293a$, height $0.93a$) in dielectric (21% fill). This structure has a 21% gap for a dielectric constant of 12.

tional gap.

The structure we propose is simply an fcc lattice (possibly distorted) of air (or low-index) cylinders in dielectric, oriented along the 111 direction. Such a structure, depicted in Fig. 7.1, results in a graphite-like [102] system of planar “slabs” (two-dimensionally periodic, finite-height) of two types: triangular lattices of air holes in dielectric and dielectric cylinders (“rods”) in air. The sculptured appearance of the rods is not important and is simply a by-product of the fabrication method (in Fig. 7.2). These slabs are stacked in a repeating, 3-layer sequence (along 111) and should be amenable to planar lithographic techniques as discussed below. The band diagram, in Fig. 7.3(a), has a complete gap of over 21% for Si:air dielectric contrast ($\epsilon = 12:1$ at $1.55\mu\text{m}$), and of over 8% even for Si:SiO₂ contrast ($\epsilon = 12:2$); the PBG persists down to ϵ contrasts of 4:1 (2:1 index contrast). The vertical transmission through roughly one period (three layers plus a hole slab) of the structure, shown in Fig. 7.3(b), is attenuated by about 20dB in the gap.

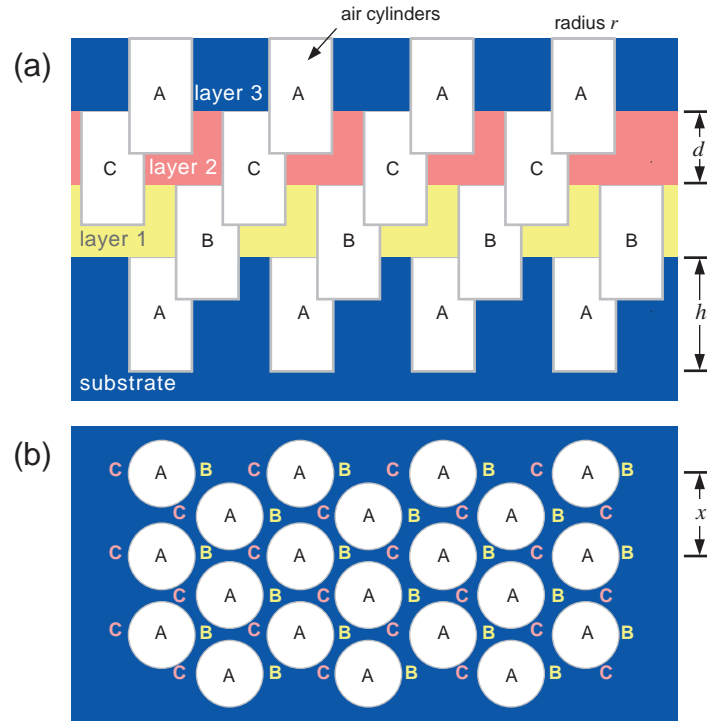


Figure 7.2: Schematic of the crystal in Fig. 7.1. (a) The vertical structure, showing the different material layers as they might be deposited during fabrication. (The layers are given different colors for clarity, but would normally be the same material and thickness.) (b) Plan view of a horizontal cross-section intersecting the “A” cylinders, with the offset locations of cylinders in other layers also labelled.

7.2 Other Photonic-Crystal Structures

Considered individually, the “photonic-crystal slab” layers that compose our structure have been the subject of much recent study [112], because they form a convenient and easily-fabricated approximation to two-dimensional photonic crystals. By themselves, however, they do not have a complete PBG and rely on index-guiding for vertical confinement, with consequential radiation losses whenever translational symmetry is broken (*e.g.*, at a waveguide bend or resonant cavity). Our structure overcomes this difficulty, and retains the advantage of a highly-symmetric cross-section—a “2d photonic crystal” in 3d. In the same system, one may then effectively work with either “rod” or “hole” 2d crystals, with the analogous characters of the modes and defects (*e.g.*, waveguides and cavities).

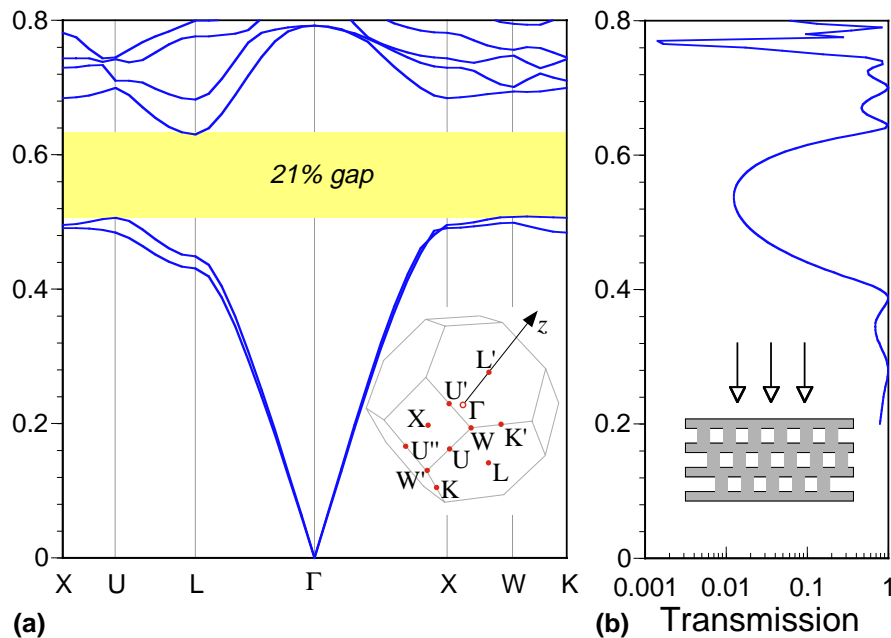


Figure 7.3: (a) Band diagram for the structure in Fig. 7.1, showing frequency in scale-invariant units vs. wavevector along important symmetry directions in the irreducible Brillouin zone. The inset shows the first Brillouin zone and its symmetry points. (b) Vertical transmission spectrum for slightly over one period (3 layers plus a capping hole slab, as shown in the inset) of this structure, showing the $\Gamma-L'$ gap.

The ideal photonic-crystal structure is one with an omnidirectional bandgap in three dimensions, and can thus be used to control light without losses regardless of propagation direction. A large gap (measured as a percentage of mid-gap frequency) is desirable because it displays the biggest optical effects, has the strongest confinement (allowing smaller crystals and devices), is robust in the presence of experimental disorder, and allows the widest bandwidth for optical devices. Most structures with complete PBGs are described by a diamond structure [101] or fcc lattice, or a distorted cousin such as an A7 or graphite structure [102], due to their “near-spherical” Brillouin zones. Although this determines the underlying topology and periodicity of the geometry, substantial innovations were required in order to design practically-realizable structures at micron length-scales. One class of systems, typified by the “inverse opal” [103]–[105] (an fcc lattice of

air spheres in dielectric), are constructed by self-assembly or similar methods—this can have unique advantages (*e.g.* inexpensive large-scale growth), but typically lacks the fine control provided by direct etching. Another PBG design, based on the diamond structure, is an fcc lattice oriented vertically in the 111 direction and formed by drilling a set of angled holes (along 110) into a bulk material [106]. A third class of designs, including ours, consists of planar layers with piecewise-constant cross-sections, which can be fabricated layer-by-layer—this provides the ability to place defects and devices in the lattice with very fine control. An often-fabricated planar-layer structure is the “layer-by-layer” [107] (or “woodpile” [108]) design: dielectric “logs” stacked in alternating perpendicular directions with a 4-layer period, forming an fcc crystal oriented in the 100 direction. Any given plane, however, does not have high rotational symmetry, meaning that complicated optical systems (with symmetric waveguides in various directions) must be formed by modifying at least two layers. A similar structure with more complicated logs (to more closely mimic a diamond structure) has also been proposed [109]. An entirely different planar-layer 3d PBG was designed using an fcc lattice stacked in the 110 direction [110], with a 23% gap for $\epsilon=12:1$ —however, it lacks planar rotational symmetry, prohibiting identical waveguides along different directions (desirable for integrated optics). There is one planar-layer PBG structure with a complete gap that does have highly-symmetric cross-sections—the simple-cubic “scaffold” lattice [111] with square rods along the edges of the cubes. (This is the exception to the rule of fcc-like lattices.) In a sense, this structure is similar to the one we propose, albeit in a simple-cubic rather than fcc lattice—it consists of alternating layers of thin slabs of square holes and thick slabs of square rods. The maximum gap achievable in the cubic scaffold for $\epsilon=12:1$ is 7%.

7.3 Characterizing the New Structure

In this section, we characterize the new structure by computing the size of its band gap, and so on, for various optimized parameters. The band diagram depicted in Fig. 7.3(a) is for air holes of radius $r=0.293a$ and height $h=0.93a$ (a is the fcc lattice constant) and a dielectric constant $\epsilon=12$, and has a 20.9% complete PBG. The bands were computed using preconditioned conjugate-gradient minimization of the Rayleigh quotient in a plane-wave basis [113]–[114]. (There are other symmetry points in the irreducible Brillouin zone of this structure that were calculated but are not shown in the band diagram, because their band edges do not determine the gap in this case.) The transmission spectrum of Fig. 7.3(b) is the result of a 3d finite-difference time-domain (FDTD) simulation [115] with a normal-incidence ($\Gamma-L'$) planewave source, absorbing boundaries above and below, and periodic boundaries at the sides.

More generally, one can apply a trigonal distortion to the fcc lattice of our structure to obtain a trigonal lattice without breaking any additional symmetry. The lattice vectors in this case become the three permutations of $\frac{a}{2}(1, 1, z)$ ($z = 0$ for fcc), with $d = \frac{a}{\sqrt{3}}\left|1 + \frac{z}{2}\right|$ and $x = \frac{a}{\sqrt{2}}|1 - z|$ as defined in Fig. 7.2. The parameter z can be varied to optimize the gap. The parameters of Fig. 7.3 were optimal for the fcc case of $z=0$; we also maximized the gap for varying z and dielectric contrast, with results in Table 7.1, “Optimal parameters and gaps for various dielectric contrasts,” on page 123. In general, the structure strongly prefers the fcc case with its nearly-spherical Brillouin zone, and distortions seem to increase the gap percentage by no more than 0.5.

Table 7.1: Optimal parameters and gaps for various dielectric contrasts

ϵ contrast	r	h	z	mid-gap v	gap size
12:1 (Si:air)	$0.285a$	$0.960a$	0.0165	$0.579 c/a$	21.4%
6:1	$0.273a$	$0.908a$	-0.00246	$0.648 c/a$	8.4%
12:2 (Si:SiO ₂)				$0.458 c/a$	
4.5:1	$0.265a$	$0.900a$	-0.00612	$0.688 c/a$	2.7%
4:1 (SiO ₂ :air)	$0.259a$	$0.895a$	-0.00778	$0.699 c/a$	0.33%

7.4 Possible Fabrication Methods

The fabrication of our structure might proceed along the following schematic lines, as depicted in Fig. 7.2. First, a layer of cylindrical holes (labelled “A”) are etched into a high-index substrate (*e.g.* Si). Next, the holes are back-filled with another material (*e.g.* SiO₂), the surface is planarized to the top of the high-index substrate, and a second layer of high-index material (“layer 1”) is grown on top. Then, the next layer of holes (“B”) is etched to the appropriate depth, but offset from the “A” holes as shown in the Fig. 7.2(b). This back-filling, planarizing, growth, and etching is repeated for the “C” holes and then for the next layer of “A” (in “layer 3”), at which point the structure repeats itself. When the desired layers are grown, the back-fill material is removed (*e.g.*, by a solvent), and a high-contrast photonic-crystal is obtained. A similar process has been successfully employed for the “layer-by-layer” structure, testifying to the feasibility of this method [116],[117]. Other variations are possible. One could leave the backfill material in the structure if it is low-index, and a complete PBG (albeit smaller, 8.4% for Si/SiO₂) can still be obtained; equivalently, a lower-index substrate could be used with air holes. Also, the layers could be fabricated individually, and then inverted and bonded together; this method has likewise been proven on the layer-by-layer structure [118],[119]. Alternatively, one may fabricate the rod and hole slabs using separate steps, requiring twice as many interlayer alignments but removing the need to etch two materials simultaneously. (It has recently come to our atten-

tion that a very similar structure was thus fabricated for $8\mu\text{m}$ wavelength, although the existence of a gap was not determined [120].) Our calculations show that the resulting extra degree of freedom, the rod radius, allows a maximum gap of over 26% for Si/air.

7.5 Summary

In summary, we have demonstrated a photonic crystal structure with a complete three-dimensional bandgap. This structure has a very large gap, is tolerant of low index contrast, is amenable to layer-by-layer fabrication, and can be thought of as a stack of 2d photonic-crystal slabs. This last feature, in the context of a large three-dimensional gap, permits simplified construction of complicated optical networks by modifying only a single layer, without breaking symmetry between different directions in the plane. The defect modes thus created are expected to have much of the simple, well-understood character of the modes in the analogous 2d photonic crystals/slabs (both rod and hole lattices)—thereby building on the large body of existing work and analyses for those basic systems, without the inherent problems of losses due to the lack of a complete PBG.

Chapter 8

Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis^{*}

8.1 Introduction

Optical systems have been the subject of enormous practical and theoretical interest in recent years, with a corresponding need for mathematical and computational tools. One fundamental approach in their analysis is eigenmode decomposition: the possible forms of electromagnetic propagation are expressed as a set of definite-frequency (time-harmonic) modes. In the absence of nonlinear effects, all optical phenomena can then be understood in terms of a superposition of these modes, and many forms of analytical study are possible once the modes are known. Of special interest are periodic (or translationally-symmetric) systems, such as photonic crystals (or waveguides), which give rise to many novel and interesting optical effects [121]. Another important basic system is that of resonant cavities, which confine light to a point-like region. There, the boundary conditions are, in principle, irrelevant if the mode is sufficiently confined—so they can be treated under the rubric of periodic structures as well via the “supercell” technique. In this chapter, we describe a fully-vectorial, three-dimensional method for computing general eigenmodes of arbitrary periodic dielectric systems, including anisotropy, based on the preconditioned block-iterative solution of Maxwell's equations in a planewave basis. A new effective dielectric tensor for anisotropic systems is introduced, and we also describe a technique for computing eigenvalues in the interior of the spectrum (*e.g.* defect modes) without computing the underlying bands. We present comparisons of different iterative solution

^{*} This chapter is based on: S. G. Johnson and J. D. Joannopoulos, “Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis,” *Optics Express* **8** (3), 173–190 (2001).

schemes, preconditioners, and other aspects of frequency-domain calculations. The result of this work is available as a free and flexible computer program downloadable from the Web [122].

There are a few common approaches to eigen-decomposition of electromagnetic systems. One, which we employ in this chapter, is to expand the fields as definite-frequency states in some truncation of a complete basis (*e.g.* planewaves with a finite cutoff) and then solve the resulting linear eigenproblem. Such methods have seen widespread use, with many variations distinguished by the critical choices of basis and eigensolver algorithm [123]–[137]. This “frequency-domain” method is discussed in further detail below. Another common technique involves the direct simulation of Maxwell’s equations over time on a discrete grid by finite-difference time-domain (FDTD) algorithms [138]—one Fourier-transforms the time-varying response of the system to some input, and the peaks in the resulting spectrum correspond to the eigenfrequencies [139]–[144]. (Care must be taken to ensure that the source is not accidentally near-orthogonal to an eigenmode.) This has the unique (and sometimes desirable) feature of finding the eigenfrequencies only—to compute the associated fields, one re-runs the simulation with a narrow-band filter for each state. (Time-domain calculations can also address problems of a dynamic nature, such as transmission processes, that are not so amenable to eigenmethods.) A third class of techniques are referred to as “transfer-matrix” methods: at a fixed frequency, one computes the transfer matrix relating field amplitudes at one end of a unit cell with those at the other end (via finite-difference, analytical, or other methods) [145]–[150]. This yields the transmission spectrum directly, and mode wavevectors via the eigenvalues of the matrix; in some sense, this is a hybrid of time- and frequency- domain. Transfer-matrix methods may be especially attractive when the structure is decomposable into a few more-easily solvable components, and also for other cases such as frequency-dependent dielectrics.

In any method, the computation is characterized by some number N of degrees of freedom (*e.g.*, the number of planewaves or grid points), and one might be interested to compare how the number of operations (the “complexity”) in each algorithm scales with N . Unfortunately, there is no simple answer. As we shall see below, the complexity in frequency-domain is $O(i_c p N \log N) + O(i_c p^2 N)$ for a planewave basis, where $p \ll N$ is the number of desired eigenmodes and i_c is the number of iterations for the eigensolver to converge. In time-domain, the complexity is $O(i_t N)$ to find the frequencies, and $O(i_t p N)$ to also solve for the fields of p modes, where i_t is the number of time steps. (Transfer-matrix methods have too many variations to consider here.) The difficulty in both cases comes from the number of iterations, which scales in different ways depending upon how the problem size is increased. In frequency-domain, i_c is hard to predict, but we shall show below that it often grows only very slowly with p and N , so we can treat it as approximately constant (often < 20). In time-domain, however, i_t must increase linearly with the spatial resolution (a certain kind of N increase) to maintain stability [138]. It also increases inversely with the required frequency resolution, by the uncertainty principle of the Fourier transform: $i_t \Delta t \Delta f \sim 1$ (where Δt is the timestep). Not only, then, is i_t a large number (typically > 1000), but it must also increase dramatically to resolve closely-spaced modes (although this can be ameliorated somewhat by sophisticated signal processing [151]); in contrast, frequency-domain methods have no special problem resolving even degenerate modes. One traditional advantage of time-domain has been its ability to extract modes in the interior of the spectrum without computing the lower-frequency states, but we will show that this is feasible in frequency-domain as well. We feel that both time- and frequency-domain methods remain useful tools to extract eigenmodes from many structures.

There is sometimes concern that discontinuities in the dielectric function may cause poor convergence in a planewave basis. As is described in Sec. 8.2.3, however, this can be

alleviated by the use of a smoothed effective dielectric tensor [125], and we demonstrate that convergence proportional to the square of the spatial resolution, Δx^2 , can be achieved even for sharply discontinuous anisotropic dielectric structures.

In the chapter that follows, we describe in greater detail our method for obtaining the eigenmodes of Maxwell's equations, dividing the discussion into two parts: Maxwell's equations and eigensolvers. First, we review how Maxwell's equations can be cast as an eigenproblem for the frequencies, discuss the choice of basis and the computation of an effective dielectric tensor, and consider the critical selection of an approximate preconditioner. Second, we describe various block-iterative algorithms for solving this eigensystem; in principle this is independent of the equations being solved, but in practice there are a number of specific considerations. Throughout, we illustrate the methods being compared with numerical results for example systems.

8.2 The Maxwell Eigenproblem

We first express Maxwell's equations as a linear eigenproblem, abstracting where possible from the differential equations in the individual field components to a higher-level view that better illuminates the overall process. To this end, we employ the Dirac notation of abstract operators \hat{A} and states $|H\rangle$ to provide a representation-independent expression for the fields and inner products, and later use ordinary matrix notation to indicate the transition to a finite problem. (The underlying equations remain fully vectorial.) This eigenproblem formulation has appeared in various forms elsewhere, and we begin by reviewing it here. The source-free Maxwell's equations for a linear dielectric $\epsilon = \epsilon(\hat{x})$ can be written in terms of only the magnetic field $|H\rangle$ [121]:

$$\hat{\nabla} \times \frac{1}{\epsilon} \hat{\nabla} \times |H\rangle = -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} |H\rangle, \quad (8.1)$$

$$\hat{\nabla} \cdot |H\rangle = 0. \quad (8.2)$$

We consider only states with definite frequency ω , *i.e.* a time-dependence $e^{-\omega t}$. Furthermore, we suppose that the system is periodic—in that case, Bloch’s theorem for periodic eigenproblems says that the states can be chosen to be of the form [152]:

$$|H\rangle = e^{i(\vec{k} \cdot \hat{x} - \omega t)} |H_{\vec{k}}\rangle, \quad (8.3)$$

where \vec{k} is the “Bloch wavevector” and $|H_{\vec{k}}\rangle$ is a periodic field (completely defined by its values in the unit cell). Eq. (8.1) then becomes the linear eigenproblem in the unit cell:

$$\hat{A}_{\vec{k}} |H_{\vec{k}}\rangle = \left(\frac{\omega}{c}\right)^2 |H_{\vec{k}}\rangle, \quad (8.4)$$

where $\hat{A}_{\vec{k}}$ is the positive semi-definite Hermitian operator:

$$\hat{A}_{\vec{k}} = (\nabla + i\vec{k}) \times \frac{1}{\epsilon} (\nabla + i\vec{k}) \times . \quad (8.5)$$

All the familiar theorems of Hermitian eigenproblems apply. Because $|H_{\vec{k}}\rangle$ has compact support, the solutions are a discrete sequence of eigenfrequencies $\omega_n(\vec{k})$ forming a continuous “band structure” (or “dispersion relation”) as a function of \vec{k} . These discrete bands (modes as a function of \vec{k}) provide a complete picture of all possible electromagnetic states of the system, but typically one is interested in only the lowest few. (For example, in a photonic crystal it is possible for there to be a band gap in the lower bands: a range of ω in which no states exist [121].) Furthermore, the modes at a given \vec{k} may be chosen to be orthonormal:

$$\langle H_{\vec{k}}^{(n)} | H_{\vec{k}}^{(m)} \rangle = \delta_{n,m}, \quad (8.6)$$

where $\delta_{n,m}$ is the Kronecker delta.

8.2.1 The choice of basis

In frequency-domain methods, Eq. (8.4) is transformed into a finite problem by expanding the states in some truncated (possibly vectorial) basis $\{|b_m\rangle\}$:

$$|H_{\vec{k}}\rangle = \sum_{m=1}^N h_m |b_m\rangle. \quad (8.7)$$

This expression becomes exact as the number N of basis functions goes to infinity, assuming a complete basis. One then has the ordinary generalized eigenproblem

$$Ah = \left(\frac{\omega}{c}\right)^2 Bh, \quad (8.8)$$

where h is a column vector of the basis coefficients h_m , and A and B are $N \times N$ matrices with entries $A_{nm} = \langle b_n | \hat{A}_k | b_m \rangle$ and $B_{nm} = \langle b_n | b_m \rangle$. It is important to note that Eq. (8.8) by itself is incomplete, however—the modes must also satisfy the “transversality” constraint of Eq. (8.2); zero-frequency spurious modes are otherwise introduced, as can be seen by taking the divergence of both sides of Eq. (8.1).

In principle, one could compute the entries of A and B and then use a standard matrix algorithm to solve Eq. (8.8) directly, and this method is sometimes employed [123],[124],[126],[127],[130]. Such a computation, however, requires $O(N^2)$ storage for the matrices and $O(N^3)$ work for the diagonalization, making it impractical for large problems. Fortunately, since only p bands are typically desired, for some $p \ll N$, iterative methods are available to compute the bands with only $O(pN)$ storage and roughly $O(p^2N)$ work; these methods are the subject of Sec. 8.3. The relevant property of iterative methods is this: they require only a fast, ideally $O(N)$, method to compute the products Ah and Bh , with no need to store the matrices explicitly.

The choice of basis functions $|b_m\rangle$, then, is determined by three factors. First, they should form an compact representation so that a reasonable N yields good accuracy. Second, a convenient and efficient method for computing Ah and Bh must be available. Third, they should be inherently transverse, or otherwise provide an inexpensive way to maintain the constraint of Eq. (8.2).

8.2.1.1 The planewave basis

We chose to use a planewave basis (for reasons described below) [123]–[127], in which $|b_m\rangle = e^{i\vec{G}_m \cdot \vec{x}}$ for some reciprocal-lattice vectors \vec{G}_m ; the truncation N is determined by

choosing a maximum cutoff for the magnitude of \vec{G}_m . Strictly speaking, a cutoff magnitude would result in a spherical volume of \vec{G} vectors, but we expand this into a parallelepiped volume so that the transformation between planewave and spatial representations takes the convenient form of a Discrete Fourier Transform (DFT). (Such an extension also removes an ambiguity of the order in which to invert and Fourier-transform ϵ [125].) The planewave set then has a duality with a spatial grid, which is often a more intuitive representation. In particular, suppose that the three primitive lattice vectors (the units of periodicity) are $\{\vec{R}_1, \vec{R}_2, \vec{R}_3\}$ and the primitive reciprocal-lattice vectors are $\{\vec{G}_1, \vec{G}_2, \vec{G}_3\}$, defined by $\vec{R}_i \cdot \vec{G}_j = 2\pi\delta_{i,j}$ [152]. Then the basis functions are $|b_{m_1, m_2, m_3}\rangle = \exp\left(\sum_j m_j \vec{G}_j \cdot \vec{x}\right)$ with $m_j = -\lceil N_j/2 \rceil + 1, \dots, \lfloor N_j/2 \rfloor$,^{*} $N = N_1 N_2 N_3$, and Eq. (8.7) for the spatial field becomes:

$$\vec{H}_{\vec{k}}\left(\sum_k n_k \vec{R}_k / N_k\right) = \sum_{\{m_j\}} \hat{h}_{\{m_j\}} \exp\left(i \sum_{j,k} m_j \vec{G}_j \cdot n_k \vec{R}_k / N_k\right) = \sum_{\{m_j\}} \hat{h}_{\{m_j\}} \exp\left(2\pi i \sum_j m_j n_j / N_j\right). \quad (8.9)$$

Here, $n_k = 0, \dots, N_k - 1$ describe spatial coordinates on an $N_1 \times N_2 \times N_3$ affine grid along the lattice directions. This is precisely a three-dimensional DFT, and can be computed by an efficient Fast Fourier Transform (FFT) algorithm [153] in $O(N \log N)$ time.

Thus, in a planewave representation, the product Ah from Eq. (8.8) can be computed in $O(N \log N)$ time by taking the curl in wavevector space (just the cross-product with $\vec{k} + \vec{G}_m$), computing the FFT, multiplying by $\tilde{\epsilon}^{-1}$, computing the inverse FFT, and taking the curl again [125]:

$$A_{nm} = -(\vec{k} + \vec{G}_n) \times \dots \text{IFFT} \dots \tilde{\epsilon}^{-1} \dots \text{FFT} \dots (\vec{k} + \vec{G}_m) \times \quad (8.10)$$

The determination of an effective inverse dielectric tensor, $\tilde{\epsilon}^{-1}$, is discussed in Sec. 8.2.3.

The matrix B is simply the identity, thanks to the orthonormality of the basis.

^{*} This is equivalent to $m_j = 0, \dots, N_j - 1$ for the DFT, in which m_j is interpreted modulo N_j , but choosing zero-centered wavevectors is important when taking derivatives of the basis.

Since the basis functions themselves are scalars in this case, the amplitudes \vec{h}_m in Eq. (8.7) must be vectors. In addition to Eq. (8.8), the field must satisfy the transversality constraint, and it is here that a key advantage of the planewave basis becomes apparent: Eq. (8.2) becomes merely a local constraint on the amplitudes, $\vec{h}_m \cdot (\vec{k} + \vec{G}_m) = 0$. For each reciprocal vector \vec{G}_m one chooses a pair $\{\hat{u}_m, \hat{v}_m\}$ of orthonormal unit vectors that are perpendicular to $\vec{k} + \vec{G}_m$, and writes the amplitude as $\vec{h}_m = h_m^{(1)}\hat{u}_m + h_m^{(2)}\hat{v}_m$. Then the basis is intrinsically transverse, and one can treat Eq. (8.8) as an ordinary eigenproblem of rank $n = 2N$ without worrying about any constraint.

8.2.1.2 Other possible bases

The planewave basis has at least two potential disadvantages: first, it corresponds to a uniform spatial grid, and may thus be a less economical representation than one based on *e.g.* a general mesh; second, the computation of the Maxwell operator A requires $O(N \log N)$ time instead of $O(N)$, although the difference may be small in practice. Both of these problems could be overcome by using a different basis—for example, a traditional finite-element basis formed of localized functions on an unstructured mesh. Such a basis, however, would make it more difficult to maintain the transversality constraint, which is why we eschew it in our implementation. (One way around this might be to replace the magnetic field with the vector potential, $\vec{H} = \vec{\nabla} \times \vec{A}$, although this introduces higher-order derivatives into the eigenproblem.) Alternatively, it is possible to solve the eigenproblem without transversality and a posteriori identify and remove the resulting spurious modes (which lie at zero frequency unless a “penalty” term is added to the eigen-equation) [128],[134].

In two dimensions, though, transversality ceases to be a problem: one simply chooses the magnetic field along z (TE fields) or the electric field along z (TM fields, for which the eigenproblem could be recast in \vec{E} or \vec{D}). This fact has been employed by various

researchers to implement finite-element or other-basis frequency-domain methods in 2D [128]–[134]. (Our implementation also supports the 2D TE/TM case, of course.)

Given the eigenmodes for the primitive cell of a lattice, it may be possible to use them to construct a localized Wannier-function representation that is a useful basis for defect-mode calculations [135]–[136], although the non-uniqueness of the Wannier functions must be resolved, *e.g.*, by fitting to the precomputed band structure in “tight-binding” fashion [136]. Such a basis is automatically divergenceless since the constituent eigenmodes are transverse. Another possibility is a tight-binding basis that is not specified explicitly, but whose matrix elements are fitted to an existing band diagram [137].

8.2.2 Inversion symmetry

In general, the basis coefficients h are complex, but additional simplifications are possible when the dielectric function possesses inversion symmetry: $\epsilon(-\hat{x}) = \epsilon(\hat{x})$. The Fourier transform of a real and even function is real and even, so it follows that the planewave representation of $\hat{A}_{\hat{k}}$, Eq. (8.10), is then a real-symmetric matrix. This means that the plane-wave amplitudes \hat{h}_m can be chosen to be purely real, resulting in a factor of two savings in storage, more than a factor of two in time (due to more-efficient FFTs and matrix operations for real data), and possibly a reduction in the number of eigensolver iterations (due to the reduced degrees of freedom). The spatial fields, in contrast, cannot generally be chosen as real—rather, with inversion symmetry they may be chosen to satisfy the property of the Fourier transform of a real function: $\vec{H}_{\hat{k}}(\hat{x}) = \vec{H}_{\hat{k}}(-\hat{x})^*$. Because inversion symmetry is extremely common in practical structures of interest, our implementation supports this optimization when the symmetry is present. For generality, however, we also handle the case of complex h for non-symmetric systems.

When ϵ is a real-symmetric tensor, corresponding to an anisotropic dielectric material, the definition and consequences of inversion symmetry are the same as above. When ϵ is a

complex-Hermitian tensor, however, corresponding to a magnetic material, however, the definition of “inversion symmetry” for the purpose of the computational savings above must be generalized slightly: $\epsilon(-\vec{x}) = \epsilon(\vec{x})^*$. With this definition, even with a complex ϵ , its Fourier transform remains real and even, and so the consequence of real \vec{h}_m continues to hold.

8.2.3 The effective dielectric tensor

When computing the product $\hat{A}_{\vec{k}}|H_{\vec{k}}\rangle$ in a planewave basis, the multiplication by ϵ^{-1} is done in the spatial domain after a Fourier transform, so one might simply use the inverse of the actual dielectric constant at that point. Unfortunately, this can lead to suboptimal convergence of the frequencies as a function of N , due to the problems of representing discontinuities in a Fourier basis. It has been shown, however, that using a smoothed, effective dielectric tensor near dielectric interfaces can circumvent these problems, and achieve accurate results for moderate N [125]. In particular, near a dielectric interface, one must average the dielectric in two different ways according to effective-medium theory, depending upon the polarization of the incident light relative to the surface unit-normal vector \hat{n} . For an electric field $\vec{E} \parallel \hat{n}$, one averages the inverse of ϵ ; for $\vec{E} \perp \hat{n}$, one takes the inverse of the average of ϵ . This results in an effective inverse dielectric tensor $\tilde{\epsilon}^{-1}$:

$$\tilde{\epsilon}^{-1} = \overline{\epsilon^{-1}}P + \bar{\epsilon}^{-1}(1 - P), \quad (8.11)$$

where P is the projection matrix onto \hat{n} : $P_{ij} = n_i n_j$. Here, the averaging is done over one voxel (cubic grid unit) around the given spatial point; if ϵ is constant, Eq. (8.11) reduces simply to ϵ^{-1} . (The original formulation was in terms of $\tilde{\epsilon}$, but is equivalent.)

We have generalized this procedure to handle the case of anisotropic (birefringent) dielectric materials, in which case the intrinsic ϵ is already a real-symmetric tensor. Or, more generally, for magnetic materials ϵ may be complex-Hermitian. In this case, an analogous equation is needed that: (i) produces a Hermitian effective inverse dielectric, (ii)

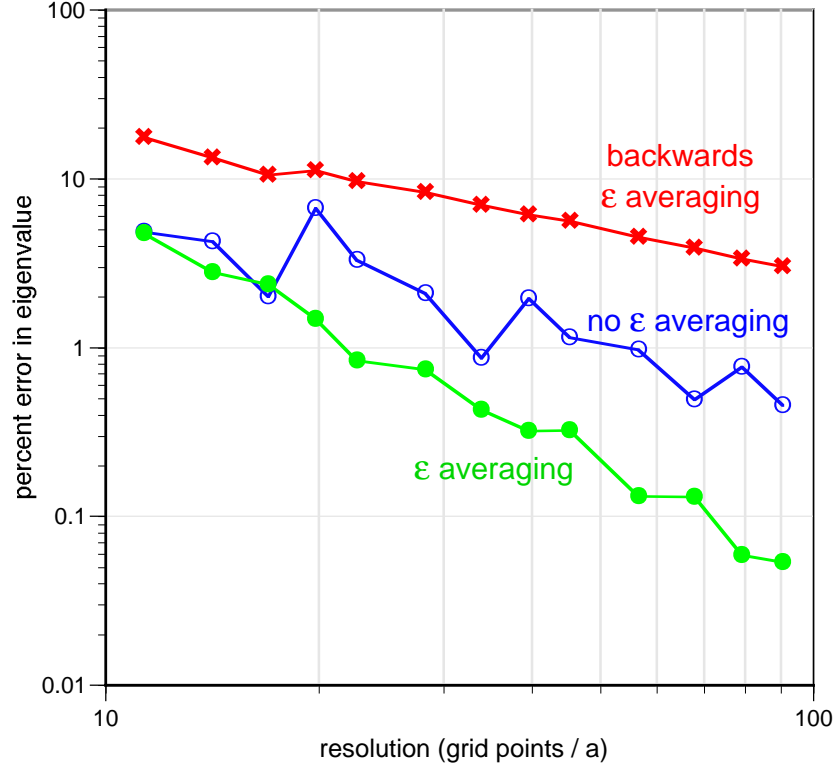


Figure 8.1: Eigenvalue convergence as a function of grid resolution (grid points per lattice constant a) for three different methods of determining an effective dielectric tensor at each point: no averaging, simply taking the dielectric constant at each grid point; averaging, the smoothed effective dielectric tensor of Eq. (8.12); and backwards averaging, the same smoothed dielectric but with the averaging methods of the two polarizations reversed.)

reduces to Eq. (8.11) for the case of scalar ϵ , (iii) yields ϵ^{-1} for constant ϵ , (iv) retains the physical justification of the different averaging methods for the different polarizations, and (v) similarly improves convergence. The expression we propose that satisfies these criteria is:

$$\tilde{\epsilon}^{-1} = \overline{\epsilon}^{-1}P + \bar{\epsilon}^{-1}(1 - P) = \frac{1}{2}(\{\overline{\epsilon}^{-1}, P\} + \{\bar{\epsilon}^{-1}, 1 - P\}), \quad (8.12)$$

where ϵ may be a tensor and $\{a, b\}$ denotes the anti-commutator $ab + ba$. Our first three conditions are manifestly satisfied. A physical justification of this formula is that a given averaging method should be used when either the field or the inverse ϵ times the field is in the appropriate direction relative to \hat{n} , hence the anti-commutators. To illustrate the con-

vergence impact of Eq. (8.12), we consider a simple example case similar to that of [127]: an fcc lattice (lattice constant a) of close-packed spherical holes in dielectric ($\epsilon = 12$), where the holes are filled by an anisotropic “liquid crystal” material with an ϵ of 2 for fields along an “extraordinary” 011 (\hat{x}) direction and 1 otherwise. We compute the frequency of the ninth band (just below the gap for air holes) at the L point as a function of grid resolution for three cases: no ϵ averaging, with averaging as in Eq. (8.12), and with averaging backwards from Eq. (8.12) (P switched with $1 - P$). The results, shown in Fig. 8.1, exhibit a significant acceleration of convergence by the averaging; conversely, the poor convergence of the backwards averaging demonstrates the importance of polarization for the smoothing method. With the averaging of Eq. (8.12), we see that the error decreases with the square of the resolution, just as for standard FDTD methods [138].

As a practical matter, it can be cumbersome to compute (or even to define) the surface normal \hat{n} for complicated three-dimensional structures, so we implement an approximation. Given a flat dielectric interface, the normal direction is exactly $\int \hat{r} \epsilon$ over a spherical surface intersecting the interface, where \hat{r} is the vector from the center of the sphere. This procedure also yields the correct normal for spherical and cylindrical interfaces, by symmetry. Therefore, we use this spherical average to define the normal direction in all cases.* In order to compute the average, we employ a 50-point 11th-degree spherical-quadrature formula [154] to numerically integrate $\hat{r} \epsilon$ over a spherical surface inscribed within the ϵ averaging voxel. (Testing this method by computing the normal vectors to a large number of random planar surfaces, we found a mean angular error of about 5° .)

* This method for defining \hat{n} can produce suboptimal results when the averaging voxel straddles two near-parallel dielectric interfaces. Preliminary investigations show that gains of a factor of two or more in eigenvalue accuracy are sometimes possible if a better approximation for \hat{n} is used in such cases.

An interesting unanswered question is whether a similar effective ϵ tensor would improve the convergence of other methods in numerical electromagnetism based on discrete grids or meshes, such as FDTD or BPM (Beam Propagation Method) algorithms. Scalar averaging methods have already been explored to improve modeling of discontinuous material interfaces in FDTD methods [155]–[156]. We have made preliminary investigations of the use of Eq. (8.12) for full-vector and wide-angle BPM [157] algorithms, comparing to the case where the smoothing is disabled, and have seen dramatic reductions in the numerical divergences that can otherwise plague high index-contrast systems with BPM.

8.2.4 Preconditioners

A critical factor in the performance of an iterative eigensolver is the choice of “preconditioning” operator. As will be explained in Sec. 8.3, our preconditioner requires us to supply an approximate inverse \tilde{A}^{-1} of A such that $\tilde{A}^{-1}h$ can be computed quickly. This choice of a “good” preconditioner is highly problem-dependent, and is unfortunately a matter of trial and error. We consider two possible preconditioners. The first is a diagonal preconditioner [125],[132],[158], inspired by the “kinetic energy” preconditioners often used in electronic calculations [159]. In this case, \tilde{A} is simply the diagonal elements of A :

$$\tilde{A}_{nm} = |\vec{k} + \vec{G}_m|^2 \delta_{n,m}, \quad (8.13)$$

where we have dropped an overall scaling factor (irrelevant to a preconditioner). The motivation for this approximation is that, for large \vec{G} components, the operator $\hat{A}_{\vec{k}}$ is dominated by the curl operations and not by the variations in ϵ . Computing $\tilde{A}^{-1}h$ is then an $O(N)$ diagonal operation.

We also consider a more accurate inverse [158]. Ideally, one would like to simply reverse each of the steps in computing Ah via the planewave representation of Eq. (8.10) to find an exact inverse. Note that a cross product of perpendicular vectors is invertible.

So, the only operation of Eq. (8.10) that is not trivially reversible is the final cross product with $\vec{k} + \vec{G}_n$, since $-i\omega\vec{E} = \varepsilon^{-1}\vec{\nabla} \times \vec{H}$ is not generally transverse (divergenceless). Since ε is normally piecewise constant and $\vec{\nabla} \cdot \varepsilon\vec{E}$, however, it is plausible that \vec{E} is “mostly” divergenceless. With this in mind, we approximate the operator $\hat{A}_{\vec{k}}$ by inserting projection operators \hat{P}_T onto the transverse field components:

$$\tilde{A} = \vec{\nabla} \times \hat{P}_T \frac{1}{\varepsilon} \hat{P}_T \vec{\nabla} \times , \quad (8.14)$$

where the rightmost \hat{P}_T is superfluous and serves only to make the operator clearly Hermitian. Both curls are now reversible, since they act on transverse planewaves. The inverse of Eq. (8.14) can be thus applied in $O(N \log N)$ time in a manner exactly analogous to Eq. (8.10): invert the $\vec{k} + \vec{G}_n$ cross product (projecting the result), FFT, multiply by ε , inverse FFT, and invert the $\vec{k} + \vec{G}_m$ cross product. This preconditioner is more expensive to compute than that of Eq. (8.13), but we shall show that it yields a substantial speedup in the eigensolver.

8.2.4.1 Removing the singularity

There is one evident obstacle in this preconditioning process—at $\vec{k} = 0$ (the “ Γ ” point), the operator $\hat{A}_{\vec{k}}$ is singular. (This will also be a problem for iterative methods such as conjugate-gradient, which are known to converge poorly for ill-conditioned matrices [167].) This difficulty is easily overcome, however, because the singular solutions are known analytically: they are the constant-field solutions at $\omega = 0$. So, at Γ we simply remove these singular solutions from the basis space (only considering planewaves with $\vec{G}_m \neq 0$), and re-insert them after we have solved for the non-singular eigenvectors.

8.3 Iterative Eigensolvers

Iterative eigensolvers are fast methods for computing the $p \ll n$ smallest (or largest, or extremal) eigenvalues and eigenvectors of an $n \times n$ generalized eigenproblem $Ay = \lambda By$, by iteratively improving guesses for the eigenvectors until convergence is

achieved. (B is the identity in our case, but we consider the full problem here for generality.) They are thus ideally suited for the problem of finding the few lowest eigenstates of Maxwell’s equations. Many iterative eigensolver algorithms have been developed, but we focus our investigations on two in particular: preconditioned conjugate-gradient minimization of the block Rayleigh quotient [125],[159]–[161], and the Davidson method [162]–[164] (an extension of Lanczos’s algorithm [165]). We choose these two because they are able to take advantage of the good preconditioners available for this problem, and because they are Krylov-subspace methods—at each step, they compute the best solution in the subspace of *all* previously tried directions (with some approximations, described below) [160],[166].

As a test case for the convergence of the algorithms discussed below, we use a 3d diamond lattice of dielectric spheres ($\epsilon = 12$) in air, which has a gap between its second and third bands [123]. Except where otherwise noted, we solve for the first five bands at the L wavevector point with a $16 \times 16 \times 16$ grid resolution in the (affine) primitive cell (~ 22.6 grid points per lattice constant), tracking the error in the eigenvalue trace versus the converged value, with the non-diagonal preconditioner of Eq. (8.14). A set of five pseudo-random fields are used as the starting point for the eigensolver, and we report the results from the case that converges in the median number of iterations. Note that the number of iterations in practice will often be less than that reported here—not only is the minimum solution tolerance usually bigger, but one typically starts with the fields from a nearby \vec{k} point (yielding a significant “head start” vs. random fields).

8.3.1 Conjugate-gradient minimization of the Rayleigh quotient

The smallest eigenvalue λ_0 and the corresponding eigenvector y_0 of a Hermitian matrix A are well-known to satisfy a variational problem—they minimize the expression:

$$\lambda_0 = \min \frac{y_0^\dagger A y_0}{y_0^\dagger B y_0} \quad (8.15)$$

known as the “Rayleigh quotient,” where \dagger denotes the Hermitian adjoint (conjugate-transpose). One can then compute this eigenpair by performing an unconstrained minimization of the Rayleigh quotient, using a method such as conjugate-gradient [167]. To find the subsequent eigenvalue and eigenvector, the minimization is repeated while maintaining orthogonality to y_0 (through B), and so on; a process known as “deflation.” As in all iterative methods, the matrix A need never be computed explicitly, only the product Ay (and By). This method [160] for computing the smallest eigenpairs of a matrix, also called a “Rayleigh-Ritz” algorithm, was employed by [125] to solve for the eigenstates of Maxwell’s equations in a planewave basis.

8.3.1.1 *The block Rayleigh quotient*

We use an extension of the conjugate-gradient Rayleigh-quotient method, adapted from [161], that solves for all of the eigenvectors at once instead of one-by-one with deflation. Such a process is called a “block” algorithm, and it has two advantages in our case: first, it can take advantage of efficient block matrix operations that have superior cache utilization on modern computers [168]–[169]; second, on parallel machines, block algorithms can communicate in larger chunks that mask latencies. Actually, as described in Sec. 8.3.4, we employ a hybrid of block algorithms and deflation, but such a generalization can only be made if one has a block method to begin with.

Let Y be an $n \times p$ matrix whose columns are the p eigenvectors with the smallest eigenvalues. Then Y minimizes the trace $\text{tr}[Y^\dagger A Y]$ subject to the orthonormality constraint $Y^\dagger B Y = I$ (where I is the identity matrix) [171]. Although it is possible to perform such a minimization directly by observing the differential geometry of the constraint man-

ifold [170], it is more convenient to introduce a rescaling, similar to that of the Rayleigh quotient, that makes the problem unconstrained.* If we write:

$$Y = Z(Z^\dagger BZ)^{-1/2} \quad (8.16)$$

for an arbitrary non-singular $n \times p$ matrix Z , then the orthonormality constraint is automatically satisfied—this is known as the “polar” or “symmetric” orthogonalization of Z , and happens to be optimal from a certain numerical-stability standpoint [172]. More importantly, we can now solve for the eigenvectors by performing an unconstrained minimization of the block Rayleigh quotient:

$$\text{tr}[Z^\dagger AZU], \quad (8.17)$$

where $U = (Z^\dagger BZ)^{-1}$. Once the minimization is completed, the eigenvectors are a superposition of the columns of Z , and can be found by diagonalizing the $p \times p$ matrix $Y^\dagger AY$ where Y is given by Eq. (8.16).

8.3.1.2 Conjugate gradient

The conjugate-gradient algorithm is a method for minimizing multidimensional functions by searching along successive “conjugate” directions, chosen so that each line minimization does not spoil the previous ones [167]. Strictly speaking, the algorithm is designed only for quadratic functions, for which it is a Krylov-subspace method. Even non-quadratic forms such as Eq. (8.17), however, are approximately quadratic near their minima and so conjugate-gradient is beneficial. The minimization direction is usually chosen starting from the steepest-descent (gradient) direction G , which in this case is given by [161]:

$$G = P_\perp AZU, \quad (8.18)$$

where P_\perp is the projection operator onto the space orthonormal to Z : $P_\perp = 1 - BZUZ^\dagger$.

The conjugate minimization direction D is then:

* We also tried using the differential-geometry conjugate-gradient algorithm of [170]. In addition to requiring more matrix operations per iteration than the method described here, it also seemed to work much more poorly with our choice of preconditioners, for unknown reasons.

$$D = \hat{K}G + \gamma D_0, \quad (8.19)$$

where \hat{K} is a preconditioning operation (discussed below), D_0 is the minimization direction from the *previous step*, and γ is either:

$$\gamma = \frac{\text{tr}[G^\dagger \hat{K}G]}{\text{tr}[G_0^\dagger \hat{K}G_0]} \quad (8.20)$$

in the Fletcher-Reeves variant or:

$$\gamma = \frac{\text{tr}[(G - G_0)^\dagger \hat{K}G]}{\text{tr}[G_0^\dagger \hat{K}G_0]} \quad (8.21)$$

in the Polak-Ribiere variant, where G_0 is the gradient direction from the previous step. Both variants involve approximations for the derivative of G [170], with Polak-Ribiere being slightly more accurate at the expense of requiring extra storage space for G_0 . If we instead choose $\gamma = 0$, the result is simply the preconditioned steepest-descent algorithm. The relative convergence of these variations is depicted in Fig. 8.2, and displays a sharp acceleration for conjugate-gradient once the error becomes small, presumably due to the function becoming approximately quadratic. (Our subsequent convergence plots use Polak-Ribiere.)

Once the conjugate direction is determined, one needs to minimize Eq. (8.17) over $Z' = Z + \alpha D$. By substituting Z' into Eq. (8.17), one finds a function in terms of α and constant $p \times p$ matrices, which we then minimize over α via a one-dimensional optimization algorithm specifically designed for use with conjugate-gradient-like methods [173]. Alternatively, one could make a two-point approximation for the second derivative of the function along the line and fit to a quadratic (*i.e.*, apply one step of Newton's method) [159],[161]. This method requires two fewer $O(p^2n)$ matrix multiplications (out of eight per iteration), but often produces somewhat slower and less reliable convergence.

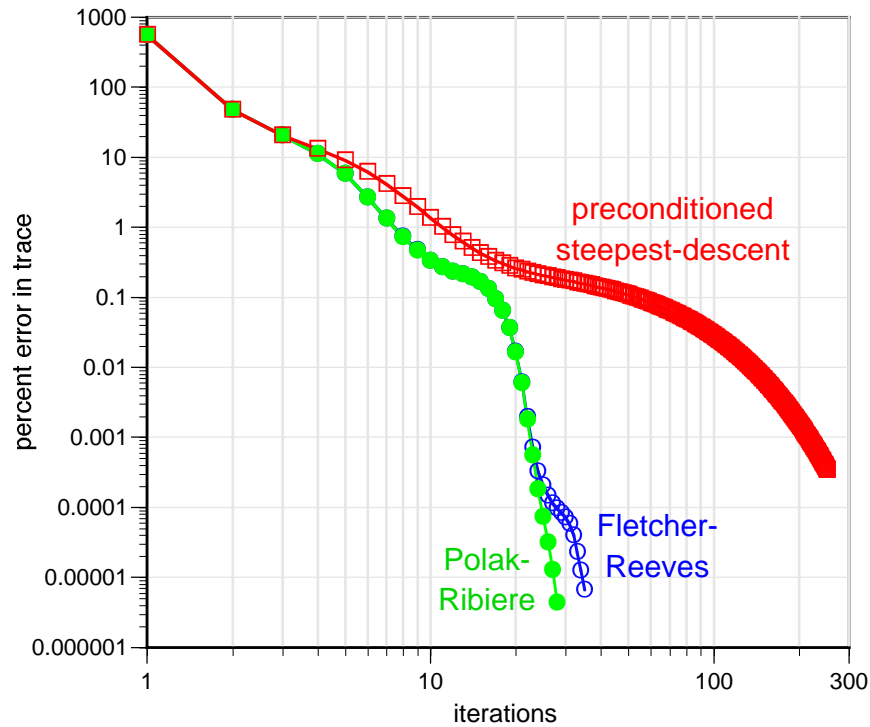


Figure 8.2: Eigensolver convergence for two variants of conjugate gradient, Fletcher-Reeves and Polak-Ribiere, along with preconditioned steepest-descent for comparison.

8.3.1.3 Preconditioning

Large gains in the rate of convergence can be achieved by choosing a proper Hermitian preconditioning operator \hat{K} in Eq. (8.19)—ideally, an approximate inverse of the Hessian (second-derivative) matrix of the objective functional, Eq. (8.17).^{*} One can think of this as an application of the multi-dimensional Newton’s method to find the zero of the gradient G : the Newton update $\hat{K}G$ of the solution guess Z is the function G divided by its derivative. Equivalently, if we are at

$$Z^* + \delta Z, \tag{8.22}$$

where Z^* is the unknown minimum, we wish to solve for $\delta Z \cong \hat{K}G$ to lowest order. Thus, we substitute Eq. (8.22) into Eq. (8.18) for the gradient and expand to first order in δZ , noting that $G^* = 0$, and find [174]:

^{*} In the literature, “preconditioner” sometimes refers instead to \hat{K}^{-1} , the approximate Hessian.

$$G \cong P_{\perp}(A\delta Z - B\delta ZUZ^{\dagger}AZ)U, \quad (8.23)$$

Suppose that Z were rotated to diagonalize the generalized eigenproblem $Z^{\dagger}AZx = Z^{\dagger}BZx\tilde{\lambda}$. Then, the second term in Eq. (8.23) would be simply $B\delta Z$ times the current eigenvalue approximations. Inverting Eq. (8.23) thus involves the inversion of $A - B\tilde{\lambda}$. At this point, however, we make an additional approximation: since the desired eigenvalues λ are small, we neglect that term and use:

$$G \cong P_{\perp}A\delta ZU. \quad (8.24)$$

An inverse of this equation is given by the Hermitian operation $\delta Z = A^{-1}GU^{-1}$, which is then further approximated by:

$$\delta Z \cong \hat{K}G = \tilde{A}^{-1}GU^{-1}, \quad (8.25)$$

where we have used an approximate inverse for A such as those in Sec. 8.2.4. The effects of the preconditioner of Eq. (8.25) with our two approximate inverses are demonstrated in Fig. 8.3, showing significant benefit from the non-diagonal preconditioner.

Additional refinements are possible that we do not consider here. One could solve Eq. (8.23) exactly by an iterative method, or at least a few iterations thereof to improve the preconditioner. Also, because of the projection operator P_{\perp} , one has a choice of how to invert Eq. (8.23) and Eq. (8.24)—it has been suggested in a similar context that one should choose δZ in the space orthogonal to Z [164]. This can be done at the expense of a few extra matrix operations, but it did not seem to improve convergence significantly in our case. Such a choice would become more important, however, if one attempted a more accurate preconditioner that inverts expressions of the form $A - B\lambda$, lest singularities arise.

8.3.2 The Davidson method

In the Davidson method, instead of iteratively improving a single (block) eigenvector approximation Y , one builds up an increasing subspace V that eventually contains the

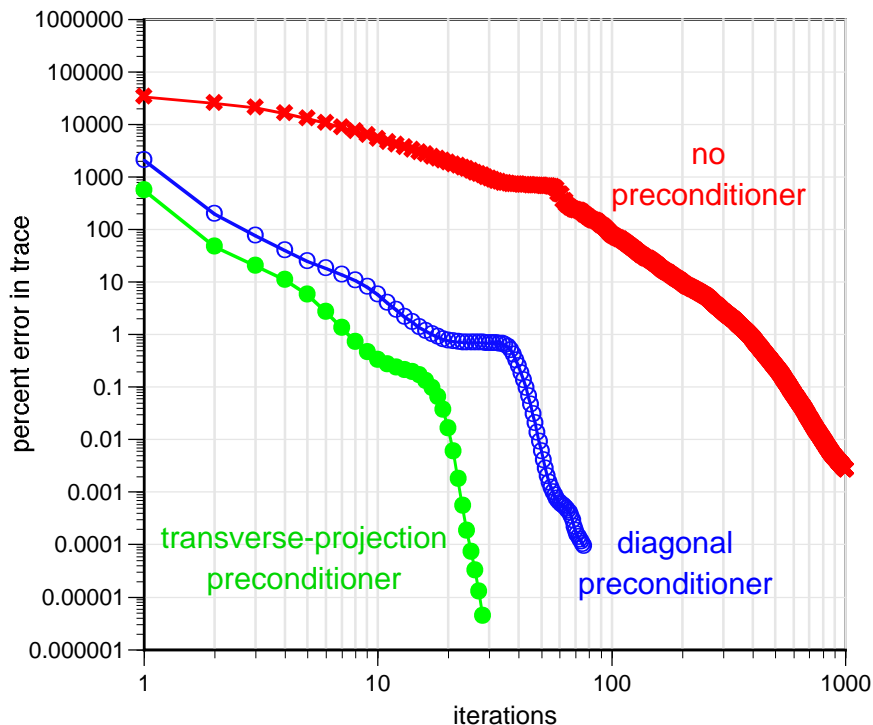


Figure 8.3: The effect of two preconditioning schemes from section 2.4, diagonal and transverse-projection (non-diagonal), on the conjugate-gradient method.

desired eigenvectors [163]. While both the Davidson and conjugate-gradient methods are Krylov-subspace algorithms, this is only exactly true of conjugate-gradient for pure quadratic forms, and of Davidson in the absence of restarting (described below). We summarize the method briefly in the following.

At each iteration, V is an $n \times q$ matrix for some $q \geq p$, whose columns span the current subspace, with $V^\dagger BV = I$. The task is to find p new vectors to add to V so that it spans a better approximation for the desired p eigenvectors. To do this, one first computes the best p eigenvectors Y so far: A is projected to the V subspace, $A_v = V^\dagger AV$, and the associated eigenvectors are computed for this $q \times q$ matrix; of these eigenvectors, one chooses the block of the p smallest eigenvalues, $L = \text{diag}(\lambda_1, \dots, \lambda_p)$, and their associated $q \times p$ eigenvectors Y_v , and computes the “Ritz” eigenvectors $Y = VY_v$. The residual (*i.e.*, error) for this eigenvalue approximation is $R = BYL - AY$. The new directions

to be added to V are then given by $D = \hat{K}R$ (orthogonalized, to maintain the orthonormality of V), where \hat{K} is a preconditioning operator.

To keep the subspace limited to a reasonable size, this process is usually restarted every few iterations by setting $V = Y$, with some tradeoff in speed of convergence. (Alternatively, one could keep the r lowest eigenvectors for some $r > p$ [164].) Restarting in this way on *every* iteration (never increasing the subspace) is a form of steepest-descent; a method of this sort was used in [132]. The ideal preconditioner \hat{K} for the Davidson method, remarkably, has been shown to be the same as that of Eq. (8.23) for trace-minimization, combined with the fact that Y in this case is an orthonormal solution to the projected eigenproblem. That is, $\hat{K}R$ should be an approximate solution for D in [164]:

$$R = P_{\perp}(AD - BDL), \quad (8.26)$$

where L is the diagonal matrix of eigenvalues from above and P_{\perp} is the projection matrix $1 - BYY^{\dagger}$ as in Sec. 8.3.1.2. Again, D should ideally be found in the space orthonormal to Y [164]; this is the ‘‘Jacobi-Davidson’’ algorithm. To solve Eq. (8.26), we make the same approximations as in Sec. 8.3.1.3 and use $\hat{K}R = \tilde{A}^{-1}R$. The Davidson method is also adaptable to non-Hermitian operators, and it has thus been employed to treat electromagnetic problems containing absorption [134].

The results of a comparison of the Davidson method described here with the block conjugate-gradient minimization of the Rayleigh quotient from the previous section are shown in Fig. 8.4. Various restarting frequencies are used, with more infrequent restarts resulting in faster convergence but also higher memory requirements and a greater computational cost per iteration (*e.g.* more than twice as much memory, and 1.6 times as many floating-point operations for eigenvector matrix products, in our period-5 restarted Davidson compared to conjugate-gradient). Overall, it does not seem to match the performance of the conjugate-gradient method in our case. Nevertheless, we believe it is important that

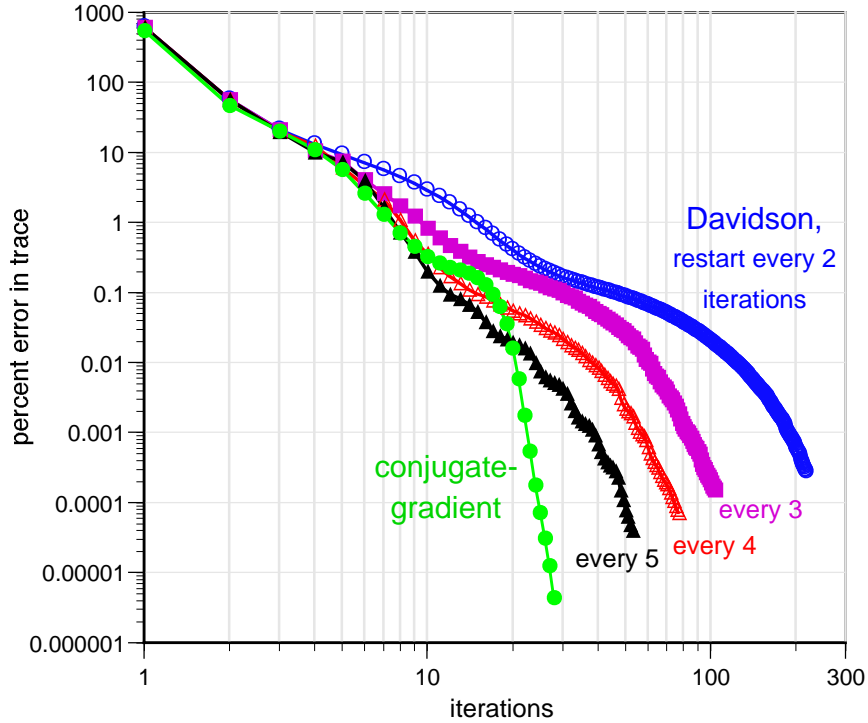


Figure 8.4: Comparison of the Davidson method with the block conjugate-gradient algorithm of Sec. 8.3.1. We reset the Davidson subspace to the best current eigenvectors every 2, 3, 4, or 5 iterations, with a corresponding increase in memory usage and computational costs.

future studies continue to investigate the Davidson method with different preconditioners and restarting schemes, especially since it seems to exhibit more regular convergence than the conjugate-gradient algorithm.

8.3.3 Interior eigenvalues

One of the most interesting aspects of photonic crystals is the possibility of localized modes associated with point defects (cavities) and line defects (waveguides) in the crystal [121],[175]. In this case, the desired modes lie in a known frequency range (the band gap) in the interior of the spectrum. Moreover, since a supercell of some sort must be used to surround the defect with sufficient bulk crystal, the underlying non-localized modes are “folded” many times (increasing their number by the number of primitive cells in the supercell). Ideally, one would like to compute only the defect modes in the band gap, with-

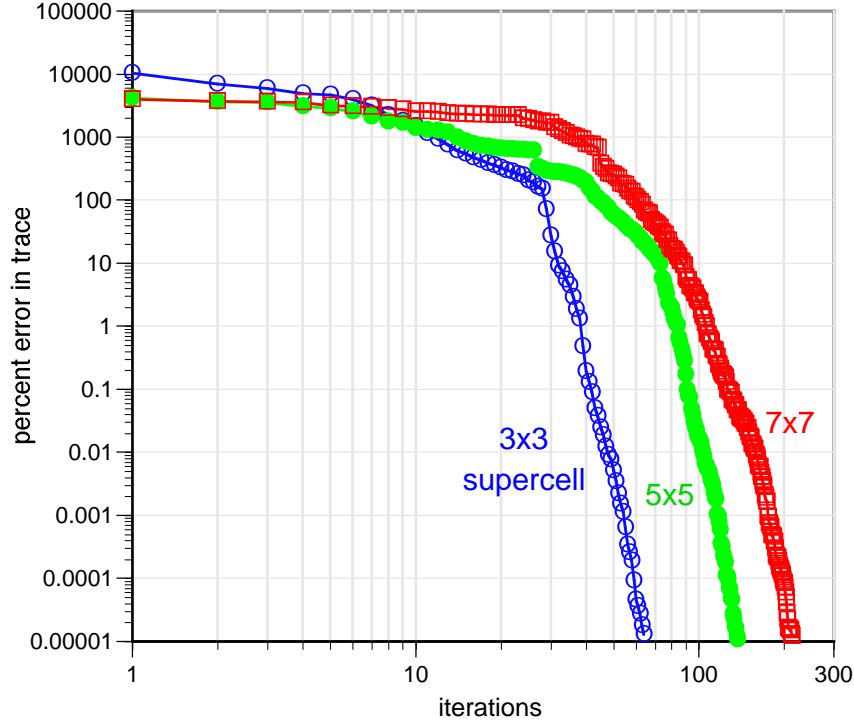


Figure 8.5: Conjugate-gradient convergence of the lowest TM eigenvalue for the “interior” eigensolver of Eq. (8.27), solving for the monopole defect state formed by one vacancy in a 2D square lattice of dielectric rods in air, using three different supercell sizes (3×3 , 5×5 , and 7×7).

out the waste of computation and memory of finding all the folded modes below them. One way to accomplish this is to use the FDTD algorithms referenced in Sec. 8.1. Here, we demonstrate that it is also possible to compute only the interior eigenvalues using iterative frequency-domain methods.

The center of the band gap is known, so we can state the problem as one of finding the eigenvalues and eigenvectors closest to the mid-gap frequency ω_m . Since the methods presented above compute the minimum eigenvalues of an operator, we can apply them directly merely by shifting the spectrum [176]:

$$\hat{A}'_{\hat{k}} = \left(\hat{A}_{\hat{k}} - \frac{\omega^2}{c^2} \right)^2 \quad (8.27)$$

This operator has the same eigenvectors as $\hat{A}_{\hat{k}}$, but its lowest eigenvalues are the ones

closest to ω_m , and thus we can compute a single defect state without computing any other eigenstates. As a preconditioner for this operator, we again approximate ω_m as small and simply use $\hat{K}G = \tilde{A}^{-2}GU^{-1}$, essentially the square of Eq. (8.25). This imperfect preconditioner and the worsened condition number of Eq. (8.27) will lead to slower eigensolver convergence, but this is generally more than compensated for by the reduction in the number of bands computed (and the resulting savings in memory and time per iteration). To demonstrate this “interior” eigensolver, we consider the case of a two-dimensional square lattice of dielectric rods in air, with a single rod removed from the crystal. Such a defect supports one confined TM mode in the gap [175], and in Fig. 8.5 we plot the eigensolver convergence of the lowest TM eigenvalue of Eq. (8.27), for 3×3 to 7×7 supercells with a resolution of 16 grid points per lattice constant. (To find the same state in a 7×7 supercell via the ordinary method, $p = 49$ bands must be computed.)

An alternative method for computing interior eigenvalues using a modified version of the Davidson method has also been suggested [164] and we intend to investigate this algorithm in future work, along with more accurate preconditioners (*e.g.* based on iterative linear solvers).

8.3.4 To block or not to block?

We have described the use of block eigensolvers to solve for all of the desired eigenstates simultaneously. The classical algorithm of computing them one by one with deflation has its advantages, however. First, it uses less memory— $O(pn)$ memory is required to store the eigenstates themselves, but only $O(n)$ memory is required for auxiliary matrices such as G and D , versus $O(pn)$ for the block method. (At least two such auxiliary matrices seem to be required for conjugate gradient, although three is more convenient, and one more for the Polak-Ribiere method.) Second, because our preconditioner is a better approximation for the lower bands (where ω is smallest), the eigensolver convergence

of the upper bands is slower—it is more efficient to compute them separately. Third, the eigensolver iterations themselves can require fewer operations for the deflation algorithm, as described below.

Since blocked algorithms maintain inherent cache-reuse advantages on modern computer architectures [168], however, it makes sense to consider a hybrid approach: we use the blocked eigensolver to solve for b bands at a time, with deflation to orthogonalize against the previous bands. The optimal choice of b depends upon the particular problem size considered and hardware/software details, but it typically ranges between five and fifteen bands. (As long as b is several times smaller than p , the memory advantages of deflation are realized.) Often, a factor of two or more in speed is gained compared to either $b = 1$ or $b = p$.

Let us consider the arithmetic complexity of the iterations. Each eigensolver iteration for b bands requires $O(knb^2)$ operations for k matrix operations such as ZU , and $O(bn\log n)$ operations for A in a planewave basis. Furthermore, orthogonalization against the q previous bands requires $O(2nbq)$ operations (once or twice per iteration, to maintain numerical stability). Assuming that all bands require the same number of iterations, this must be repeated $\frac{p}{b}$ times. Thus, each eigensolver iteration requires a total of $O(knpb) + O(pn\log n)$ operations for the bands themselves, plus $O(np(p - b))$ operations for the orthogonalization. So, if k is roughly greater than the number of orthogonalizations per iteration ($k \sim 8$ in our implementation), then deflation requires fewer operations per iteration than computing all of the bands at once. Of course, the actual running time scales in a more complicated way, due to the cache, cpu pipeline, and other issues.

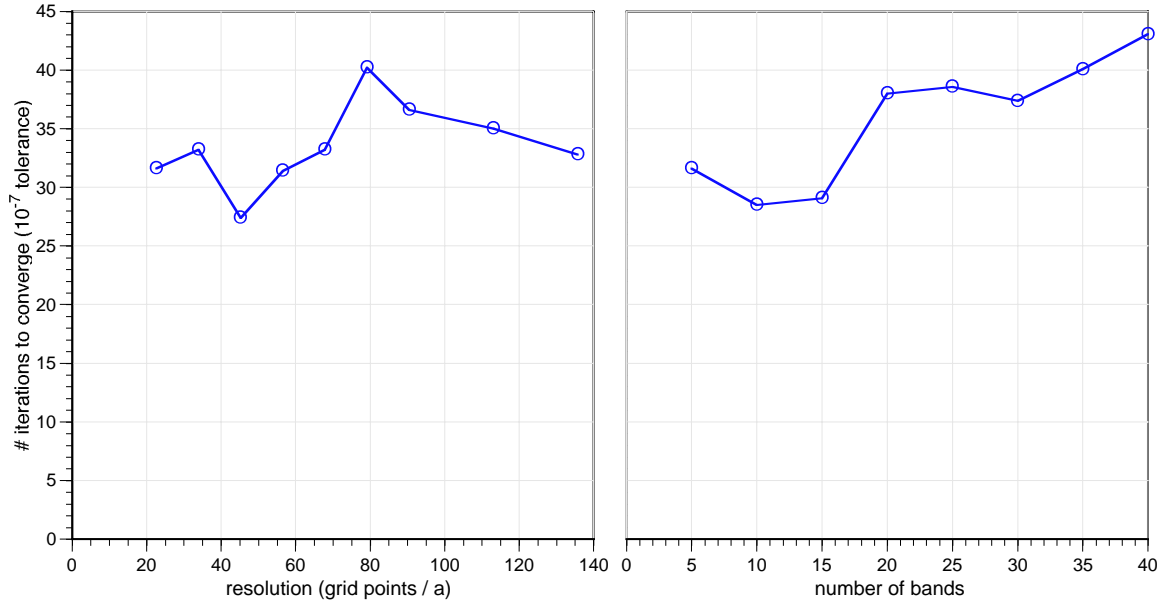


Figure 8.6: Scaling of the number of conjugate-gradient iterations required for convergence (to a fractional tolerance of 10^{-7}) as a function of the spatial resolution (in grid points per lattice constant, with a corresponding planewave cutoff), or the number p of bands computed.

8.3.5 Scaling

One final question that we wish to address is that of scaling: how does the eigensolver convergence rate scale with the size of the problem? Specifically, we consider the convergence of the block conjugate-gradient algorithm for the diamond structure as we increase either the resolution or the number of bands. In both cases, we plot the mean number of iterations (over five runs with random starting fields) for the eigensolver to converge to within a fractional tolerance of 10^{-7} (which is probably lower than the tolerance that would be used in practice). In the case of the increased number of bands, we solve for the bands in blocks of $b = 5$ bands, as described in Sec. 8.3.4, and use the average number of iterations per block. The results, plotted in Fig. 8.6, demonstrate that the number of iterations increases only very slowly as the problem size is increased. (We suspect that the slowed convergence for the increased number of bands results largely from the worsening of the small- ω approximation in our preconditioner.) In contrast, FDTD algorithms must

scale their number of iterations (or, strictly speaking, Δt) linearly with the spatial resolution in order to maintain stability.

8.4 Summary

We have presented efficient preconditioned block-iterative algorithms for computing eigenstates of Maxwell's equations for periodic dielectric systems using a planewave basis. Such methods, combined with appropriate effective dielectric tensors for accuracy, interior eigensolvers to compute only the desired modes in defect systems, and a flexible and freely-available implementation, provide an attractive way to perform eigen-analyses of diverse electromagnetic systems.

Chapter 9

Concluding Remarks

9.1 Looking Back

We began this thesis by introducing a way to intersect waveguides with nearly zero crosstalk, demonstrating the power of simple symmetry and resonance concepts in designing photonic components. The motivation for this work was to enable the creation of integrated networks of optical components, and especially networks in which one does not have to worry about the third dimension—a completely planar design is now possible without regard for crossing “wires.” Because we live in a three-dimensional universe, however, that third dimension, like an unwelcome guest, continues to bring worries to our door: what do we do about vertical confinement?

The subsequent chapters proposed two possible answers to that question. First, we considered how to achieve vertical confinement in the simplest possible way, by photonic-crystal slabs: truncating the height of a two-dimensional crystal and employing vertical index-guiding. We showed how this does, in fact, work...at a price: a complete band gap is sacrificed. Nevertheless, we were able to construct lossless waveguides and low-loss resonant cavities, thereby establishing the foundations for a wide class of useful devices. Second, we proposed a new three-dimensional photonic crystal with a *complete* band gap and a behavior capturing the essence of our favorite two-dimensional structures...at a price: increased complexity of fabrication, with multiple layers that need to be carefully aligned. The total absence of radiation losses and the associated headaches of avoiding them, however, make true photonic crystals like ours an almost irresistible target for future experimental and theoretical work.

Meanwhile, we addressed the question of how to analyze these kinds of optical systems on a computer. We developed new methods to ease the computational burden on theorists, allowing a greater variety of structures to be explored with less effort, and pulled together a wide range of developments in numerical algorithms. Moreover, we tackled a less mathematical but more practical difficulty—that of the enormous effort required to implement such computations in versatile software—by releasing the fruits of our labors for free use and adaptation by the scientific community.

9.2 Looking Forward

We've seen that the real world has properties that we sometimes dislike—it is three-dimensional—but we've found ways to circumvent such problems. Future efforts need to take these considerations further, and examine other ways in which reality falls short of theory. The most pressing general problem of this sort is disorder, the random imperfections that plague any experimental structure. Until now, the most common attitude towards disorder has been to ignore it; after all, if it is much tinier than the wavelength, its effects are small, and we have had other fish to fry. In any practical system, however, especially those involving high- Q resonators or slow-light waveguides, scattering from imperfections is likely to be a major limitation, and quantitative predictions of its effects are desperately needed. Other problems in this category include those of understanding the quantitative effects of symmetry-breaking in slabs, finite-size effects in three-dimensional crystals, material absorption, and thermal or nonlinear sensitivities. Our next thesis will cover these topics in greater detail.

*Enough of Science and of Art,
Close up those barren leaves;
Come forth, and bring with you a heart
That watches and receives.*

William Wordsworth, "The Tables Turned"

References

- [1] J. Clerk Maxwell, “A dynamical theory of the electromagnetic field,” *Philosophical Transactions of the Royal Society of London* **155**, 459–512 (1865). Abstract: *Proceedings of the Royal Society of London* **13**, 531–536 (1864).
- [2] F. Bloch, “Über die quantenmechanik der electronen in kristallgittern,” *Z. Physik* **52**, 555–600 (1928).
- [3] G. Floquet, “Sur les équations différentielles linéaires à coefficients périodiques,” *Ann. École Norm. Sup.* **12**, 47–88 (1883).
- [4] Lord Rayleigh, “On the maintenance of vibrations by forces of double frequency, and on the propagation of waves through a medium endowed with a periodic structure,” *Philosophical Magazine* **24**, 145–159 (1887).
- [5] Lord Rayleigh, “On the remarkable phenomenon of crystalline reflexion described by Prof. Stokes,” *Philosophical Magazine* **26**, 256–265 (1888).
- [6] Lord Rayleigh, “On the reflection of light from a regularly stratified medium,” *Proceedings of the Royal Society of London* **93**, 565–577 (1917).
- [7] G. W. Hill, “On the part of the motion of the lunar perigee which is a function of the mean motions of the sun and moon,” *Acta Math.* **8**, 1–36 (1886). (This is a reprint of an article published in 1877.)
- [8] J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light* (Princeton, September 1995).
- [9] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, “Photonic crystals: putting a new twist on light,” *Nature* **386**, 143 (1997).
- [10] K. Aretz, H. Bülow, “Reduction of Crosstalk and Losses of Intersecting Waveguide,” *Electronics Letters* **25**, no. 11, pp. 730–731 (May 1989).
- [11] M. G. Daly, P. E. Jessop, D. Yevick, “Crosstalk Reduction in Intersecting Rib Waveguides,” *J. Lightwave Tech.* **14**, no. 7, pp. 1695–1698 (July 1996).
- [12] H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, Englewood Cliffs, NJ, 1984), chapter 7.
- [13] See e.g. M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, NY, 1964).
- [14] J. D. Joannopoulos, P. R. Villeneuve, S. Fan, “Photonic Crystals: putting a new twist on light,” *Nature* **386**, pp. 143–149 (March 1997).
- [15] P. R. Villeneuve, S. Fan, J. D. Joannopoulos, “Microcavities in photonic crystals: Mode symmetry, tunability, and coupling efficiency,” *Phys. Rev. B* **54**, no. 11, pp. 7837 (1996).
- [16] See e.g. K. S. Kunz, R. J. Luebbers, *The Finite-Difference Time-Domain Methods* (CRC Press, Boca Raton, FL, 1993).
- [17] S. Fan, J. N. Winn, A. Devenyi, J. C. Chen, R. D. Meade, J. D. Joannopoulos, “Guided and Defect Modes in Periodic Dielectric Waveguides,” *J. Opt. Soc. Am. B* **12**, pp. 1267 (1995).
- [18] See e.g. H. A. Haus, Y. Lai, “Theory of cascaded quarter wave shifted distributed feedback resonators,” *IEEE Journal of Quantum Electronics* **28**, no. 1, pp. 205–213 (January 1992).

- [19] E. Yablonovitch, T. J. Gmitter, and K. M. Leung, *Phys. Rev. Lett.* **67**, 2295 (1991).
- [20] H. S. Sözüer and J. P. Dowling, *J. Mod. Opt.* **41**, 231 (1994).
- [21] K. Ho, C. Chan, C. Soukoulis, R. Biswas, and M. Sigalas, *Solid State Comm.* **89**, 413 (1994).
- [22] S. Fan, P. R. Villeneuve, R. Meade, and J. D. Joannopoulos, *Appl. Phys. Lett.* **65**, 1466 (1994).
- [23] C. C. Cheng and A. Scherer, *J. Vac. Sci. Tech. B* **13**, 2696 (1995).
- [24] S. Noda, N. Yamamoto, and A. Sasaki, *Jpn. J. Appl. Phys.* **35**, L909 (1996).
- [25] S. G. Romanov, N. P. Johnson, A. V. Fokin, V. Y. Butko, H. M. Yates, M. E. Pemble, and C. M. S. Torres, *Appl. Phys. Lett.* **70**, 2091 (1997).
- [26] G. Feiertag, W. Ehrfeld, H. Freimuth, H. Kolle, H. Lehr, M. Schmidt, M. M. Sigalas, C. M. Soukoulis, G. Kiriakidis, and T. Pedersen, *Appl. Phys. Lett.* **71**, 1441 (1997).
- [27] S. Y. Lin, J. G. Fleming, D. L. Hetherington, B. K. Smith, R. Biswas, K. M. Ho, M. M. Sigalas, W. Zubrzycki, S. R. Kurtz, J. Bur, *Nature* **394**, 251 (1998).
- [28] R. D. Meade, A. Devenyi, J. D. Joannopoulos, O. L. Alerhand, D. A. Smith, and K. Kash, *J. Appl. Phys.* **75**, 4753 (1994).
- [29] P. L. Gourley, J. R. Wendt, G. A. Vawter, T. M. Brennan, and B. E. Hammons, *Appl. Phys. Lett.* **64**, 687 (1994).
- [30] S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and E. F. Schubert, *Proc. SPIE* **3002**, 67 (1997).
- [31] M. Boroditsky, R. Coccioli, and E. Yablonovitch, *Proc. SPIE* 3283, 1 (1998).
- [32] P. R. Villeneuve, S. Fan, S. G. Johnson, and J. D. Joannopoulos, *IEE Proc. Optoelec.* **145**, 384 (1998).
- [33] R. Coccioli, M. Boroditsky, K. W. Kim, Y. Rahmat-Samii, and E. Yablonovitch, *IEE Proc. Optoelec.* **145**, 391 (1998).
- [34] B. D'Urso, O. Painter, J. O'Brien, T. Tombrello, A. Yariv, and A. Scherer, *J. Opt. Soc. Am. B* **15**, 1155 (1997).
- [35] M. Kanskar, P. Paddon, V. Pacradouni, R. Morin, A. Busch, J. F. Young, S. R. Johnson, J. Mackenzie, and T. Tiedje, *Appl. Phys. Lett.* **70**, 1438 (1997).
- [36] R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, *Phys. Rev. B* **48**, 8434 (1993). Erratum: S. G. Johnson, *Phys. Rev. B* **55**, 15942 (1997).
- [37] R. A. Waldron, "Theory and potential applications of backward waves in nonperiodic inhomogenous waveguides, *Proc. IEEE* **111**, (1964).
- [38] M. Ibanescu, Y. Fink, S. Fan, E. L. Thomas, and J. D. Joannopoulos, "An all-dielectric coaxial waveguide," *Science* **289**, 415–419 (2000).
- [39] Y. Chen, G. Faini, H. Launois, and J. Etrillard, *Superlattices and Microstructures* **22**, 109 (1997).
- [40] T. Kraus, R. De La Rue, and S. Band, *Nature* **383**, 699 (1996).
- [41] D. Labilloy, H. Benisty, C. Weisbuch, T. F. Krauss, R. M. De La Rue, V. Bardinal, R. Houdre, U. Oesterle, D. Cassagne, and C. Jouanin, *Phys. Rev. Lett.* **79**, 4147 (1997).
- [42] A. Mekis, J. C. Chen, I. Kurland, S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, *Phys. Rev. Lett.* **77**, 3787 (1996).

- [43] S. Fan, P. R. Villeneuve, J. D. Joannopoulos, H. A. Haus, *Phys. Rev. Lett.* **80**, 960 (1998).
- [44] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, *Nature* **386**, 143 (1997).
- [45] R. D. Meade, A. Devenyi, J. D. Joannopoulos, O. L. Alerhand, D. A. Smith, and K. Kash, *J. Appl. Phys.* **75**, 4753 (1994).
- [46] H. Benisty, *J. Appl. Phys.* **79**, 7483 (1996).
- [47] A. Mekis, S. Fan, and J. D. Joannopoulos, *Phys. Rev. B* **58**, 4809 (1998).
- [48] A. Mekis, J. C. Chen, I. Kurland, S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, *Phys. Rev. Lett.* **77**, 3787 (1996).
- [49] S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, *Phys. Rev. Lett.* **80**, 960 (1998).
- [50] S. G. Johnson, C. Manolatou, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and H. A. Haus, *Optics Letters* **23**, 1855 (1998).
- [51] P. L. Gourley, J. R. Wendt, G. A. Vawter, T. M. Brennan, and B. E. Hammons, *Appl. Phys. Lett.* **64**, 687 (1994).
- [52] T. Kraus, R. De La Rue, and S. Band, *Nature* **383**, 699 (1996).
- [53] S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and E. F. Schubert, *Phys. Rev. Lett.* **78**, 3294 (1997).
- [54] B. D'Urso, O. Painter, J. O'Brien, T. Tombrello, A. Yariv, and A. Scherer, *J. Opt. Soc. Am. B* **15**, 1155 (1997).
- [55] M. Kanskar, P. Paddon, V. Pacradouni, R. Morin, A. Busch, J. F. Young, S. R. Johnson, J. Mackenzie, and T. Tiedje, *Appl. Phys. Lett.* **70**, 1438 (1997).
- [56] M. D. B. Charlton, S. W. Roberts, and G. J. Parker, *Materials Sci. and Eng. B* **B49**, 155 (1997).
- [57] Y. Chen, G. Faini, H. Launois, and J. Etrillard, *Superlattices and Microstructures* **22**, 109 (1997).
- [58] D. Labilloy, H. Benisty, C. Weisbuch, T. F. Krauss, R. M. De La Rue, V. Bardinal, R. Houdre, U. Oesterle, D. Cassagne, and C. Jouanin, *Phys. Rev. Lett.* **79**, 4147 (1997).
- [59] P. R. Villeneuve, S. Fan, S. G. Johnson, and J. D. Joannopoulos, *IEE Proc. Optoelec.* **145**, 384 (1998).
- [60] R. Coccioli, M. Boroditsky, K. W. Kim, Y. Rahmat-Samii, and E. Yablonovitch, *IEE Proc. Optoelec.* **145**, 391 (1998).
- [61] M. Boroditsky, R. Coccioli, and E. Yablonovitch, *Proc. SPIE* **3283**, 1 (1998).
- [62] S. G. Johnson, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, *Phys. Rev. B.* **60**, 5751 (1999).
- [63] H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, Englewood Cliffs, NJ, 1984), chapter 7.
- [64] C. Manolatau, S. G. Johnson, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, *J. Lightwave Tech.* **17**, no. 9 (1999), in press.
- [65] J. S. Foresi, P. R. Villeneuve, J. Ferrera, E. R. Thoen, G. Steinmeyer, S. Fan, J. D. Joannopoulos, L. C. Kimerling, H. I. Smith, and E. P. Ippen, *Nature* **390**, 143 (1997).
- [66] S. Y. Lin, E. Chow, V. Hietch, P. R. Villeneuve, J. D. Joannopoulos, *Science* **282**, 274 (1998).

- [67] E. Yablonovitch, T. J. Gmitter, and K. M. Leung, *Phys. Rev. Lett.* **67**, 2295 (1991).
- [68] H. S. Sözüer and J. P. Dowling, *J. Mod. Opt.* **41**, 231 (1994).
- [69] K. Ho, C. Chan, C. Soukoulis, R. Biswas, and M. Sigalas, *Solid State Comm.* **89**, 413 (1994).
- [70] S. Fan, P. R. Villeneuve, R. Meade, and J. D. Joannopoulos, *Appl. Phys. Lett.* **65**, 1466 (1994).
- [71] T. Baba, N. Fukaya, and J. Yonekura, *Electron. Lett.* **35**, p. 654 (1999).
- [72] B. Temelkuran and E. Ozbay, *Appl. Phys. Lett.* **74**, p. 486 (1999).
- [73] S. Kuchinsky, D. C. Allan, N. F. Borrelli, and J.-C. Cotteverte, *Optics Communications* **175**, 147 (2000).
- [74] J. C. Knight, J. Broeng, T. A. Birks, and P. St. Russell, *Science* **282**, 1476 (1998).
- [75] J. K. Ranka, R. S. Windeler, A. J. Stentz, *CLEO '99 Postdeadline Papers CPD8-1* (May 23–29, 1999) (Baltimore, MD).
- [76] O. Hanaizumi, Y. Ohtera, T. Sato, and S. Kawakami, *Appl. Phys. Lett.* **74**, 777 (1999).
- [77] A. Edelman and S. T. Smith, *BIT* **36**, 494 (1996).
- [78] R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, *Phys. Rev. B* **48**, 8434 (1993). Erratum: S. G. Johnson, *Phys. Rev. B* **55**, 15942 (1997).
- [79] K. M. Leung and Y. F. Liu, *Phys. Rev. Lett.* **65**, 2646 (1990).
- [80] Z. Zhang and S. Satpathy, *Phys. Rev. Lett.* **65**, 2650 (1990).
- [81] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, “Photonic crystals: putting a new twist on light,” *Nature* **386**, 143 (1997).
- [82] S. Fan, P. R. Villeneuve, J. D. Joannopoulos, and E. F. Schubert, “High extraction efficiency of spontaneous emission from slabs of photonic crystals,” *Phys. Rev. Lett.* **78**, 3294 (1997).
- [83] P. R. Villeneuve, S. Fan, S. G. Johnson, and J. D. Joannopoulos, “Three-dimensional photon confinement in photonic crystals of low-dimensional periodicity,” *IEE Proc. Optoelec.* **145**, 384 (1998).
- [84] R. Coccioli, M. Boroditsky, K. Kim, Y. Rahmat-Samii, and E. Yablonovitch, “What is the smallest possible electromagnetic mode volume in a dielectric cavity?” *IEE Proc. Optoelec.* **145**, 391 (1998).
- [85] O. J. Painter, A. Husain, A. Scherer, J. D. O’Brien, I. Kim, and P. D. Dapkus, “Defect modes of a two-dimensional photonic crystal in an optically thin dielectric slab,” *J. Opt. Soc. Am. B* **16**, 275 (1999).
- [86] S. G. Johnson, S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, “Guided modes in photonic crystal slabs,” *Phys. Rev. B.* **60**, 5751 (1999).
- [87] H. Benisty, D. Labilloy, C. Weisbuch, C. J. M. Smith, T. F. Krauss, D. Cassagne, A. Beraud, and C. Jouanin, “Radiation losses of waveguide-based two-dimensional photonic crystals: positive role of the substrate,” *Appl. Phys. Lett.* **76**, 532 (2000).
- [88] S. Kuchinsky, D. C. Allan, N. F. Borrelli, and J.-C. Cotteverte, “3d localization in a channel waveguide in a photonic crystal with 2d periodicity,” *Opt. Commun.* **175**, 147 (2000).
- [89] S. G. Johnson, P. R. Villeneuve, S. Fan, and J. D. Joannopoulos, “Linear waveguides in photonic-crystal slabs,” *Phys. Rev. B*, in press (2000).

- [90] P. R. Villeneuve, S. Fan, J. D. Joannopoulos, K.-Y. Lim, G. S. Petrich, L. A. Kolodziejski, R. Reif, "Air-bridge microcavities," *Appl. Phys. Lett.* **67**, 167 (1995).
- [91] J. D. Joannopoulos, S. Fan, A. Mekis, and S. G. Johnson, "Novelties of light with photonic crystals," *Proc. NATO ASI 2000: Photonic Crystals and Light Localization*, in press.
- [92] H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, Englewood Cliffs, NJ, 1984), chapter 7.
- [93] C. Manolatou, S. G. Johnson, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, "High-density integrated optics," *J. Lightwave Tech.* **17**, 1682 (1999).
- [94] A. W. Snyder, "Radiation losses due to variations of radius on dielectric or optical fibers," *IEEE Trans. Microwave Theory Tech.* **MTT-18**, 608 (1970).
- [95] J. D. Jackson, *Classical Electrodynamics, 3rd ed.* (Wiley, New York, NY, 1999), chapter 9.
- [96] See, e.g., K. S. Kunz and R. J. Luebbers, *The Finite-Difference Time-Domain Methods* (CRC, Boca Raton, FL, 1993).
- [97] V. A. Mandelshtam and H. S. Taylor, "Harmonic inversion of time signals and its applications," *J. Chem. Phys.* **107**, 6756 (1997). Erratum: *ibid.*, **109**, 4128 (1998).
- [98] E. Yablonovitch, *Phys. Rev. Lett.* **58**, 2059 (1987).
- [99] S. John, *Phys. Rev. Lett.* **58**, 2486 (1987).
- [100] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, *Nature* **386**, 143 (1997).
- [101] K. M. Ho, C. T. Chan, C. M. Soukoulis, *Phys. Rev. Lett.* **65**, 3152 (1990).
- [102] C. T. Chan, S. Datta, K. M. Ho, and C. M. Soukoulis, *Phys. Rev. B* **50**, 1988 (1994).
- [103] B. T. Holland, C. F. Blanford, A. Stein, *Science* **281**, 538 (1998).
- [104] J. E. G. J. Wijnhoven, W. L. Vos, *Science* **281**, 802 (1998).
- [105] K. Busch and S. John, *Phys. Rev. E* **58**, 3896 (1998).
- [106] E. Yablonovitch, T. J. Gmitter, and K. M. Leung, *Phys. Rev. Lett.* **67**, 2295 (1991).
- [107] K. Ho, C. Chan, C. Soukoulis, R. Biswas, and M. Sigalas, *Solid State Comm.* **89**, 413 (1994).
- [108] H. S. Sözüer and J. P. Dowling, *J. Mod. Opt.* **41**, 231 (1994).
- [109] K. M. Leung, *Phys. Rev. B* **56**, 3517 (1997).
- [110] S. Fan, P. R. Villeneuve, R. Meade, and J. D. Joannopoulos, *Appl. Phys. Lett.* **65**, 1466 (1994).
- [111] H. S. Sözüer and J. W. Haus, *J. Opt. Soc. Am. B* **10**, 296 (1993).
- [112] For recent work, see, e.g., *Photonic Crystals and Light Localization*, edited by C. M. Soukoulis, Proceedings of the NATO ASI on Photonic Band Gap Materials, Limin Hersonissou, Crete, Greece, June 19-30, 2000 (NATO ASI Series, Kluwer Academic Publishers, The Netherlands, 2001).
- [113] R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, *Phys. Rev. B* **48**, 8434 (1993). Erratum: S. G. Johnson, *Phys. Rev. B* **55**, 15942 (1997).
- [114] S. G. Johnson and J. D. Joannopoulos, *The MIT Photonic-Bands Package* (<http://ab-initio.mit.edu/mpb/>).

- [115] See, *e.g.*, K. S. Kunz and R. J. Luebbers, *The Finite-Difference Time-Domain Methods* (CRC, Boca Raton, FL, 1993).
- [116] E. Ozbay, A. Abeyta, G. Tuttle, M. Tringides, R. Biswas, C. T. Chan, C. M. Soukoulis, and K. M. Ho, *Phys. Rev. B* **50**, 1945 (1994).
- [117] S. Y. Lin, J. G. Fleming, D. L. Hetherington, B. K. Smith, R. Biswas, K. M. Ho, M. M. Sigalas, W. Zubrzycki, S. R. Kurtz, J. Bur, *Nature* **394**, 251 (1998).
- [118] S. Noda, N. Yamamoto, and A. Sasaki, *Jpn. J. Appl. Phys.* **35**, L909 (1996).
- [119] S. Noda, K. Tomoda, N. Yamamoto, A. Chutinan, *Science* **289**, 604 (2000).
- [120] J. G. Fleming and S. Y. Lin, private communication.
- [121] See, *e.g.*, J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, “Photonic crystals: putting a new twist on light,” *Nature* **386**, 143–149 (1997).
- [122] S. G. Johnson and J. D. Joannopoulos, The MIT Photonic-Bands Package home page <http://ab-initio.mit.edu/mpb/>.
- [123] K. M. Ho, C. T. Chan, and C. M. Soukoulis, “Existence of a photonic gap in periodic dielectric structures,” *Phys. Rev. Lett.* **65**, 3152–3155 (1990).
- [124] H. S. Sözüer and J. W. Haus, “Photonic bands: convergence problems with the plane-wave method,” *Phys. Rev. B* **45**, 13962–13972 (1992).
- [125] R. D. Meade, A. M. Rappe, K. D. Brommer, J. D. Joannopoulos, and O. L. Alerhand, “Accurate theoretical analysis of photonic band-gap materials,” *Phys. Rev. B* **48**, 8434–8437 (1993). Erratum: S. G. Johnson, *ibid* **55**, 15942 (1997).
- [126] T. Suzuki and P. K. L. Yu, “Method of projection operators for photonic band structures with perfectly conducting elements,” *Phys. Rev. B* **57**, 2229–2241 (1998).
- [127] K. Busch and S. John, “Liquid-crystal photonic-band-gap materials: the tunable electromagnetic vacuum,” *Phys. Rev. Lett.* **83**, 967–970 (1999).
- [128] J. Jin, *The Finite-Element Method in Electromagnetics* (Wiley, New York, 1993), Chap. 5.7.
- [129] A. Figotin, Y. A. Godin, “The computation of spectra of some 2D photonic crystals,” *J. Comput. Phys.* **136**, 585–598 (1997).
- [130] W. C. Sailor, F. M. Mueller, and P. R. Villeneuve, “Augmented-plane-wave method for photonic band-gap materials,” *Phys. Rev. B* **57**, 8819–8822 (1998).
- [131] W. Axmann and P. Kuchment, “An efficient finite element method for computing spectra of photonic and acoustic band-gap materials: I. Scalar case,” *J. Comput. Phys.* **150**, 468–481 (1999).
- [132] D. C. Dobson, “An efficient method for band structure calculations in 2D photonic crystals,” *J. Comput. Phys.* **149**, 363–376 (1999).
- [133] D. Mogilevtsev, T. A. Birks, and P. St. J. Russell, “Localized function method for modeling defect modes in 2D photonic crystals,” *J. Lightwave Tech.* **17**, 2078–2081 (1999).
- [134] S. J. Cooke and B. Levush, “Eigenmode solution of 2-D and 3-D electromagnetic cavities containing absorbing materials using the Jacobi-Davidson algorithm,” *J. Comput. Phys.* **157**, 350–370 (2000).
- [135] K. M. Leung, “Defect modes in photonic band structures: a Green’s function approach using vector Wannier functions,” *J. Opt. Soc. Am. B* **10**, 303–306 (1993).

- [136] J. P. Albert, C. Jouanin, D. Cassagne, and D. Bertho, “Generalized Wannier function method for photonic crystals,” *Phys. Rev. B* **61**, 4381–4384 (2000).
- [137] E. Lidorikis, M. M. Sigalas, E. N. Economou, and C. M. Soukoulis, “Tight-binding parameterization for photonic band-gap materials,” *Phys. Rev. Lett.* **81**, 1405–1408 (1998).
- [138] See, *e.g.*, K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Methods* (CRC, Boca Raton, Fla., 1993).
- [139] C. T. Chan, S. Datta, Q. L. Yu, M. Sigalas, K. M. Ho, C. M. Soukoulis, “New structures and algorithms for photonic band gaps,” *Physica A* **211**, 411–419 (1994).
- [140] C. T. Chan, Q. L. Lu, and K. M. Ho, “Order-N spectral method for electromagnetic waves,” *Phys. Rev. B* **51**, 16635–16642 (1995).
- [141] S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, “Large omnidirectional band gaps in metallo-dielectric photonic crystals,” *Phys. Rev. B* **54**, 11245–11251 (1996).
- [142] K. Sakoda and H. Shiroma, “Numerical method for localized defect modes in photonic lattices,” *Phys. Rev. B* **56**, 4830–4835 (1997).
- [143] J. Arriaga, A. J. Ward, and J. B. Pendry, “Order N photonic band structures for metals and other dispersive materials,” *Phys. Rev. B* **59**, 1874–1877 (1999).
- [144] A. J. Ward and J. B. Pendry, “A program for calculating photonic band structures, Green’s functions and transmission/reflection coefficients using a non-orthogonal FDTD method,” *Comput. Phys. Comm.* **128**, 590–621 (2000).
- [145] P. Yeh, *Optical Waves in Layered Media* (Wiley, New York, 1988).
- [146] J. B. Pendry and A. MacKinnon, “Calculation of photon dispersion relations,” *Phys. Rev. Lett.* **69**, 2772–2775 (1992).
- [147] P. M. Bell, J. B. Pendry, L. M. Moreno, and A. J. Ward, “A program for calculating photonic band structures and transmission coefficients of complex structures,” *Comput. Phys. Comm.* **85**, 306–322 (1995).
- [148] J. M. Elson and P. Tran, “Dispersion in photonic media and diffraction from gratings: a different modal expansion for the R-matrix propagation technique,” *J. Opt. Soc. Am. A* **12**, 1765–1771 (1995).
- [149] J. M. Elson and P. Tran, “Coupled-mode calculation with the R-matrix propagator for the dispersion of surface waves on truncated photonic crystal,” *Phys. Rev. B* **54**, 1711–1715 (1996).
- [150] J. Chongjun, Q. Bai, Y. Miao, and Q. Ruhu, “Two-dimensional photonic band structure in the chiral medium—transfer matrix method,” *Opt. Commun.* **142**, 179–183 (1997).
- [151] V. A. Mandelshtam and H. S. Taylor, “Harmonic inversion of time signals,” *J. Chem. Phys.* **107**, 6756–6769 (1997). Erratum: *ibid*, 109, 4128 (1998).
- [152] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt Saunders, Philadelphia, 1976).
- [153] M. Frigo and S. G. Johnson, “FFTW: an adaptive software architecture for the FFT,” in *Proc. 1998 IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing* (Institute of Electrical and Electronics Engineers, New York, 1998), 1381–1384.

- [154] A. H. Stroud, *Approximate Calculation of Multiple Integrals* (Prentice-Hall, Englewood Cliffs, NJ, 1971).
- [155] J. Nadobny, D. Sullivan, P. Wust, M. Seebass, P. Deuffhard, and R. Felix, “A high-resolution interpolation at arbitrary interfaces for the FDTD method,” *IEEE Trans. Microwave Theory Tech.* **46**, 1759–1766 (1998).
- [156] P. Yang, K. N. Liou, M. I. Mishchenko, and B.-C. Gao, “Efficient finite-difference time-domain scheme for light scattering by dielectric particles: application to aerosols,” *Appl. Opt.* **39**, 3727–3737 (2000).
- [157] F. Ma, C. L. Xu, and W. P. Huang, “Wide-angle full vectorial beam propagation method,” *IEE Proc. Optoelec.* **143**, 139–143 (1996).
- [158] R. D. Meade, private communications.
- [159] M. C. Payne, M. P. Tater, 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**, 1045–1097 (1992).
- [160] See, e.g., A. Edelman and S. T. Smith, “On conjugate gradient-like methods for eigen-like problems,” *BIT* **36**, 494–509 (1996).
- [161] S. Ismail-Beigi and T. A. Arias, “New algebraic formulation of density functional calculation,” *Comp. Phys. Commun.* **128**, 1–45 (2000).
- [162] E. R. Davidson, “The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices,” *Comput. Phys.* **17**, 87–94 (1975).
- [163] M. Crouzeix, B. Philippe, and M. Sadkane, “The Davidson Method,” *SIAM J. Sci. Comput.* **15**, 62–76 (1994).
- [164] G. L. G. Sleijpen and H. A. van der Vorst, “A Jacobi-Davidson iteration method for linear eigenvalue problems,” *SIAM J. Matrix Anal. Appl.* **17**, 401–425 (1996).
- [165] B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980).
- [166] H. A. van der Vorst, “Krylov subspace iteration,” *Computing in Sci. and Eng.* **2**, 32–37 (2000).
- [167] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, London, 1981).
- [168] J. J. Dongarra, J. Du Croz, I. S. Duff, and S. Hammarling, “A set of Level 3 Basic Linear Algebra Subprograms,” *ACM Trans. Math. Soft.* **16**, 1–17 (1990).
- [169] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users’ Guide* (SIAM, Philadelphia, 1999).
- [170] A. Edelman, T. A. Arias, and S. T. Smith, “The geometry of algorithms with orthogonality constraints,” *SIAM J. Matrix Anal. Appl.* **20**, 303–353 (1998).
- [171] A. H. Sameh and J. A. Wisniewski, “A trace minimization algorithm for the generalized eigenvalue problem,” *SIAM J. Numer. Anal.* **19**, 1243–1259 (1982).
- [172] B. Philippe, “An algorithm to improve nearly orthonormal sets of vectors on a vector processor,” *SIAM J. Alg. Disc. Meth.* **8**, 396–403 (1987).

- [173] J. J. More and D. J. Thuente, “Line search algorithms with guaranteed sufficient decrease,” *ACM Trans. Math. Software* **20**, 286–307 (1994).
- [174] S. Ismail-Beiji, private communications.
- [175] P. R. Villeneuve, S. Fan, and J. D. Joannopoulos, “Microcavities in photonic crystals: mode symmetry, tunability, and coupling efficiency,” *Phys. Rev. B* **54**, 7837–7842 (1996).
- [176] L.-W. Wang and A. Zunger, “Solving Schrodinger’s equation around a desired energy: application to Silicon quantum dots,” *J. Chem. Phys.* **100**, 2394–2397 (1994).