

AB INITIO THEORETICAL STUDIES OF TRANSITION-METAL,  
MOLECULAR, AND PHOTONIC BAND-GAP MATERIALS

by

ANDREW MARSHALL RAPPE

A. B. Chemistry and Physics, Harvard College  
(1986)

Submitted to the Department of Physics and Chemistry in Partial Fulfillment of the  
Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN PHYSICS AND CHEMISTRY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1992

© 1992 Massachusetts Institute of Technology  
All rights reserved

Signature of Author

\_\_\_\_\_  
Department of Physics and Chemistry  
May 4, 1992

Certified by

\_\_\_\_\_  
John D. Joannopoulos  
Professor, Department of Physics  
Thesis Supervisor

Accepted by

\_\_\_\_\_  
Robert Silbey, Chair  
Departmental Graduate Committee  
Department of Physics and Chemistry

ARCHIVES  
MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

JUL 30 1992

LIBRARIES

AB INITIO THEORETICAL STUDIES OF TRANSITION-METAL, MOLECULAR,  
AND PHOTONIC BAND-GAP MATERIALS

by

ANDREW MARSHALL RAPPE

Submitted to the Department of Physics and Chemistry at the Massachusetts Institute of Technology on May 4, 1992 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Physics and Chemistry

**ABSTRACT**

The first part of this thesis reports on improvements made to the plane-wave pseudopotential method for performing electronic structure calculations within the local density approximation of density functional theory. Pseudopotentials are designed which require as few plane-wave basis functions as possible for the accurate description of any element, within the conventional plane-wave pseudopotential method. Calculations are performed using these methodological improvements. The structural properties of many organic molecules are computed accurately without any experimental input. Also, two proposed relaxed structures of a grain boundary in gold are studied, and the more favorable structure is identified.

The second part of this thesis covers investigations into the electromagnetic properties of materials with periodically modulated dielectric constants called photonic crystals. Essentially exact solutions of Maxwell's equations are computed, allowing the prediction of a number of new physical phenomena. This theoretical research predicts the existence of localized modes about defects in photonic crystals, the existence of surface electromagnetic modes on photonic crystals, and the existence of photonic crystals with two-dimensional periodicity containing an in-plane photonic band gap. These results have been confirmed by experiments, some of which are reported in the thesis as well.

Thesis Supervisor: Dr. John D. Joannopoulos

Title: Professor of Physics

This doctoral thesis has been examined by a Committee of the interdepartmental committee in Physics and Chemistry as follows:

Professor Robert Silbey

---

Chair

Professor John D. Joannopoulos

---

Thesis Supervisor

Professor Keith Nelson

---



Acknowledgments:

I wish to thank all the members of Professor Joannopoulos' research group, past and present, for providing a stimulating and supportive research environment. In particular, I thank Karin Rabe and Thimios Kaxiras for showing me how exciting, challenging and fun research in condensed matter physics could be. I am grateful to Bob Meade for sharing his enthusiasm for life and sense of humor. I also thank my collaborators Arnaldo, Paul, Mark, and Karl for the challenging problems we have tackled together.

I am grateful to Professor Joannopoulos for choosing me to be in his research group. I greatly value his efforts to make me a better scientist and a better person. His high moral character, his dedication to excellent science, and his easygoing manner have been an inspiration to me.

I thank my good friends Ben, Kamala, Jeff, Michael, Scott, Len, Estelle, Fred, Heidi, Wick, Franklin, and Kipper for their constant support along this long road. Playing bridge, rooting for Harvard's sports teams, and just visiting with me really helped me keep my life together. I hope that we will remain friends forever.

But I owe my success to my family and to my beloved. Mom and Linda, you stood by me when it really counted. I could never replace you and I feel fortunate to have you in my life. I am proud to call you my family. Barri, there is no measure of my love for you or my gratitude for the many ways you have helped me to reach this day. I love you.

## TABLE OF CONTENTS

Introduction	5
Chapter 1: Optimized Pseudopotentials	14
Chapter 2: The Design of Convergent and Transferable <i>Ab Initio</i> Pseudopotentials	27
Chapter 3: A Test of the Utility of Plane Waves for the Study of Molecules from First Principles	40
Chapter 4: <i>Ab Initio</i> Study of a Grain Boundary in Gold	51
Chapter 5: Photonic Bound States in Periodic Dielectric Materials	62
Chapter 6: Donor and Acceptor Modes in Photonic Band Structure	73
Chapter 7: The Existence of a Photonic Band Gap in Two Dimensions	85
Chapter 8: Measurement of Photonic Band Structure in a Two-Dimensional Periodic Dielectric Array	97
Chapter 9: Electromagnetic Bloch Waves at the Surface of a Photonic Crystal	109

## **Introduction**

The wish to understand how the world works is a fundamental human yearning. An expression of this understanding is the ability to predict the behavior of one's surroundings and then experimentally to observe such behavior in accord with experiment. The fundamental equations governing quantum-mechanical and electromagnetic systems have been known for nearly a century. Solution of these equations should allow one in principle to predict accurately the electronic and electromagnetic properties of any real system. However, the exact solution of these equations is in general so demanding that it cannot be accomplished exactly for any but the simplest systems. As a result, judicious approximation of the governing equations and accurate solution of them has become an active area of research in the physical sciences. This thesis will chronicle my contributions to the advancement of computational methodology and my application of these new methods in order to elucidate electronic and electromagnetic properties of real materials.

The interaction of quantum-mechanical particles via a Coulomb potential is accurately formulated by quantum electrodynamics, but the use of this formalism for the study of real materials is currently out of the question due to its complexity. To simplify calculations, the nuclear degrees of freedom can be treated classically with very little loss of accuracy. The electrons are so much lighter than the nuclei that they respond nearly instantaneously to changes in nuclear positions (Born-Oppenheimer approximation). Now, the quantum-mechanical portion of the calculation is the determination of electronic states in an external potential, in conjunction with the classical motion of the nuclei according to the forces placed on them.

Several approaches exist to solve for electronic states essentially exactly within the Born-Oppenheimer approximation. For example, the quantum Monte Carlo method directly applies the many-electron Schrodinger or Dirac equation, and samples the solution statistically. The only error involved is the sampling error. This method yields

extraordinarily accurate results for a modest number of electrons<sup>1</sup>, but the computational effort rises very swiftly as the number of electrons increases and the symmetry of the system decreases. Therefore, it is not the method of choice for the study of large or complicated systems. The same comment applies to the configuration interaction (CI) method. By including all electron-electron correlations explicitly, its computational effort scales exponentially with the number of electrons in the system.

Because the above methods do not provide information about large and complicated systems of interest, other methods involving further approximations have been developed. The Hartree-Fock method approximates the many-electron wavefunction as a Slater determinant of one-electron wavefunctions. This insures that the wavefunction will have the proper fermionic character, but it neglects many other electron-electron interactions. Corrections to Hartree-Fock have been developed including Moller-Plesset perturbation theory. These methods in general provide less accurate results than CI or quantum Monte Carlo, but they are much more computationally tractable. As a result, Hartree-Fock and corrections to it are the basis for most quantum chemistry investigations<sup>2</sup> of molecular structure and reactivity from first principles.

Another direction for computational effort has been toward semi-empirical approaches. The embedded atom method<sup>3</sup> and the MOPAC<sup>4</sup> package are examples of these. These methods rely on inputs from first-principles calculations and from experiment to produce results. The chief difficulty with these methods is that while they accurately reproduce results for the systems for which they were parametrized, they do not have a good track record for predicting the properties of new systems. This severely limits their utility.

---

<sup>1</sup>J. B. Anderson, C. A. Traynor, and B. M. Boghosian, preprint (1991); S. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. Lett. **61**, 1631 (1988).

<sup>2</sup>W. J. Hehre, L. Radom, P. v.R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986). *B*

<sup>3</sup>M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984). *J*

<sup>4</sup>J. J. P. Stewart QCPE **455** *6*, 1986. *J*

preprint

The use of density functional theory for the computation of the properties of real materials quantum-mechanically has been steadily gaining in popularity in the last two decades. This approach relies on the density functional theorem<sup>5</sup>, which demonstrates that a one-to-one mapping between ground-state electronic charge densities and total energies exists. Therefore, it is possible in principle to insert a charge density into a functional which returns the total energy corresponding to that charge density as accurately as if the many-electron Dirac or Schrodinger equation had been solved. However, the density functional theorem only proves the existence of such a functional and does not provide its form explicitly. A number of tractable approximations to the density functional have been formulated. The earliest and simplest, Thomas-Fermi theory, proposes a very simple density functional to replace the kinetic energy term of the Schrodinger equation. It is just not accurate enough for study of real materials.

The Kohn-Sham density functional<sup>6</sup> is more complicated and realistic than the Thomas-Fermi method. The density functional calculations reported in this thesis are based on the Kohn-Sham approach. The kinetic energy corresponding to a given charge density is written as the kinetic energy that a system of non-interacting electrons which have the given charge density would have. The interaction of the external potential with the charge density is included straightforwardly, and a Hartree term to include the bulk of the interactions between the electrons is also added. To this a final term, called the exchange-correlation energy, is added. If this term could be determined exactly, the Kohn-Sham density functional would yield exact total energies in the Born-Oppenheimer approximation. But, as mentioned above, this term is not known exactly.

Several approximations to the exchange-correlation term have been developed for the Kohn-Sham equation, and improvements to this term are an active area of research.

---

<sup>5</sup>P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964). J

<sup>6</sup>W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965). J

One of the most successful density functional exchange-correlation terms<sup>7</sup> parametrizes a series of quantum Monte Carlo calculations performed<sup>8</sup> on electrons in a uniform external potential. Because the external potential is uniform for these calculations, the parametrized density functional can be taken to be local in real space. Thus, use of a potential of this type constitutes the local density approximation (LDA). LDA calculations give the same results as quantum Monte Carlo results when the external potential is uniform. This suggests a natural extension to LDA, namely the addition of terms to the exchange-correlation functional which depend on the gradient of the density<sup>9</sup>, to provide more accurate results when the density is changing rapidly. These methods appear to have considerable merit for molecular systems. Another extension to LDA is the local spin density approximation (LSD)<sup>10</sup>, which includes the effects of spin into the calculation.

Despite the physical reasonableness of the Kohn-Sham equation, the use of approximations to the exchange-correlation energy functional is an uncontrolled approximation whose effect it is difficult to evaluate from first principles. The reason that density functional theory in general and LDA in particular have become widespread in their use is that countless calculations have been performed and the results give good agreement with experiment for a wide variety of systems. For example, LDA calculations of lattice parameters of crystals can generally be relied upon to fall within 1-2% of experiment. This level of accuracy, although significantly lower than quantum Monte Carlo and CI, is comparable with perturbation-theory-corrected Hartree Fock and much higher than Hartree Fock or semi-empirical methods. In addition, density functional calculations scale much better with problem size than any of the methods which provide similar accuracy. Accordingly, the electronic structure calculations discussed in this thesis were all performed

---

<sup>7</sup>J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

<sup>8</sup>D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

<sup>9</sup>For current examples, see A. D. Becke, Phys. Rev. A **38**, 3098 (1988); J. P. Perdew, Phys. Rev. B **33**, 8822, (1986).

<sup>10</sup>U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).

within the LDA Kohn-Sham formalism.

A number of implementations of LDA exist, and each has its advantages. The linear augmented plane wave method<sup>11</sup> (LAPW) and the linear muffin-tin orbital method<sup>12</sup> (LMTO) have been very successful for small to moderate size systems. For investigations of the behavior of the core electrons in solids, the use of one of these approaches is mandatory. Increasingly, however, the calculations on systems containing large numbers of atoms are being performed with the pseudopotential method<sup>13</sup>. This approach uses the observation that core electrons react very slightly to changes in an atom's environment, whereas valence electrons undergo significant changes. The pseudopotential replaces the Coulomb potential of the nucleus and the core electrons, and it is designed so that the ground eigenstates of the pseudopotential (pseudowavefunctions) closely resemble the valence eigenstates of the Coulomb potential. In most pseudopotential implementations<sup>14,15</sup>, the pseudopotential is chosen so that the pseudowavefunctions are nodeless and match the valence eigenstates of the Coulomb potential outside the core region, for distances from the nucleus greater than some value  $r_c$ . The pseudowavefunctions and all-electron eigenstates also have the same eigenvalues. Use of the pseudopotential method allows one to avoid recalculating the core eigenstates in the solid, and this is a significant advantage of the pseudopotential method.

Many basis sets have been used with pseudopotentials, including localized orbitals<sup>16</sup>, mixed bases of plane waves and localized orbitals<sup>17,18</sup>, and finite elements on variable grids<sup>19</sup>. In particular, the plane-wave pseudopotential method<sup>20</sup> currently enjoys

---

<sup>11</sup>L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **33**, 823 (1986). *J*

<sup>12</sup>O. K. Andersen, Phys. Rev. B **12**, 3060 (1975). *J*

<sup>13</sup>J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287, 880 (1959). *J*

<sup>14</sup>D. R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979). *J*

<sup>15</sup>L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982). *J*

<sup>16</sup>J. R. Chelikowsky and M. Y. Chou, Phys. Rev. B **38**, 7966 (1988). *J*

<sup>17</sup>M. H. Kang, R. C. Tatar, E. J. Mele and P. Soven, Phys. Rev. B **35**, 5457 (1987). *J*

<sup>18</sup>S. G. Louie, K. -M. Ho and M. L. Cohen, Phys. Rev. B **19**, 1774 (1979). *J*

<sup>19</sup>J. Bernholc, to be published. *X*

<sup>20</sup>J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C: Solid State Phys. **12**, 4409 (1979). *J*

widespread popularity. This is because the plane-wave basis set has a number of valuable computational features. It is very straightforward to improve a finite plane-wave basis set systematically. Another advantage is that the set of all plane waves is complete, whereas approaches which include localized orbitals can have problems with over-completeness in certain implementations. Also, the overlap integrals between plane-wave basis functions are simple to calculate. Furthermore, the simplicity of plane waves allows the use of fast Fourier transform algorithms and iterative diagonalization techniques<sup>21,22,23</sup> to enhance efficiency. As a result, the plane-wave pseudopotential method is currently the most efficient way to study the structural and electronic properties of complicated real materials from first principles.

The first half of the thesis addresses the efficiency of pseudopotentials within the plane-wave pseudopotential method. Chapter 1 explains the features of a pseudopotential which predict its efficiency for plane-wave pseudopotential calculations, and it outlines a procedure for the creation of efficient optimized pseudopotentials for any element which require as few basis functions as possible. Chapter 1 also contains a demonstration of the efficiency gain when this procedure is followed. Chapter 2 describes an augmentation to the procedure in which the scattering properties of the pseudopotential and its efficiency can be simultaneously improved. Chapter 3 explains the application of these optimized pseudopotentials to a wide variety of organic molecules. Chapter 4 shows the application of the optimized pseudopotentials to a gold grain boundary. Neither organic molecules nor transition metals had been treated with plane waves and first-principles pseudopotentials prior to this work. The development of the optimized pseudopotentials improved the efficiency of the plane-wave pseudopotential method sufficiently to allow these systems to be addressed.

---

<sup>21</sup>R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985). J

<sup>22</sup>M. P. Teter, M. C. Payne, and D. C. Allan, Phys. Rev. B **40**, 12255 (1989). J

<sup>23</sup>M. J. Gillan, J. Phys.: Condensed Matter **1**, 689 (1989). J

One future direction for research is the application of optimized pseudopotentials to other classes of systems which are not tractable without these pseudopotentials. Work is already ongoing to study cuprate superconductors and carbon and gold defects in silicon. A more distant goal is to revamp the pseudopotential formalism to make the pseudopotential approximation more systematically improvable.

Maxwell's equations provide an exact description of the electromagnetic behavior of materials. The only difficulty with their use is that to provide the positions of all the electronic charges in a real material, one would have to solve the electronic system quantum-mechanically before beginning an electromagnetic analysis, with all the problems discussed above. However, for electromagnetic radiation with a wavelength much longer than the interatomic spacing, this exact specification of charges is neither necessary nor desirable. The electric and magnetic fields can be averaged over a region large compared to the interatomic spacing but very small compared to the wavelength. The response of the materials to electric and magnetic fields can be expressed as a scalar dielectric constant and scalar diamagnetic constant. For many materials this yields essentially exact results. For other materials, called nonlinear materials, these scalars must be supplemented with higher-order tensor terms.

Recently, attention has been focused on materials which have periodically modulated dielectric constants, called photonic crystals. For these periodic or nearly-periodic systems, a plane-wave expansion of the solutions to Maxwell's equations in macroscopic media brings the advantages discussed above. Theoretical<sup>24</sup> and experimental<sup>25</sup> results have revealed structures made of materials with two different dielectric constants arranged in a 3D periodic array which have a range of frequencies for which no extended electromagnetic modes exist, called a photonic band gap. Chapter 5 explains the theoretical prediction that the introduction of a defect into such a structure can

---

<sup>24</sup>K. M. Ho, C. T. Chan, C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990).

<sup>25</sup>E. Yablonovitch and T. J. Gmitter, Phys. Rev. Lett. **63**, 1950 (1989). *J*

give rise to an exponentially localized electromagnetic state without the use of any metallic shielding. Chapter 6 reports an experimental confirmation of such localized photon modes, and compares experimental results with theoretical calculations. Chapter 7 discusses the development of materials with 2D finite symmetry which possess an in-plane photonic band gap. It also compares theoretical predictions and experimental results for one particular structure. Chapter 8 is a detailed experimental confirmation of the predictions pertaining to 2D photonic crystals. Chapter 9 discusses the prediction of the existence of surface modes on photonic crystals under certain circumstances.

A major focus of future work in this area is the study of more complicated electromagnetic systems. Work is in progress to recast the Maxwell's equations in a form which requires significantly less computer time and memory to achieve results. In addition, I plan to continue to study these photonic crystals theoretically and experimentally. Future directions will include the modeling of systems containing metal and dielectric, dissipative systems, and systems containing nonlinear materials.

**Chapter 1: Optimized Pseudopotentials  
in collaboration with  
K. M. Rabe and E. Kaxiras**

Theoretical studies based on *ab initio* electronic structure calculations are useful for elucidating many materials problems in the physics of solids. Specialized basis sets (i.e. other than plane waves) have proven to be effective for calculations of total energies and band structures of solids.<sup>1</sup> The use of such a set is vital if the behavior of the core electrons of the solid is to be considered explicitly. However, the behavior of the core electrons frequently has a minimal effect on properties of interest in the solid, and the implicit inclusion of the core states into a pseudopotential opens up the possibility of using a plane-wave basis for electronic structure calculations. This choice of basis has many advantages for a wide range of calculations, including studies of forces on ions<sup>2</sup>, phonon frequencies, corrections to quasiparticle self-energies<sup>3</sup>, and dielectric matrices.<sup>4</sup> The plane-wave basis provides controlled convergence and high numerical accuracy. In addition, using this basis for the single-particle states allows for a convenient and unbiased representation of the charge density. Furthermore, when a plane-wave basis is used, the equation-of-motion method of Car and Parrinello<sup>5</sup> can be implemented to increase calculational efficiency tremendously. The method is especially efficient when used with the excellent fast Fourier-transform algorithms which are currently available.

The main drawback to using plane waves as a basis is that real materials require many basis functions for their description. The problem of convergence becomes critical for materials which have sharply peaked valence states. In describing these substances, a prohibitive number of plane waves would be required, making the use of the conventional plane-wave approach impossible. As a result, almost all *ab initio* calculations involving transition metals and first-row nonmetals have been performed with specialized basis sets.<sup>1</sup>

---

<sup>1</sup>For example, see L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **33**, 823 (1986); C. T. Chan, D. Vanderbilt, S. G. Louie, and J. R. Chelikowsky, Phys. Rev. B **33**, 7941 (1986); M. H. Kang, R. C. Tatar, E. J. Mele, and Paul Soven, Phys. Rev. B **35**, 5457 (1987); O. K. Andersen, Phys. Rev. B **12**, 3060 (1975), and references therein.

<sup>2</sup>J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).

<sup>3</sup>R. W. Godby, M. Schluter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

<sup>4</sup>M. S. Hybertsen and S. G. Louie, Phys. Rev. B **37**, 2733 (1988).

<sup>5</sup>R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

Optimizing the convergence with basis size of quantities of interest, including total energy, would allow the benefits of the plane-wave basis to be derived in as wide a class of systems as possible. As a result, several approaches have been proposed to achieve this goal. Most of these<sup>6,7,8</sup> focus on improving the norm-conserving pseudopotentials<sup>9</sup> which have been very widely used in the literature to address problems in the theory of solids. Each of these approaches was either designed for a specific case<sup>6</sup> or did not significantly affect convergence of total energies.<sup>7,8</sup> The purpose of this paper is to isolate those features of a pseudopotential which are responsible for the convergence of the total energy in a plane-wave basis. Having done that, we then formulate a criterion which can be used to create a pseudopotential with optimal convergence and which should be applicable to any element.

We have found that the convergence of the total energy of solids with plane-wave basis size mirrors the convergence of the total energies of the isolated pseudoatoms which comprise the solid. In physical terms, we can understand this correspondence by dividing materials into four classes from the point of view of convergence. In the first class, valence states in the atom and solid are very similar. For this class, which includes noble gas solids and ionic solids, correspondence between atomic and solid convergence should be excellent. The second class contains solids in which the valence electrons are distributed nearly uniformly throughout the solid, such as free-electron metals. For these materials, the solid valence states will require many fewer plane waves for convergence than the atomic states. Consequently, they do not present a convergence problem, and their convergence does not need to be optimized. The third class contains materials such as transition metals which combine characteristics of the first two classes. In this class, some valence states in the solid are free-electron-like and do not present a convergence problem,

---

<sup>6</sup>For example, see Y. Bar-Yam, S. T. Pantelides, and J. D. Joannopoulos, Phys. Rev. B **39**, 3396 (1989).

<sup>7</sup>For example, see D. Vanderbilt, Phys. Rev. B **32**, 8412 (1985), and references therein. *J*

<sup>8</sup>L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982). *J*

<sup>9</sup>D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979). *J*

and others are highly peaked around each atomic site and resemble the corresponding atomic eigenstates strongly. Because the latter states control convergence and because they are similar in the atom and solid, materials in the third class have convergence properties which correspond closely to those of the constituent atoms. In the fourth class are the systems which exhibit covalent bonding. Here, no direct relationship between atom and solid exists, since the states in the solid do not resemble the atomic states. However, the bonding states in the solid are confined in interatomic regions of approximately the same volume as the atomic states. Therefore, we expect convergence of atomic and solid total energies to be similar for covalent systems as well.

We have checked these claims and found close correspondence between the convergence of the atomic and solid total energies for a number of elements which are particularly problematic to treat with standard plane-wave approaches because of slow convergence. Specifically, we examined ionic, transition metal, and covalent systems including oxygen, copper, and carbon. All the results are qualitatively very similar. Below we discuss in detail and exhibit results for copper, which is about as difficult a case as exists for the plane-wave method.

The fact that atomic and solid total energies do converge similarly for all solids for which convergence is a problem presents a general criterion for the generation of optimally convergent pseudopotentials: to make the total energies of solids converge optimally, it is sufficient in principle to create pseudopotentials which give optimal total-energy convergence for the constituent atoms. In practice, we implement this criterion with two approximations; we require the potential to be continuous, and we optimize only the atomic kinetic-energy convergence. It can be proven<sup>10</sup> using scaling arguments that total-energy convergence and kinetic-energy convergence are very similar in the limit of large cutoff energies. By requiring the pseudopotential to be continuous, we lower the cutoff energy

---

<sup>10</sup>A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos (unpublished).

above which this is true. In fact, we have found kinetic-energy convergence to mirror total-energy convergence for all continuous pseudopotentials which we have examined, for all cutoff energies of interest. One or both of these approximations may be eliminated to obtain marginally better results. The method we have chosen has the merits of straightforwardness of implementation and numerical stability.

Our goal is to create an atomic pseudowavefunction which has a continuous pseudopotential and kinetic-energy convergence to a certain level of tolerance in as small a basis as possible. A pseudowavefunction gives rapid kinetic-energy convergence when the high Fourier components of the pseudowavefunction contain very little kinetic energy. If one simply minimizes the kinetic energy of the pseudowavefunction, as proposed by Cohen and Heine,<sup>11</sup> no control is exercised over the *distribution* of the remaining kinetic energy among the Fourier components of the pseudowavefunction. To achieve optimal convergence, the kinetic energy in the high Fourier components of the pseudowavefunction must be minimized directly. To accomplish this, it is sufficient to minimize the portion of the kinetic energy in the atomic pseudowavefunction above some cutoff  $q_c$  in Fourier space while ensuring that the pseudowavefunction has two continuous derivatives. We can begin with any pseudowavefunction  $F(r)$  which matches the atomic eigenfunction smoothly at  $r=r_c$ . To it we add a correction function  $C(r)$ , which optimizes the convergence of the original wavefunction. The result is  $\Psi$ , the optimized wavefunction, which equals  $F+C$  for  $r \leq r_c$  and which equals the all-electron wavefunction for  $r > r_c$ . We find it efficient and convenient to have an analytic form for  $F$ . In practice, we write  $F$  as a sum of four Bessel functions whose wave vectors  $q_i'$  are chosen so that their logarithmic derivatives match that of the all-electron wavefunction,  $\phi_l(r)$ , at  $r_c$ :

$$F(r) = \sum_{i=1}^4 \alpha_i j_i(q_i' r) ; \frac{j_i'(q_i' r_c)}{j_i(q_i' r_c)} = \frac{\phi_i'(r_c)}{\phi_i(r_c)} \quad (1)$$

---

<sup>11</sup>M. H. Cohen and V. Heine, Phys. Rev. 122, 1821 (1961).

The coefficients of these four functions are chosen to normalize the wavefunction and to make it continuous with two continuous derivatives at  $r_c$ . Because normalization is a quadratic constraint on the  $\alpha_i$ 's, a sum of three Bessel functions cannot always be found which satisfies these constraints. Accordingly, a fourth Bessel function is included whose coefficient is chosen arbitrarily within the range of values which permits the constraints to be met. To the expression  $F$ , we add  $C(r)$ , which we expand in Bessel functions whose wave vectors  $q_i$  are chosen so that the functions have a node at  $r_c$ .

$$C(r) = \sum_{i=1}^N \beta_i j_i(q_i r) ; j_i(q_i r_c) = 0 \quad (2)$$

Lagrange multipliers enforce the constraints of normalization and continuity of two derivatives at  $r_c$ , and the coefficients  $\beta_i$  of these Bessel functions are chosen to minimize the kinetic energy beyond the cutoff  $q_c$ , which is expressed as follows:

$$-\int_0^\infty d^3r \Psi_l^*(r) \nabla^2 \Psi_l(r) - \int_0^{q_c} d^3q q^2 |\Psi_l(q)|^2 \quad (3)$$

The pseudopotential is calculated by directly inverting the Schrodinger equation.

This scheme for generating pseudopotentials has two parameters which may be chosen differently depending upon the application. The parameter  $r_c$  may be varied to insure transferability of the pseudopotentials. For any given  $r_c$ , our method provides the wavefunction which converges most rapidly. The other parameter,  $q_c$ , is determined iteratively. A  $q_c$  is chosen, and the amount of kinetic energy beyond  $q_c$  is minimized. Then  $q_c$  is varied until exactly the tolerated quantity of kinetic energy is found beyond  $q_c$  after minimization.

We now illustrate this approach for the case of copper. We begin with an all-electron calculation of an atom of  $Cu^{1+}$  ( $3d^9 4s^{0.75} 4p^{0.25}$ ). We pick the cutoff radius

$r_c=1.97 a_0$  to be smaller than half the interatomic spacing. We optimize convergence for a tolerance of 1 mRy, and this leads to  $q_c=7.14$ . In the expression for F,  $\alpha_4$  is fixed equal to 0.5. Five Bessel functions are included in the correction function C for the potential we present. (Extending the set of Bessel functions in C up to twenty gave only marginal improvement of convergence.) The coefficients and wave vectors for F and C are contained in Table I.

---

Function F			Function C		
i	$\alpha_i$	$q_i'$	i	$\beta_i$	$q_i$
1	1.619452	2.278679	1	0.203543	2.926964
2	2.436893	3.921348	2	-0.448616	4.618887
3	1.744898	5.536289	3	-0.827052	6.258186
4	0.500000	7.142447	4	-0.169339	7.879067
			5	0.016011	9.491198

---

TABLE I. Expansion of the optimized d pseudowavefunction for  $r < r_c$ . Coefficients and wave vectors of the spherical Bessel functions of angular momentum 2 are provided in atomic units for functions F and C, which are defined in the text.

As a test of the improvement achieved, we compare the convergence of our optimized potential to a pseudopotential created by the standard method of Hamann, Schluter, and Chiang (HSC).<sup>9</sup> We optimized only the d pseudopotential, because the d electrons clearly govern convergence, even after optimization. The s and p wavefunctions have  $r_c$ 's of 2.6  $a_0$ . We choose the HSC d wavefunction so that the HSC and new wavefunctions match the all-electron wavefunction exactly for the same range of radius,  $r_c=1.97 a_0$ . The d wavefunctions of these two potentials and their Fourier transforms are compared in Fig. 1. Notice that although the two pseudowavefunctions are quite similar in real space, the new wavefunction has a Fourier transform which converges much more

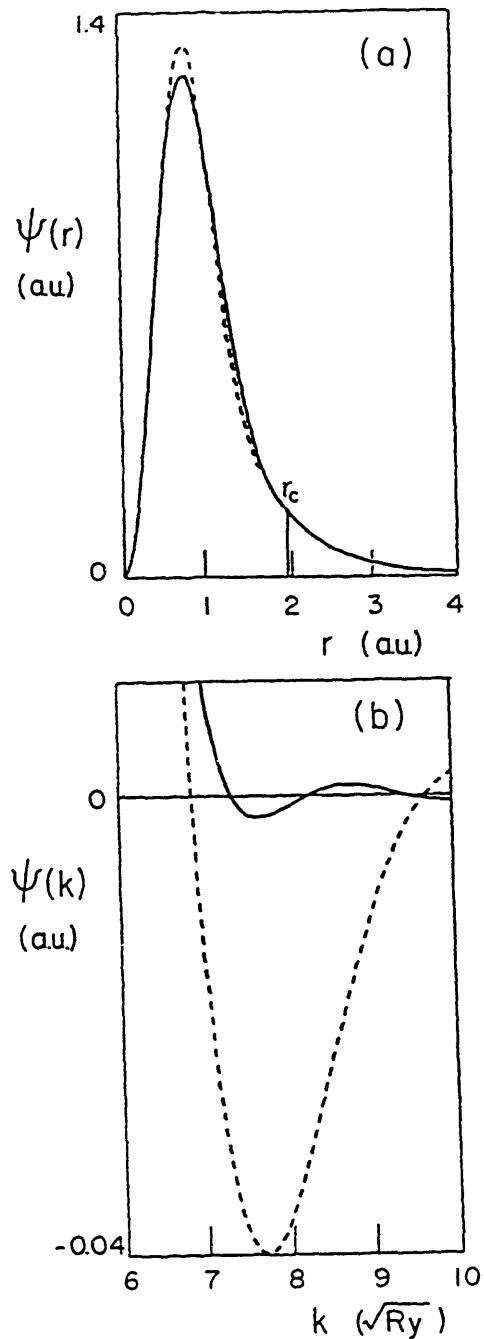


FIG. 1. Pseudowavefunctions (a) in real space and (b) in Fourier space, for the copper 3d eigenstate, generated by the HSC method (dashed line) and by the present approach (solid line). The pseudowavefunctions are normalized to unity.

rapidly. The ionic pseudopotentials which give rise to these wavefunctions are compared in Fig. 2. The new potential is deeper and somewhat less convergent in Fourier space than the HSC potential. If potential well-depth or potential Fourier-transform convergence were used as criteria<sup>7</sup>, the HSC potential would be considered preferable to the optimized potential. However, the criterion of atomic total-energy convergence indicates that the optimized potential is far superior. In Fig. 3, we show by solid lines the atomic total energy as a function of wave vector for both pseudopotentials. This calculation predicts that the total energy of solid copper will converge to within 1 mRy at a plane-wave cutoff energy of 50 Ry (500 plane waves for fcc copper) with the new pseudopotential, whereas the same level of convergence for the HSC case requires a cutoff of 115 Ry (1700 plane waves for fcc copper).

In order to verify the predictions of the atomic calculations, these pseudopotentials were then used to calculate the total energy of fcc copper at the experimental lattice constant. For these calculations we used 256  $\mathbf{k}$  points in Brillouin zone (10 in the irreducible sector).<sup>12</sup> The results are plotted as dots in Fig. 3. The close correlation between atomic and solid total-energy curves both for the standard HSC potential and for our new potential shows that the criterion of atomic total-energy convergence does indeed have great predictive power in general. In addition, the fact that a potential which was created to optimize atomic total-energy convergence also exhibits a large improvement in convergence of the solid total energy, shows that we have gone beyond identifying a valid criterion and have produced and exhibited a potential which possesses markedly improved convergence.

Using this new pseudopotential, we performed the first *ab initio* plane-wave basis determination of the structural properties of a noble metal. We computed the total energy of fcc copper at 15 atomic volumes ranging from a 27% compression to a 34% expansion

---

<sup>12</sup>H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976). J

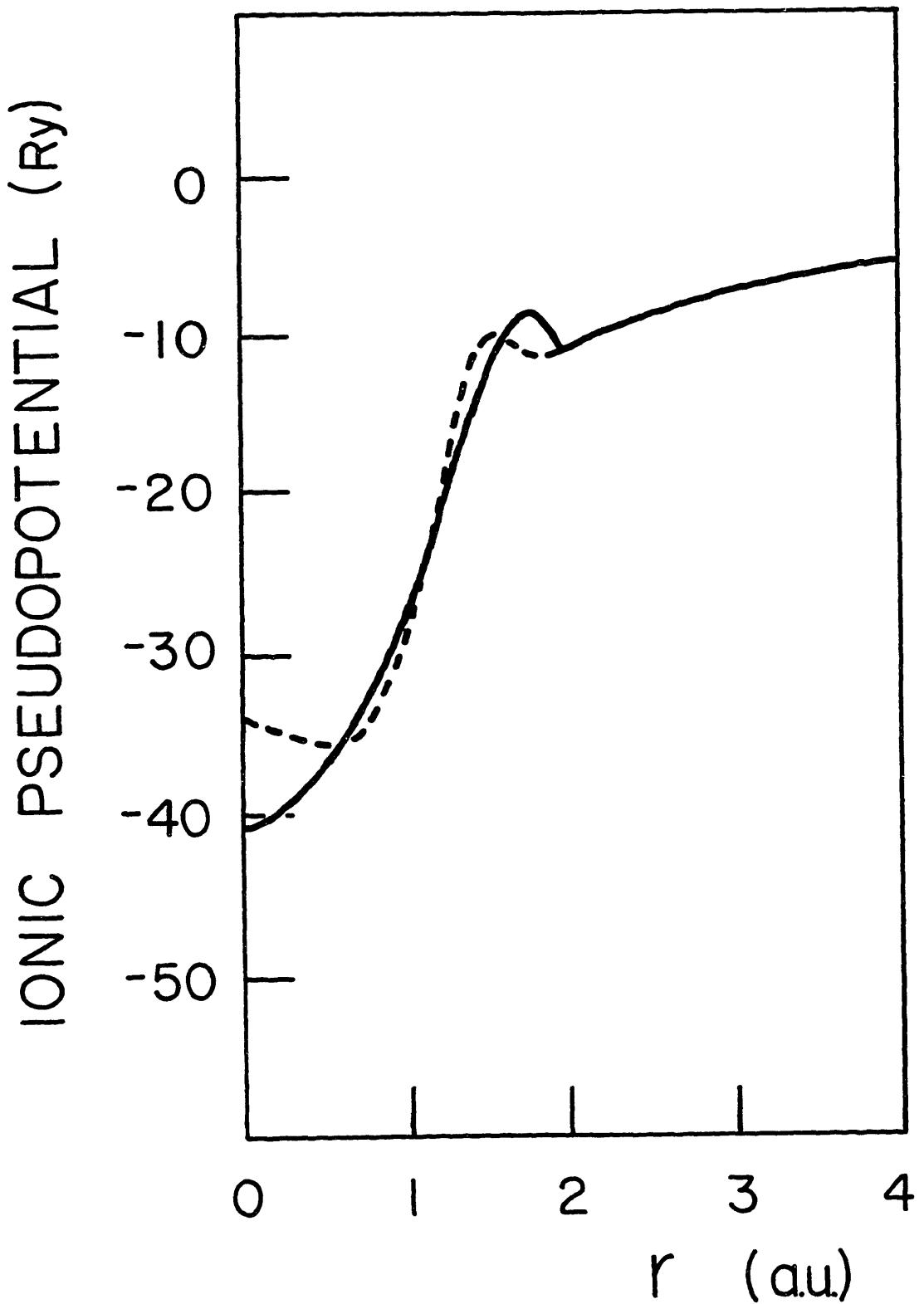


FIG. 2. Pseudopotentials for the copper 3d eigenstate using the HSC method (dashed line) and the present approach (solid line).

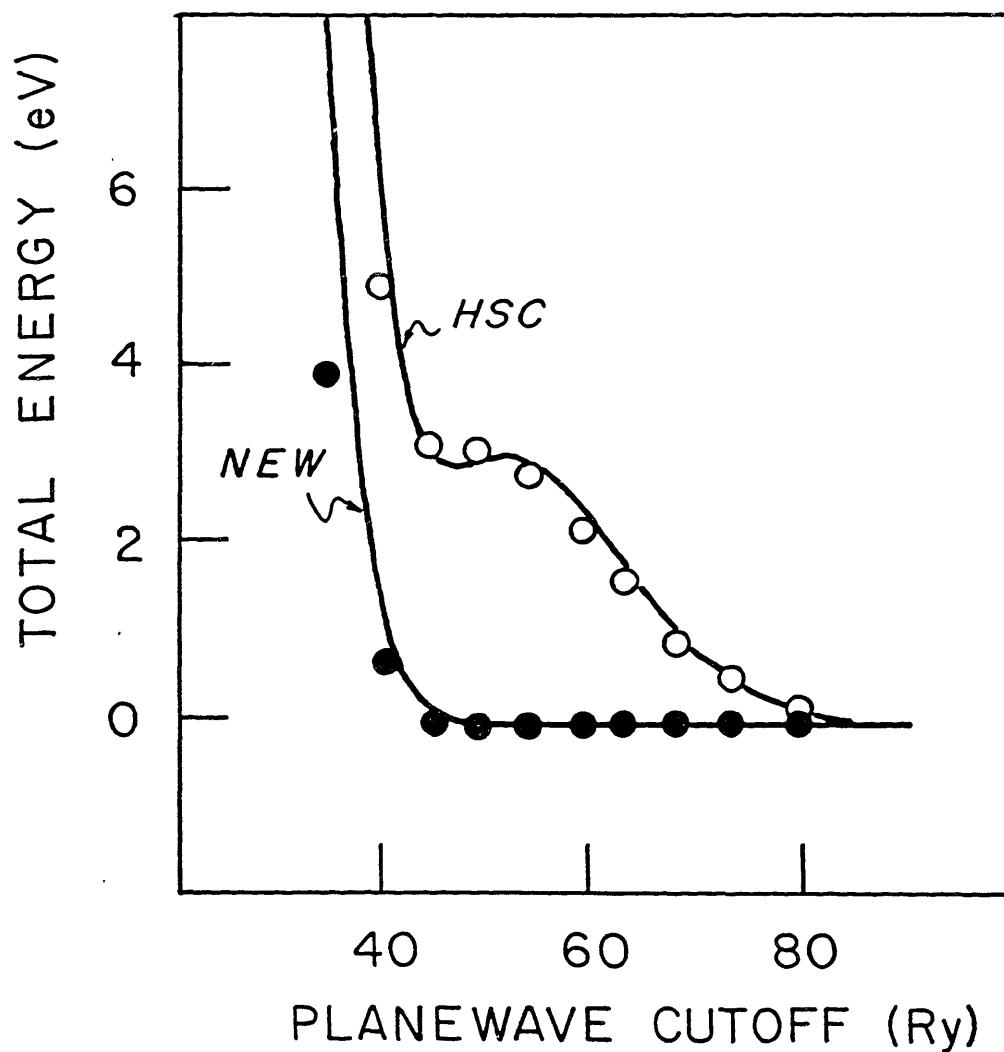


FIG. 3. Atomic (solid lines) and fcc solid (dots) total energies as a function of cutoff energy for copper in the HSC and present approaches. The zero of atomic total energy for each pseudopotential was chosen to be the total atomic energy at a cutoff energy of 324 Ry. The zero of solid total energy was chosen for each pseudopotential so that the atomic and solid total energies coincide at a cutoff energy of 80 Ry.

relative to the calculated equilibrium volume. For these calculations, we used 1372  $\mathbf{k}$  points in the Brillouin zone (44 in the irreducible sector) and a cutoff energy of 50 Ry, the value predicted by the atomic calculation. The enlarged  $\mathbf{k}$  point set was essential to describe the shape of the Fermi surface realistically. Using 50 Ry as a cutoff energy insures that all total energies are converged to within 1 mRy. This tolerance is an upper bound to the convergence of all energy differences, and energy differences between similar structures will probably be converged to a much more stringent tolerance. Accordingly, a much smaller cutoff energy may be used for many applications. The equilibrium lattice constant, cohesive energy, and bulk modulus of fcc copper which result from these calculations agree very well with experiment. Experimental and calculated bulk properties of fcc copper are compared in Table II.

	Experiment	Present	% Difference
Lattice constant (a. u.)	6.81	6.84	+ 0.4
Cohesive energy (Ry/atom)	0.257	0.292	+13.6
Bulk modulus (MBar)	1.42	1.63	+ 14.8

TABLE II. Comparison of the bulk properties of fcc copper as obtained by the present method of calculation to experimental values.

In conclusion, we have addressed the problem of improving convergence in a plane-wave basis. We identify the essential features of a pseudopotential which determine its convergence properties. From this, we develop a criterion which we use to create optimally convergent pseudopotentials. Using this principle, it should be possible to optimize the convergence characteristics for any atom, without sacrificing accuracy. This makes it possible to scrutinize solids containing transition metals and first-row nonmetals with the plane-wave pseudopotential formalism. As a result of this work, we envision that

the simplicity and versatility of calculations using a plane-wave basis can now be extended to a wide range of ionic, covalent, and transition metal systems.

We thank Dr. Oscar Alerhand for stimulating discussions. This work was supported in part by the Office of Naval Research Contract No. N000 14-86-K-0158 and the U. S. Joint Services Electronics Program Contract No. DAAL-03-86-K-0002. One of us (A. M. R.) would like to acknowledge partial financial support from the Office of Naval Research.

**Chapter 2: The Design of Convergent and Transferable *Ab Initio*  
Pseudopotentials**

The sharply-peaked atomic wavefunctions of the first-row nonmetals and transition metals are very difficult to describe in a plane-wave basis set. For this reason, specialized basis sets have been the method of choice for most *ab initio* descriptions of materials containing these elements. Because the plane-wave pseudopotential method has many positive features, there is great interest in developing pseudopotentials which have sufficiently good convergence and transferability characteristics to allow the study of these "problem" elements.

Our work<sup>1</sup> has demonstrated that a key element in the creation of tractable pseudopotentials is the control of the convergence of the total energy of the free atom described by a pseudopotential, i.e. the pseudoatom. This is because the convergence of a pseudoatom's total energy and that of solids containing the pseudoatom are remarkably similar in a wide range of materials, as discussed in Chapter 1. Furthermore, our work shows that the convergence of the kinetic energy of the pseudoatom is an excellent predictor of the convergence of the total energy of the pseudoatom, and therefore of the total energy of solids containing that pseudoatom as well.

The convergence of the kinetic energy of the pseudowavefunction is a powerful predictor of the convergence characteristics of the pseudopotential in solid-state environments because its validity is independent of the method chosen for insuring transferability. To illustrate this fact, we demonstrate the use of this atomic criterion to generate high-quality pseudopotentials within three different pseudopotential formalisms. First, the use of the criterion with simple norm conservation is illustrated for the case of copper. Next, the criterion is used in conjunction with a transferability criterion involving explicit optimization of logarithmic derivatives at several energies. Finally, the use of the criterion within the nonlocal generalized eigenvalue formalism of Vanderbilt<sup>2</sup> is discussed.

---

<sup>1</sup>A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990). 

<sup>2</sup>D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990). 

Because the convergence of the kinetic energy of the pseudowavefunction mirrors the convergence characteristics of the pseudopotential in the solid, it is desirable to create a pseudowavefunction with rapid kinetic-energy convergence. Here, we demonstrate a particularly simple way to achieve this goal. We create a pseudowavefunction whose Fourier transform possesses a minimal amount of kinetic energy beyond a cutoff  $q_c$ . This will provide good convergence using a plane-wave cutoff energy equal to  $(q_c)^2$ , using Rydberg energy units. This wavefunction also matches the all-electron wavefunction outside a real-space cutoff radius  $r_c$ , conserves norm, and has the same eigenvalue as the all-electron reference state.

The method of optimizing the pseudowavefunction described here gives the same results as the method of Chapter 1. This is because the same quantity is minimized in both cases, and the same constraints are applied. Here, a simpler basis set is used to span the space of possible pseudowavefunctions, and this makes the present method conceptually and computationally simpler. The wavefunction is constructed as follows. Inside  $r_c$ , the wavefunction  $\psi_l$  is expanded in a complete set of spherical Bessel functions of the proper angular momentum whose logarithmic derivatives at  $r_c$  match that of the reference wavefunction  $\phi_l$ .

$$\psi_l(r) = \theta(r_c - r) \sum_{i=1}^N \alpha_i j_l(q_i r) + \theta(r - r_c) \phi_l(r) \quad (1)$$

where  $N \sim 10$  and the  $q_i$  are wave vectors chosen to match

$$\frac{j_l'(q_i r_c)}{j_l(q_i r_c)} = \frac{\phi_l'(r_c)}{\phi_l(r_c)} \quad (2)$$

The kinetic energy  $KE_{q>q_c}$  of the wavefunction beyond  $q_c$  is conveniently expressed as the difference between the kinetic energy in real space and the kinetic energy from 0 to  $q_c$  in Fourier space.

$$KE_{q>q_c} = - \int_0^\infty d^3r \psi_l^*(r) \nabla^2 \psi_l(r) - \int_0^{q_c} d^3q q^2 |\psi_l(q)|^2 \quad (3)$$

This is the quantity we wish to minimize. All of the integrals remain the same throughout the calculation and can therefore be tabulated. Defining the following integrals,

$$A_i = \int_0^{r_c} dr r^2 [j_l(q_i r)]^2 \quad (4)$$

$$B_i(q) = \sqrt{\frac{2}{\pi}} \int_0^{r_c} dr r^2 j_l(q_i r) j_l(q r) \quad (5)$$

$$C(q) = \sqrt{\frac{2}{\pi}} \int_{r_c}^\infty dr r^2 j_l(q r) \phi_l(r) \quad (6)$$

$$D_{ij} = \int_0^{q_c} dq q^4 B_i(q) B_j(q) \quad (7)$$

$$E_i = \int_0^{q_c} dq q^4 B_i(q) C(q) \quad (8)$$

$$F = \int_{r_c}^\infty dr r^2 \phi_l^*(r) \nabla^2 \phi_l(r) \quad (9)$$

$$G = \int_0^{q_c} dq q^4 [C(q)]^2 \quad (10)$$

the expression for  $KE_{q>q_c}$  may be rewritten

$$KE_{q>q_c} = \left[ \left( \sum_{i=1}^N \alpha_i^2 q_i^2 A_i \right) - F \right] - \left[ \left( \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j D_{ij} \right) + 2 \sum_{i=1}^N \alpha_i E_i + G \right] \quad (11)$$

We now wish to minimize  $KE_{q>q_c}$  subject to the constraints of norm conservation and wavefunction smoothness at  $r_c$ . This minimization can be performed rapidly by various

scientific computer library minimization subroutines. The calculation takes us about five CPU seconds on a Cray-2.

The application of this procedure to copper greatly improves the convergence properties of the total energy for both a single copper pseudoatom and for copper fcc solid relative to the same calculations performed using a standard norm-conserving pseudopotential of the Hamann, Schluter, Chiang type (HSC)<sup>3</sup> in exactly the same way that the method described in Chapter 1 did. Table 1 compares the bulk properties of fcc copper as calculated using the new pseudopotential to the results of an LAPW calculation<sup>4</sup> and to experimental values. The comparison to experiment must be taken simply as a context because (for simplicity) neither of the theoretical calculations includes relativistic effects, and both theoretical results include spin-polarization effects only as a correction added to the atomic total energy. Even so, the agreement between the theoretical results is quite good, suggesting that the convergence and transferability of copper are satisfactory for the study of copper (and presumably other transition metals) with these pseudopotentials. The comparison between the current pseudopotential method and LAPW results for bulk fcc copper is made in more detail in Figure 1, which shows the total energy versus volume curves obtained by each method.

---

	Experiment	This work	LAPW
Lattice constant ( $\text{\AA}$ )	3.60	3.62	3.57
Cohesive energy (eV/atom)	3.50	3.97	4.19
Bulk Modulus (MBar)	1.42	1.63	1.72

---

TABLE I. A comparison of the bulk properties of fcc copper as calculated by optimized pseudopotential and LAPW methods with corresponding experimental values.

---

<sup>3</sup>D. R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

<sup>4</sup>Z. W. Lu, S.-H. Wei, and A. Zunger, Phys. Rev. B **41**, 2699 (1990), and S.-H. Wei, private communication.

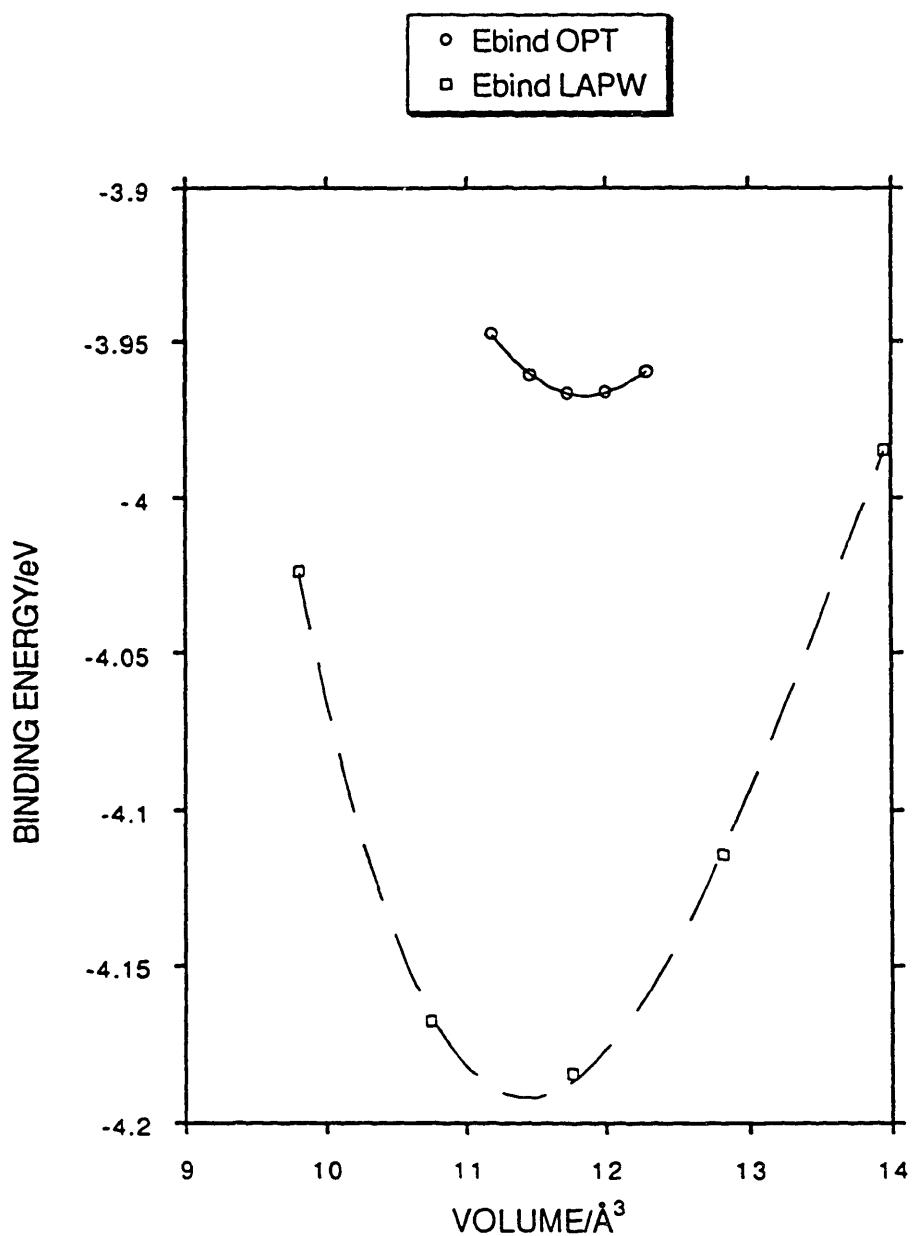


FIG. 1. Graph of energy versus volume for fcc copper as calculated by the pseudopotential method with an optimized pseudopotential and by the LAPW method.

The aforementioned procedure is most effective when the transferability needs of the material are modest. In the particular example of copper, we obtained good agreement with theory and experiment using a very large cutoff radius  $r_c$  of  $1.97 a_0$ . This is acceptable because the copper atoms remain quite far apart in fcc copper, and the d electrons do not participate strongly in bonding. In real space, one can view this as a small region of significant overlap of d wavefunctions centered on different atoms. In energy space, one can see this same phenomenon as a very small bandwidth for the d-like states. As a result of this small interaction between neighboring d orbitals, a large core sphere may be pseudized without affecting the important solid-state properties<sup>5,6</sup>, and only a very small range of energies about the valence eigenvalue need have correct logarithmic derivatives for these properties to be obtained.

For other elements, this prescription is less effective. The pseudopotential generating method described above insures a certain level of transferability due to the norm-conservation condition and the matching of the valence wavefunction outside  $r_c$ . However, for elements which participate in strong covalent bonds, a larger bandwidth results, and matching of the logarithmic derivative over a wider range is required.

Fortunately, a straightforward way of simultaneously controlling convergence and transferability by optimizing pseudowavefunction kinetic energy and logarithmic derivatives is readily available. We minimize the following expression:

$$\beta \left( - \int_0^{\infty} d^3r \psi_l^*(r) \nabla^2 \psi_l(r) - \int_0^{q_c} d^3q q^2 |\psi_l(q)|^2 \right) + (1-\beta) \left( \sum_{i=1}^M (\chi(E_i) - \gamma(E_i))^2 \right) \quad (12)$$

Here,  $\chi$  represents the logarithmic derivative for the all-electron potential as a function of energy for some choice of diagnostic radius. The function  $\gamma$  is the logarithmic derivative for the pseudopotential at the same radius.

---

<sup>5</sup>J. R. Chelikowsky and M. Y. Chou, Phys. Rev. B **38**, 7966 (1988). *J*

<sup>6</sup>Y. Bar-Yam, S. T. Pantelides, and J. D. Joannopoulos, Phys. Rev. B **39**, 3396 (1989). *J*

For beta equal to unity, the pseudopotential with the fastest convergence results. For beta equal to zero, the pseudopotential which best fits the all-electron logarithmic derivative at a series of chosen energies is found. For beta between zero and one, potentials with improved transferability *and* convergence are generated. The value of beta phenomenologically represents the relative difficulties caused by convergence and transferability error in a particular element.

The energies which were used for fitting the logarithmic derivative for the p potential in carbon are shown in Figure 2. The pseudopotential logarithmic derivative curve at the cutoff radius of  $1.4 a_0$  agrees well with the all-electron curve throughout the bandwidth of diamond. It agrees with the all-electron result better than the HSC potential used for comparison does. Figure 3 shows the superior convergence properties of the new carbon pseudopotential relative to the HSC potential. In addition, the excellent agreement between atom and solid gives further support to the validity of the atomic convergence criterion in covalent materials. Thus, the enhanced prescription for pseudopotential generation achieves improvement of convergence and transferability for a difficult test case. Note that there is a marked similarity between this procedure for beta equal zero and the procedure of Teter and Allan<sup>7</sup>. They make no explicit provision for the control of convergence characteristics, but the method for the control of transferability is quite similar.

This prescription has the advantage that it creates a single potential for each angular momentum which has nearly the best overall characteristics of convergence and transferability possible. For many elements, this allows in-depth analysis of materials which is not possible or much more difficult using standard norm-conserving pseudopotentials. We believe that little improvement over the current approach is possible within the framework of a single potential per angular momentum. The Vanderbilt

---

<sup>7</sup>M. Teter and D. C. Allan, private communication.

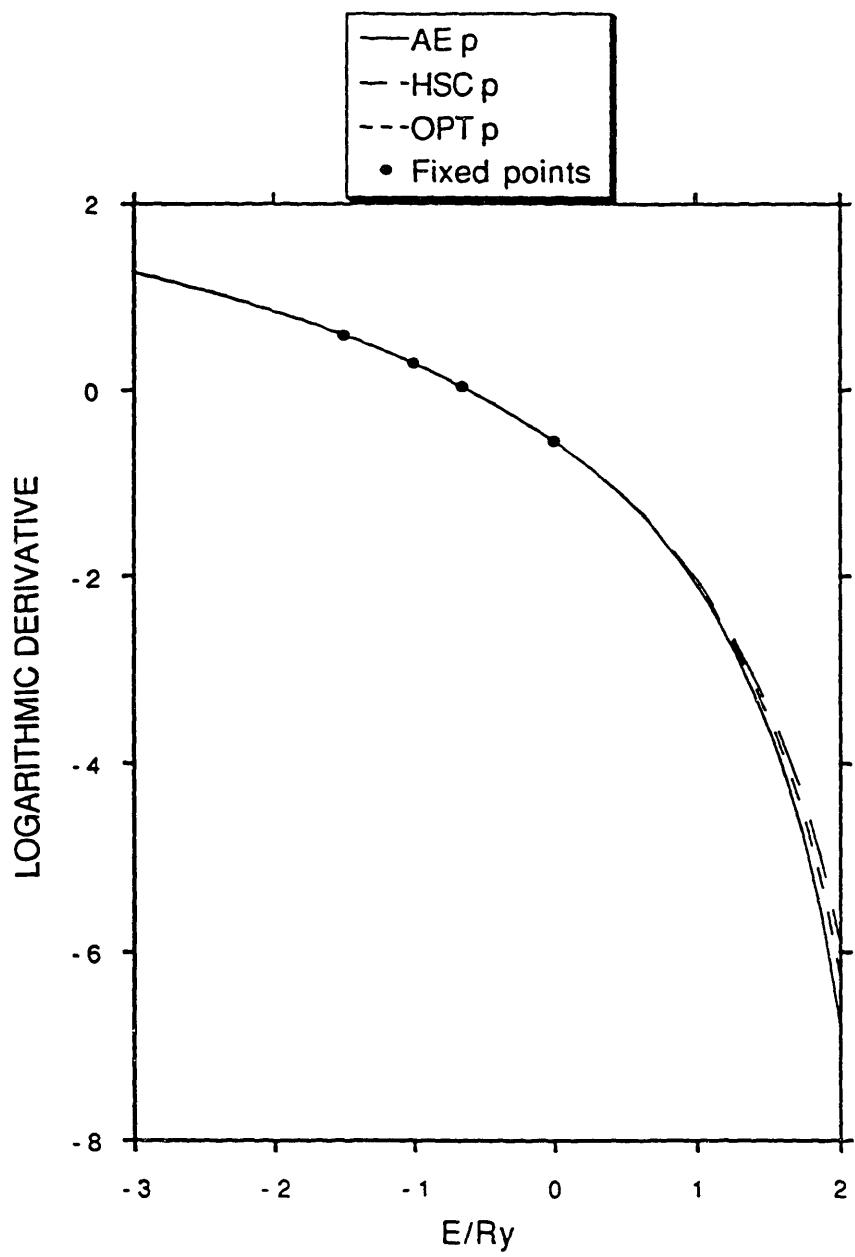


FIG. 2. Logarithmic derivatives of carbon all-electron potential and HSC and optimized pseudopotentials at  $r_{\log} = 1.4 a_0$ .

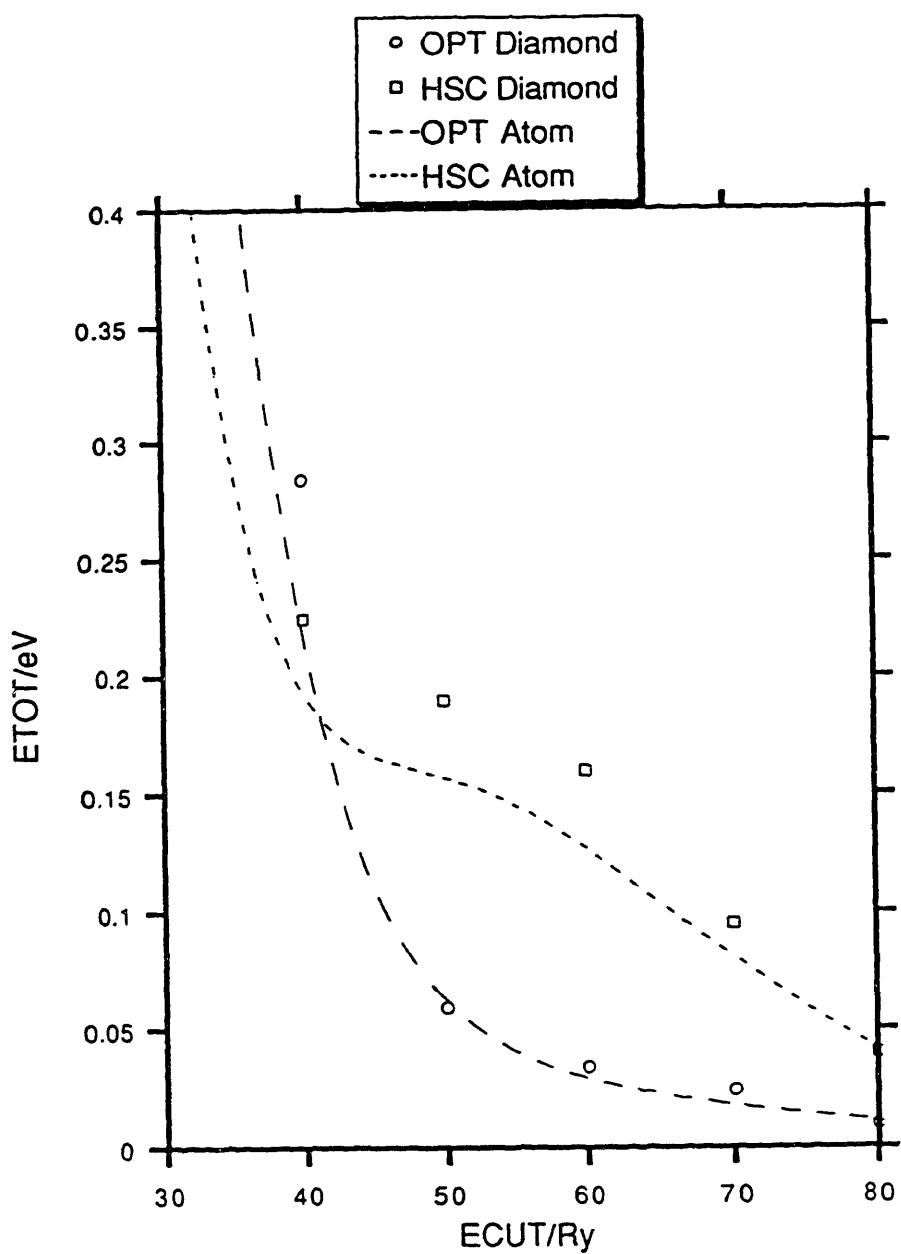
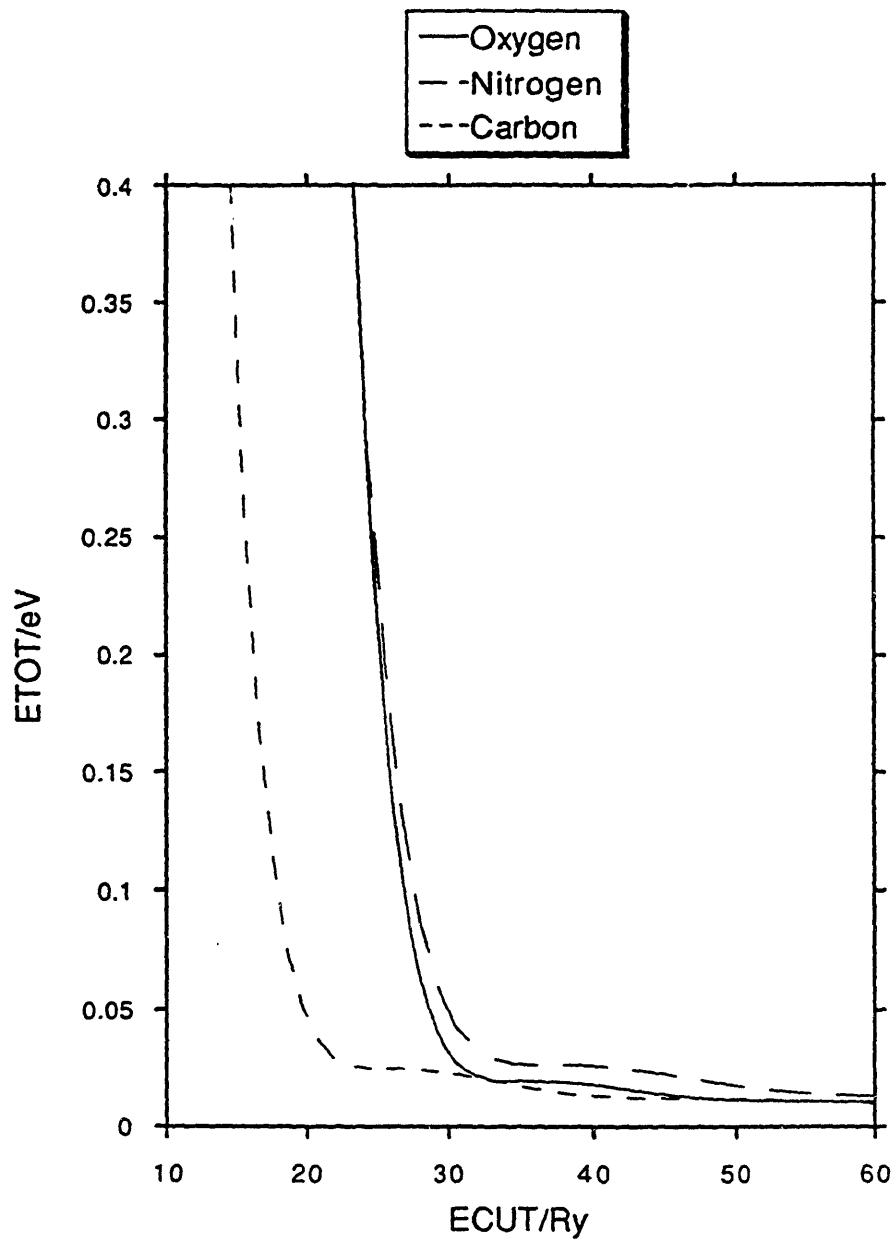


FIG. 3. Convergence of the total energy of carbon as a free pseudoatom and in diamond using HSC and optimized norm-conserving pseudopotentials.

generalized eigenvalue approach is very promising because of its ability to reduce the plane-wave requirements of calculations in return for an increase in the number of nonlocal potentials. This is accomplished by the removal of the norm conservation constraint on the wavefunctions through the use of a generalized eigenvalue formulation designed to correct the charge deficiencies. We can use the same criterion of convergence of atomic kinetic energy to design pseudopotentials with the best convergence in this formalism as well. In fact, a simple computational implementation involves using our optimization procedure without the norm conservation condition. The convergence properties of pseudopotentials for carbon, nitrogen, and oxygen are shown in Figure 4. The cutoff radii for these elements are  $1.4 \text{ } a_0$ ,  $1.25 \text{ } a_0$ , and  $1.35 \text{ } a_0$  respectively. A comparison of Figure 4 with Figure 3 shows the enormous gains in convergence which are possible when norm conservation is removed as a constraint.

This paper has demonstrated techniques for creating pseudopotentials which possess excellent convergence and transferability properties. We stress the importance of the convergence of the kinetic energy of the pseudoatom as an atomic convergence criterion for the convergence of solids containing the pseudoatom. Regardless of the procedure used to insure good transferability, utilization of the atomic convergence criterion consistently leads to excellent convergence. The results of atomic and solid-state calculations using these new pseudopotentials are exhibited, showing that excellent convergence and transferability are consistently obtained when the atomic convergence criterion is linked with an appropriate transferability criterion. These pseudopotentials facilitate the study of many elements in the periodic table, including transition metals and first-row nonmetals.

The authors should like to acknowledge the many contributions of Karin Rabe and Efthimios Kaxiras in the early stages of this work. The authors also wish to thank Su-Huai Wei for providing LAPW data for copper. This research was supported in part by the Office of Naval Research Contract No. N000 14-86-K-0158, the U. S. Joint Services



**FIG. 4.** Convergence of the total energies of pseudoatoms of the first-row nonmetals in non-norm-conserving formalism.

Electronics Program Contract No. DAAL-03-86-K-0002, and the U. S. Air Force Office of Scientific Research Grant No. AFOSR-90-0276. One of us (A. M. R.) acknowledges the support of the Joint Services Electronics Program.

**Chapter 3: A Test of the Utility of Plane Waves for the Study of  
Molecules from First Principles**

in collaboration with

**P. A. Bash**

Density functional theory<sup>1,2</sup> has recently shown significant promise<sup>3,4</sup> as an analytical tool for studying chemical systems from first principles. Currently, the calculations of properties of molecules containing first-row elements are most commonly performed using a basis set of localized orbitals for wavefunction expansions. The purpose of this paper is to report results which indicate the viability of the plane-wave basis set as an alternative choice for this expansion. First we will describe the salient features of this computational method, and then we will exhibit structural results for a number of small molecules.

The plane-wave basis set has a number of extremely desirable features. The set of plane waves is complete and orthonormal. In addition, plane waves can be indexed by a single vector index, the wave vector. These two properties make finite sets of plane waves systematically improvable in a straightforward manner. A calculation is performed using a basis set of all plane waves such that each plane wave's wave vector has a magnitude less than some value, the plane wave cutoff. Then the basis set is improved by increasing the plane-wave cutoff until the results converge. This systematic improvability provides another advantage, the numerical stability of results. Another benefit of using plane waves is that *a priori* knowledge of the electronic distribution is not required for the generation of the basis set. Finally, there are extremely reliable and efficient methods for finding electronic eigenstates and atomic positions<sup>5,6</sup>. The Car and Parrinello molecular dynamics method<sup>7,8</sup>, because it allows simultaneous updates to wavefunctions and atoms, speeds up calculations significantly relative to traditional matrix diagonalization approaches. The fast

---

<sup>1</sup>P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964). *J*

<sup>2</sup>W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965). *J*

<sup>3</sup>For a current survey see: *Density Functional Methods in Chemistry*, (Springer-Verlag, New York, 1991). *J*

<sup>4</sup>R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, (Oxford, New York, 1989). *B*

<sup>5</sup>J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C: Solid State Phys. **12**, 4409 (1979). *J*

<sup>6</sup>For reviews, see J. D. Joannopoulos in *Physics of Disordered Materials*, (Plenum Publishing, New York, 1985), and W. Pickett in *Computer Physics Reports* **9**, 115 (1989). *B*

<sup>7</sup>R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985). *J*

<sup>8</sup>M. C. Payne, J. D. Joannopoulos, D. C. Allan, M. P. Teter, and D. H. Vanderbilt, Phys. Rev. Lett. **56**, 2656 (1986). *J*

Fourier-transform algorithm improves the scaling of many parts of the calculation, and the conjugate gradient technique<sup>9,10</sup> for minimizing the electronic energy makes possible the study of large molecules, whereas other minimization techniques exhibit instability as the cell size becomes very large. The scaling behavior of the various parts of the calculation is shown in Table I. A gauge of the computational time for this method is the time required to calculate the total energy and quantum-mechanical forces on the ions in a particular geometry. For the formamide molecule in a 7 Å box with a plane-wave cutoff of 544.232 eV (40 Rydbergs), this calculation requires approximately 450 CPU seconds on a Cray-2.

Step	Scaling
Fourier transform of wavefunctions	$N_b N_{pw} \ln(N_{pw})$
acceleration of coefficients	$N_b N_{pw} \ln(N_{pw})$
forces on ions	$N_b N_i N_{pw}$
orthogonalization	$N_b N_b N_{pw}$

TABLE I. Scaling properties of the LDAP-PW method. Here  $N_b$  is the number of filled bands (orbitals),  $N_i$  is the number of ions, and  $N_{pw}$  is the number of plane waves.  $N_{pw}$  is roughly 100 times larger than  $N_b$ , which is twice the size of  $N_i$ .

The chief difficulty with using plane waves as a basis set is that sharply-peaked functions require huge numbers of planes waves to converge. For this reason, expression of core electrons using plane waves is extremely difficult. Because valence electrons are orthogonalized to the core, even the expression of valence electrons is very difficult when core electrons are included in the calculation. This problem has been surmounted by the use of pseudopotentials<sup>11,12</sup>. A pseudopotential is a weaker potential than the Coulomb potential, and it includes the effects of the core electrons implicitly. To generate a pseudopotential, a density functional calculation is normally performed on a free atom with

<sup>9</sup>M. P. Teter, M. C. Payne, and D. C. Allan, Phys. Rev. B **40**, 12255 (1989).

<sup>10</sup>T. A. Arias, M. C. Payne, and J. D. Joannopoulos, preprint. *J*

<sup>11</sup>J. C. Phillips, Phys. Rev. **112**, 685 (1958). *J*

<sup>12</sup>V. Heine, and M. L. Cohen, Solid State Physics **24**, (1970). *J*

specified electron fillings of the orbitals. A pseudopotential is then constructed so that outside the core region, its lowest eigenfunctions, called pseudowavefunctions, are identical to the valence wavefunctions of the atomic calculation<sup>13</sup>, and corresponding states have identical eigenvalues. The pseudopotential is also designed to reproduce the scattering properties of the atomic potential as accurately as possible<sup>14</sup>. Once it is designed with reference to an atomic density functional calculation, the pseudopotential is used without modification in any chemical environment. The pseudopotential approximation is motivated by the observation that valence electrons account for all chemical properties of atoms to a very large extent. Some core relaxation is neglected by this approach, but generalizations of the pseudopotential approximation can include this effect, and for many chemical systems of interest, the core relaxation effect is indeed negligible. Thus, the pseudopotential approximation greatly reduces the plane-wave requirements, making density functional calculations with plane waves tractable and leaving the vital chemical interactions intact.

Although pseudopotentials provide a huge gain in computational efficiency, the elements in the first row of the periodic table are still very difficult to describe with plane waves, and for this reason very few plane-wave pseudopotential calculations have been performed on materials containing these elements. Recently, significant advances have been made toward making pseudopotentials for these elements which give converged results with a lower plane-wave cutoff<sup>15,16</sup>. We use pseudopotentials which are designed to give the most convergent results possible and still provide accuracy<sup>17,18</sup> using the efficient nonlocal pseudopotential form of Kleinman and Bylander<sup>19</sup>. These

---

<sup>13</sup>D. R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979). *J*

<sup>14</sup>D. C. Allan and M. P. Teter, J. Am. Ceram. Soc. **73**, 3247 (1990). *J*

<sup>15</sup>R. Chelikowsky and M. Y. Chou, Phys. Rev. B **38**, 7966 (1988). *J*

<sup>16</sup>Y. Bar-Yam, S. T. Pantelides, and J. D. Joannopoulos, Phys. Rev. B **39**, 3396 (1989). *J*

<sup>17</sup>A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990). *J*

<sup>18</sup>A. M. Rappe and J. D. Joannopoulos, *Computer Simulation in Materials Science* (Kluwer, the Netherlands, 1991), pp. 409-422. *J*

<sup>19</sup>L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982). *J*

pseudopotentials increase computational efficiency sufficiently to allow calculation of the properties of fairly large molecules. Recent generalizations of the pseudopotential form<sup>20</sup> hold great promise that even more efficient pseudopotentials will be available in the near future.

Another difficulty with plane waves is that they enforce a periodicity not present in most molecules. A plane-wave basis set is directly applicable to systems which exhibit discrete translational symmetry, so the plane-wave pseudopotential method has been used extensively for crystalline materials. The supercell method greatly expands the class of materials which can be studied with plane waves. In this method, an aperiodic system is placed at each site of a three-dimensional lattice of points, thereby restoring periodicity. In the limit that the lattice points are extremely far apart, the aperiodic system is modeled accurately. Computationally, one systematically increases the lattice spacing and filters out the effects of the induced periodicity<sup>21</sup>.

Because the plane-wave pseudopotential method has a number of positive features as described above, we wish to examine whether it is a viable alternative method for performing high-level *ab initio* quantum chemistry calculations<sup>22,23,24,25,26</sup>. This involves testing the ability of pseudopotentials to provide accurate results for first-row atoms, the ability of plane waves to describe the valence wavefunctions efficiently, and the ability of the supercell method to mimic an isolated system. All of these computational aspects are completely independent of the exchange-correlation functional, and so for simplicity, we use a local density approximation<sup>27,28</sup> to this functional. For more realistic

---

<sup>20</sup>For examples, see D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990); K. Laasonen, R. Car, C. Lee, and D. Vanderbilt, Phys. Rev. B **43**, 6796 (1991);<sup>J</sup> R. D. King-Smith, M. C. Payne, and J. S. Lin, preprint (1991). <sup>pre-print</sup>

<sup>21</sup>G. A. Baraff and M. A. Schluter, Phys. Rev. B **19**, 4965 (1979). <sup>J</sup>

<sup>22</sup>D. Hohl, R. O. Jones, R. Car, and M. Parrinello, Chem. Phys. Lett. **139**, 540 (1987). <sup>J</sup>

<sup>23</sup>D. Hohl, R. O. Jones, R. Car, and M. Parrinello, J. Chem. Phys. **89**, 6823 (1988). <sup>J</sup>

<sup>24</sup>D. Hohl, R. O. Jones, R. Car, and M. Parrinello, J. Am. Chem. Soc. **111**, 825 (1989). <sup>J</sup>

<sup>25</sup>R. O. Jones, and D. Hohl, J. Am. Chem. Soc. **112**, 2590 (1990). <sup>J</sup>

<sup>26</sup>D. C. Allan and M. P. Teter, private communication (1990). <sup>J</sup>

<sup>27</sup>J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981). <sup>J</sup>

<sup>28</sup>D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980). <sup>J</sup>

calculations, we must use a more accurate functional, but the performance of the LDAP-PW (local density approximation, pseudopotential, plane-wave) method will permit us to gauge the effectiveness of pseudopotentials, plane waves, and supercells.

We generated pseudopotentials according to the methods described above, and tested their convergence properties. All pseudopotential calculations of structural parameters were very well converged using a plane-wave cutoff of 544.232 eV. For example, increasing the cutoff to 816.348 eV changed the  $\langle\text{OCN}\rangle$  of formamide by less than 0.1 degrees, and the  $r(\text{CO})$  and  $r(\text{CN})$  distances changed by 0.001 Å and 0.013 Å respectively. The accuracy of the pseudopotentials was also tested by calculating their scattering properties and by comparing energy differences between different atomic electronic filling states with the same energy differences evaluated by all-electron density functional calculations. An exception is to be noted in the case of hydrogen. Because it has no core electrons, we used the Coulomb potential to represent the hydrogen nucleus.

We tested the supercell method in two ways. We calculated structural properties of a particular molecule with a certain lattice spacing, and then repeated the calculation with a larger spacing. We found that as long as the lattice is large enough so that the van der Waals radii of the molecule do not overlap with those of the neighboring molecules, the results are completely insensitive to lattice spacing. For example, increasing the box size of formamide from a 7 Å cube to a 9 Å cube changed the  $\langle\text{OCN}\rangle$  by 0.1 degrees, the  $r(\text{CO})$  distance by 0.001 Å, and the  $r(\text{CN})$  distance by 0.004 Å. In addition, we studied the effect of changing the orientation of the molecules with respect to the translation axes. Here again, as long as the molecules are well-separated, the change in structural parameters when the molecules are reoriented is minimal. Again taking the formamide example, we find that a 30 degree rotation of the molecule changed the bond angle by less than 0.1 degrees and the bond distances by less than 0.001 Å. One difficulty arises when molecules are placed on a lattice. Even though no chemical bonding takes place between neighboring molecules, and the bonds and angles are unaffected by periodicity, the long-range dipole-

dipole interactions slow down the calculations. All the molecules on the lattice rotate to make their dipole moments align. These large distance (sometimes several angstroms) motions to satisfy spurious intermolecular forces are much more time-consuming computationally than the desired adjustments to attain the proper intramolecular geometry. We have tested that the calculated properties of the molecule are the same in any orientation. Therefore, we inhibit rotation of the molecule as a whole by the following procedure. We compute the quantum-mechanical forces on all the nuclei. We keep one atom fixed at a point. We constrain a second atom to move only on a particular line which passes through the first atom, and force a third atom to move only within a plane which passes through the first two atoms. All other atoms are free to follow the quantum-mechanical forces without restriction. This allows all interatomic distances and angles to optimize freely, but it prevents rigid rotation of the molecule.

To determine whether the LDAP-PW method and generalizations of it hold promise for quantum chemistry, we calculated structural properties of a number of small molecules. The calculations were done at a 544.232 eV plane-wave cutoff (with the exception of H<sub>2</sub> which was done at 816.35 eV) and supercell sizes of 5.5Å × 5.5Å × 5.5Å (4800 plane waves), 7Å × 7Å × 7Å (9900 plane waves), or 11Å × 7Å × 5Å (11,100 plane waves) depending on the molecule studied. The electronic degrees of freedom were relaxed until the total energy of the molecule was within 10<sup>-4</sup> eV/molecule of the converged value. Molecular geometries were considered to be converged when forces on the atoms were less than 0.01 eV/Å; optimizations beyond this level changed structures by less than 1%. Initial starting configurations for H<sub>2</sub>O, CH<sub>4</sub>, and NH<sub>3</sub>, were such that the hydrogen atoms were placed at right angles to one another and a distance of about 1.0Å from their associated heavy atom. The initial coordinates for all other molecules were taken from AM1 semi-empirical calculations using the MOPAC<sup>29</sup> program.

---

<sup>29</sup>J. J. P. Stewart QCPE 455 6, 1986. *J*

Molecule	Parameter	Exp.	MP2	$\Delta^1$	LDAP-PW	$\Delta^2$	$\Delta_G$
H <sub>2</sub>	r(HH)	0.742	0.738	0.004	0.776	0.034	0.038
O <sub>2</sub>	r(OO)	1.208	1.242	0.034	1.183	0.025	0.059
N <sub>2</sub>	r(NN)	1.098	1.131	0.033	1.099	0.001	0.032
P <sub>2</sub>	r(PP)	1.893	1.936	0.043	1.875	0.018	0.061
CO	r(CO)	1.128	1.151	0.023	1.138	0.010	0.013
CO <sub>2</sub>	r(CO)	1.162	-	-	1.170	0.008	-
H <sub>2</sub> O	r(OH)	0.958	0.969	0.011	0.976	0.018	0.007
	<(HOH)	104.5	104.0	0.5	104.0	0.5	0.0
CH <sub>4</sub>	r(CH)	1.092	1.090	0.002	1.102	0.010	0.012
NH <sub>3</sub>	r(NH)	1.012	1.017	0.005	1.028	0.016	0.011
	<(HNH)	106.7	106.3	0.4	105.7	1.0	0.6
PH <sub>3</sub>	r(PH)	1.420	1.415	0.005	1.440	0.020	0.025
	<(HPH)	93.3	94.6	1.3	91.6	1.7	3.0
H <sub>2</sub> O <sub>2</sub>	r(OO)	1.452	1.467	0.015	1.414	0.038	0.053
	r(OH)	0.965	0.976	0.011	0.982	0.017	0.006
	<(OOH)	100.0	98.7	1.3	100.2	0.2	1.5
	$\omega$ (HOOH)	119.1	121.3	2.2	118.7	0.4	2.6
N <sub>2</sub> H <sub>2</sub>	r(NN)	1.252	1.267	0.015	1.233	0.019	0.034
	r(NH)	1.028	1.036	0.008	1.052	0.024	0.016
	<(NNH)	106.9	105.4	1.5	107.2	0.3	1.8
H <sub>2</sub> CO	r(CO)	1.208	1.221	0.013	1.205	0.003	0.016
	r(CH)	1.116	1.104	0.012	1.135	0.019	0.031
	<(HCH)	116.5	115.6	0.9	116.4	0.1	0.8
C <sub>2</sub> H <sub>2</sub>	r(CC)	1.203	1.218	0.015	1.196	0.007	0.022
	r(CH)	1.061	1.066	0.005	1.058	0.003	0.008
HNC	r(NC)	1.169	1.187	0.018	1.178	0.009	0.009
	r(HN)	0.994	1.002	0.008	1.018	0.024	0.016
HNO	r(NO)	1.212	1.237	0.025	1.188	0.024	0.049
	r(HN)	1.063	1.058	0.005	1.093	0.030	0.035
	<(HNO)	108.6	107.3	1.3	109.2	0.6	1.9

TABLE II. See caption next page.

Molecule	Parameter	Exp	MP2	$\Delta_e^1$	LDAP-PW	$\Delta_e^2$	$\Delta_G$
Formamide (HCONH <sub>2</sub> )	r(CO)	1.193	1.224	0.031	1.217	0.024	0.007
	r(CN)	1.376	1.361	0.015	1.365	0.011	0.004
	<(OCN)	124.7	124.8	0.1	124.9	0.2	0.1
Glycine (NH <sub>2</sub> CH <sub>2</sub> CO <sup>1</sup> O <sup>2</sup> H)	r(C=O <sup>1</sup> )	-	1.218	-	1.204	-	0.014
	r(CO <sup>2</sup> )	-	1.358	-	1.345	-	0.013
	r(CC)	-	1.515	-	1.535	-	0.020
	r(NC)	-	1.451	-	1.460	-	0.009
	<(O <sup>1</sup> CO <sup>2</sup> )	-	123.2	-	123.7	-	0.5
	<(O <sup>2</sup> CC)	-	111.4	-	111.9	-	0.5
	<(O <sup>1</sup> CC)	-	125.4	-	124.4	-	1.0
	<(CCN)	-	114.9	-	114.8	-	0.1
	r(PO)	-	1.510	-	1.482	-	0.028
$\text{H}_2\text{PO}_4^{-1}$	r(P-OH)	-	1.680	-	1.640	-	0.040
	<(HO-P-OH)	-	101.1	-	101.2	-	0.1
	<(O-P-OH)	-	106.9	-	107.1	-	0.2
	<(O-P-O)	-	126.2	-	125.2	-	1.0
3-hydroxy-acrylaldehyde (HOCHCHCHO)	r(O-H)	0.969	0.994	0.025	1.079	0.110	0.085
	r(O...H)	1.680	1.694	0.014	1.418	0.262	0.276

TABLE II. A comparison of LDAP-PW geometry optimizations with experiment and with MP2. All experimental and MP2 values are taken from reference 32 except those of glycine,  $\text{H}_2\text{PO}_4^{-1}$ , and 3-hydroxy-acrylaldehyde<sup>30,31</sup> which were optimized using Gaussian 90 at MP2 / 6-31G\*, MP2 / 6-31+G\*, and MP2 / 6-31G\*\* respectively.  $\Delta_e^1$  and  $\Delta_e^2$  are the deviations of the MP2 and LDAP-PW calculations from experiment, and  $\Delta_G$  is the deviation of LDAP-PW from the MP2 values.

<sup>30</sup>J. S. Binkley, M. J. Frisch, and H. F. Schaefer III, Chem. Phys. Lett. 126, 1 (1986). *J*

<sup>31</sup>S. L. Baughcum, R. W. Duerst, W. F. Rowe, Z. Smith, and E. B. Wilson, J. Am Chem. Soc. 103, 6296 (1981). *J*

Table II lists the results of the calculations and compares them with both experimental values and MP2 / 6-31G\*, MP2 / 6-31+G\* or MP2 / 6-31G\*\* Hartree-Fock<sup>32</sup> calculations. Overall agreement of LDAP-PW with experiment and MP2 calculations is excellent. Average deviations from experiment and MP2 are 0.030Å and 0.034Å for bond lengths and 0.5 degrees and 1.0 degrees for bond angles. These results also compare favorably with localized orbital density functional methods. Reference 3 lists several practical examples. If the hydrogen-bonded system 3-hydroxy-acrylaldehyde is not included in the averages, the bond-length deviations become 0.017Å and 0.024Å respectively. Upon closer examination, certain trends become apparent within the data contained in Table II. The most striking pattern is that an LDAP-PW calculation of a bond distance between two heavy (non-hydrogen) atoms is almost always closer to experiment than the corresponding MP2 result. Conversely, bonds involving at least one hydrogen are almost always calculated more accurately using MP2. In addition, bond angles which include zero or one hydrogen atoms are almost always more accurate using LDAP-PW, while those including two hydrogens are usually better reproduced by MP2. We believe that these effects are due to the use of the LDA rather than a more accurate exchange-correlation functional. The small size of hydrogen makes its charge density change over a short length scale. Another trend which we have observed is that long bonds are reproduced more poorly than short bonds by the LDAP-PW. As the atoms in a bond separate, they act more like free atoms, and the LDA does a poorer job on free atoms than on covalent materials<sup>33</sup>. This effect is apparent in the O-O bond of H<sub>2</sub>O<sub>2</sub> and the hydrogen bond in 3-hydroxy-acrylaldehyde. The most striking deviation from MP2 and from experimental results is the internal hydrogen bond in 3-hydroxy-acrylaldehyde. The LCGTO-MCP-LSD method gave a O-H distance of 1.228Å for this molecule (referred to as malonaldehyde in reference 3). Our result of 1.418Å is closer to experiment, but it is

---

<sup>32</sup>W. J. Hehre, L. Radom, P. v.R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986). *B*

<sup>33</sup>S. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. Lett. **61**, 1631 (1988). *J*

still 0.262Å from the experimental value of 1.68Å. The use of the plane-wave basis set provides accurate solution of the density functional equations, and the inadequacy in the local density approximation appears to be the main barrier to more highly accurate calculations. Work is in progress to incorporate non-local corrections<sup>34,35</sup>.

These results show that the LDAP-PW method can produce accurate structural parameters for molecules. The data suggest that the method which uses plane waves, pseudopotentials, and supercells may be a useful alternative to standard *ab initio* quantum mechanical methods for calculations on chemical systems. Future work will include more realistic density functionals and tests of energetics of chemical transformations. In addition, research is continuing to develop more efficient energy minimization algorithms and superior pseudopotentials. These changes will improve the efficiency and accuracy of the method for the study of periodic and aperiodic chemical and biochemical systems.

This research was supported in part by the Florida State University Supercomputer Computations Research Institute which is partially funded by the U. S. Department of Energy through Contract No. DE-FC05-85ER250000. We would also like to acknowledge computer time provided by the Florida State University Computing Center, the M. I. T. Supercomputer Facility, and the San Diego Supercomputer Center. One of us (A. M. R.) is supported by a Joint Services Electronics Program Fellowship.

---

<sup>34</sup>A. D. Becke, Phys. Rev. A **38**, 3098 (1988) *J*

<sup>35</sup>J. P. Perdew, Phys. Rev. B **33**, 8822, (1986). *J*

**Chapter 4: *Ab Initio* Study of a Grain Boundary in Gold  
in collaboration with  
M. Needels and P. D. Bristowe**

It is now well established that *ab initio* molecular dynamics can be used effectively to study the bulk and defect properties of sp-bonded materials<sup>1,2,3,4,5,6</sup>. The principal reason for this success lies in the ability to describe these materials accurately using "soft" pseudopotentials and a limited plane-wave basis set. Clearly, it would be desirable to extend this technique to other materials while still retaining the advantages of pseudopotentials and plane waves. Transition metals are of interest because of their many metallurgical applications but, as is well known, they are difficult to describe using conventional pseudopotential theory. In particular, they require an extremely large number of plane waves for total-energy convergence because of the presence of d orbitals which are strongly localized but make an important contribution to the bonding<sup>7</sup>. Recently, however, a new approach to constructing *ab initio* pseudopotentials has been proposed which minimizes these problems. In this approach, the freedom in constructing the pseudopotential is explicitly used to optimize the total-energy convergence of the solid<sup>8</sup>. Specifically, by optimizing the atomic kinetic energy and requiring the potential to be continuous, a soft pseudopotential can be constructed which is more rapidly convergent than a conventional pseudopotential. This has been demonstrated for the case of copper in which the lattice constant, cohesive energy, and bulk modulus of that metal have been computed efficiently and accurately<sup>8</sup>.

In the present paper, we apply the optimized pseudopotential approach in an *ab initio* molecular dynamics calculation of a grain boundary in the transition metal, gold. Grain boundaries are important interfacial defects in metals and control many of the physical properties of the material. Our objective is twofold. First, we wish to show that

---

<sup>1</sup>M. C. Payne, P. D. Bristowe, and J. D. Joannopoulos, Phys. Rev. Lett. **58**, 1348 (1987). *J*

<sup>2</sup>M. Needels, M. C. Payne, and J. D. Joannopoulos, Phys. Rev. Lett. **58**, 1765 (1987). *J*

<sup>3</sup>D. C. Allan and M. P. Teter, Phys. Rev. Lett. **59**, 1136 (1987). *J*

<sup>4</sup>G. W. Fernando, G. -X. Qian, M. Weinert, and J. W. Davenport, Phys. Rev. B **40**, 7985 (1989). *J*

<sup>5</sup>C. Woodward, B. I. Min, R. Benedek, and J. Garner, Phys. Rev. B **39**, 4853 (1989). *J*

<sup>6</sup>R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985). *J*

<sup>7</sup>D. C. Allan and M. P. Teter, J. Am. Ceram. Soc. **73**, 3247 (1990). *J*

<sup>8</sup>A. M. Rappe, K. M. Rabe, E. Kaxiras and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990). *J*

large complex transition metal systems can be investigated using plane-wave basis sets and, second, that an energetically preferred model for a specific twist grain boundary can be determined and compared to previous experimental and theoretical studies. The results represent the first *ab initio* investigation of a grain boundary in a transition metal.

The computational methodology is as follows. We first tried using a preconditioned conjugate gradient algorithm<sup>9</sup>. However, this approach failed miserably since the total energy could not be converged to within ten eV because it began to oscillate randomly. Therefore, we used a steepest-descent equation-of-motion algorithm<sup>10</sup>. We used an "optimized" non-local, norm-conserving, scalar-relativistic pseudopotential of the Kleinman-Bylander form<sup>11</sup>. We constructed the pseudopotential by solving the Dirac equation for the all-electron problem to obtain spin-up and spin-down pseudopotentials for the s, p, and d angular momentum wavefunctions<sup>12</sup>. The potentials are obtained by pseudizing the valence major components<sup>13,14</sup> and minimizing the kinetic energy beyond a cutoff<sup>8,15</sup> of 30 Ry. The scalar-relativistic potentials for each angular momentum are obtained by averaging the spin-up and spin-down potentials by their respective multiplicities<sup>14</sup>. These "averaged" potentials had the same level of convergence as the spin-dependent pseudopotentials.

The pseudopotential was tested in a number of ways. First, total-energy differences between the atomic ground state configuration and several excited-state configurations were calculated for the all-electron atom using the Dirac equation and the pseudoatom using the Schrodinger equation. These energy differences agreed to within

---

<sup>9</sup>M. P. Teter, M. C. Payne and D. C. Allan, Phys. Rev. B**40**, 12255 (1989). *J*

<sup>10</sup>M. C. Payne, J. D. Joannopoulos, D. C. Allan, M. P. Teter and D. H. Vanderbilt, Phys. Rev. Lett. **56**, 2656 (1986). *J*

<sup>11</sup>L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982). *J*

<sup>12</sup>The core radii are 2.25  $a_0$  for all three angular momentum states; the reference configuration is  $\{\text{Xe}\}5\text{d}^10\text{s}^{0.5}\text{p}^{0.5}$ .

<sup>13</sup>D. R. Hamann, M. Schluter and C. Chiang, Phys. Rev. Lett. **43**, 1491 (1979). *J*

<sup>14</sup>G. B. Bachelet and M. Schluter, Phys. Rev. B **25**, 2103 (1982). *J*

<sup>15</sup>A. M. Rappe and J. D. Joannopoulos, *Computer Simulation in Materials Science* (Kluwer, the Netherlands, 1991), pp. 409-422. *J*

1% in all cases. Second, the bulk properties of gold were calculated and are displayed in Table I. The lattice constant and bulk modulus agree well with experimental and other theoretical results<sup>16</sup>. The cohesive energy exceeds the experimental value, as expected for any LDA calculation. However, our result agrees well with other theoretical values.

	Experiment	Present	% Difference
Lattice constant ( $\text{\AA}$ )	4.079	4.086	+ 0.2
Cohesive energy (eV/atom)	3.81	4.11	+ 7.9
Bulk modulus (MBar)	1.732	1.95	+ 12.6

TABLE I. Comparison of the bulk properties of gold as calculated in the present work to experimental values. (plane-wave cutoff = 30 Ry; no. of  $\mathbf{k}$  points = 110)

The other aspects of the calculation were also tested. The supercell contained two grain boundaries to maintain periodicity in the directions normal to the boundary plane. Each grain consisted of six (002) planes of atoms with five atoms per plane. The separation of the grain boundaries was tested, using the embedded atom method<sup>17</sup>, by comparing grain boundary energies between the 12-layer supercell and an 80-layer cell. The energy difference was small. This can be understood in terms of the small intrinsic dislocation spacing of the boundary we are considering. Within elasticity theory, the dislocation strain field decays on a length scale approximately equal to this spacing. Convergence with respect to basis set size was checked by increasing the cutoff from 30 Ry, which we used for the grain boundary calculation (18000 plane waves), to 50 Ry. The total energy of a one-atom unit cell of crystalline gold changed by less than 0.02 eV, indicating that the calculation was very well converged with respect to the basis set size. Brillouin zone averaging for the grain boundary was obtained by summing over three

<sup>16</sup>N. Takeuchi, C. T. Chan, and K. M. Ho, Phys. Rev. B **40**, 1565 (1989). *S*

<sup>17</sup>M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984). *S*

special  $\mathbf{k}$  points in the irreducible 1/16 of the Brillouin zone. The possibility of partially filled and/or degenerate states occurring near the Fermi level was treated using a Gaussian spreading scheme<sup>4</sup> with a width of 0.05 eV. *A priori*, we estimated the accuracy of this  $\mathbf{k}$  point set to be somewhere between that using 28 and 60  $\mathbf{k}$  points in the bulk irreducible Brillouin zone. To compute the energy of the grain boundaries, we used a reference total energy of bulk gold obtained by using an identical  $\mathbf{k}$  point sampling in a 10-atom unit cell of bulk gold, which gives 24  $\mathbf{k}$  points in the irreducible 1/8 of the Brillouin zone, and an energy of -890.212 eV/atom. As can be seen from Table II, with this level of  $\mathbf{k}$  point sampling, the absolute total energy is converged to a few meV/atom.

---

No. of $\mathbf{k}$ points in irreducible Brillouin Zone	Energy/Atom (eV)
19	-890.201
28	-890.192
44	-890.211
60	-890.215
85	-890.212
110	-890.215

---

TABLE II. Bulk  $\mathbf{k}$  point convergence test. (plane-wave cutoff = 30 Ry)

The grain boundary of interest is the  $\Sigma = 5$  ( $36.9^\circ$ ) [001] twist boundary in gold where  $\Sigma$  is the inverse density of coincidence sites.<sup>18</sup> The perfect geometrical configuration for this boundary is formed by rotating one fcc crystal with respect to another by  $36.9^\circ$  about [001]. This boundary has been studied extensively using X-ray diffraction techniques<sup>19,20</sup> and classical atomistic simulation methods<sup>20,21</sup>. The X-ray investigations have independently determined two distinct atomic models for the boundary which are

<sup>18</sup>W. Bollmann, *Crystal Defects and Crystalline Interfaces* (Springer, New York, 1970). ↗

<sup>19</sup>M. R. Fitzsimmons and S. L. Sass, *Acta Metall.* **36**, 3101 (1988). ↗

<sup>20</sup>I. Majid, P. D. Bristowe, and R. W. Balluffi, *Phys. Rev. B* **40**, 2779 (1989). ↗

<sup>21</sup>P. D. Bristowe and A. G. Crocker, *Philos. Mag. A* **38**, 487 (1978). ↗

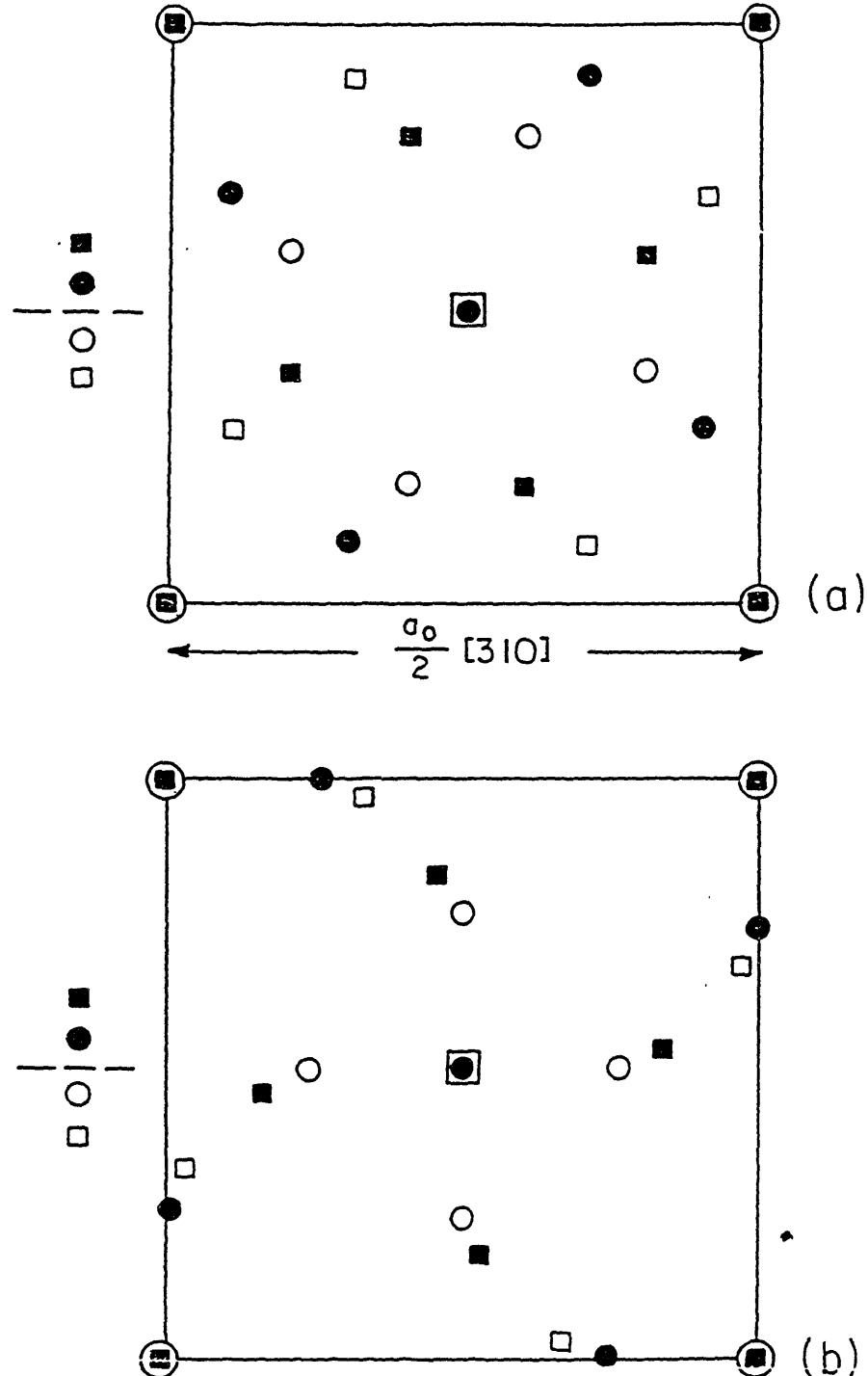


FIG. 1. Two models for the  $\Sigma = 5$  ( $36.9^\circ$ ) [001] twist boundary in gold:  
(a) MBB model, (b) FS model. Structures are shown in projection along  
[001] with two planes on either side of the boundary. Stacking sequence of  
planes is indicated by the vertical column of open and closed symbols.  
Note that the MBB model is indistinguishable from the perfect unrelaxed  
geometry on the scale of this figure.

shown in projection in Fig. 1. They are labeled the FS (Fitzsimmons-Sass) model and the MBB (Majid-Bristowe-Balluffi) model. Note that the MBB model is almost indistinguishable from the perfect geometrical configuration on the scale of this figure. Both models can be characterized by rotational-like relaxations around a symmetry axis passing through the center of the unit cell, but they differ in the magnitude of these relaxations. The maximum relaxation exhibited by the FS model (relative to the unrelaxed geometrical configuration) is about  $0.7\text{\AA}$  which is approximately seven times as large as the corresponding relaxation in the MBB model. It is not immediately obvious which model should have the lower energy. However, it is noted that the FS model has 80 bond lengths that are distorted by more than 5% whereas the MBB model has only 12.

The computed energies for the two twist boundary models are shown in Table III. It is clear that the MBB model is energetically preferred over the FS model by more than a factor of ten. The forces on the atoms were also computed and found to be two orders of magnitude higher in the FS model compared to the MBB model. For the MBB model, the atomic forces in the plane of the boundary were all less than  $0.1 \text{ eV}/\text{\AA}$  and those normal to the boundary were less than  $0.26 \text{ eV}/\text{\AA}$ . This level of residual force is considered quite satisfactory for this kind of calculation and therefore the structure is thought to be relatively stable. On the other hand, the forces acting on the atoms in the FS model are so large than the structure must be unstable. In Figure 2 we illustrate the direction and magnitude of these forces as vectors projected onto the model structure.

The energy of both model structures have been computed previously using semi-empirical methods<sup>20,22</sup>. For comparison with the present calculations we present in Table III the results of the embedded atom (EAM) calculations<sup>20</sup>. It is seen that these calculations also yield a much higher energy for the FS model (by a factor of about seven) compared to the MBB model. In addition, the magnitude of the MBB model energy is remarkably close

---

<sup>22</sup>J. Budai, P. D. Bristowe, and S. L. Sass, *Acta Metall.* **31**, 699 (1983).

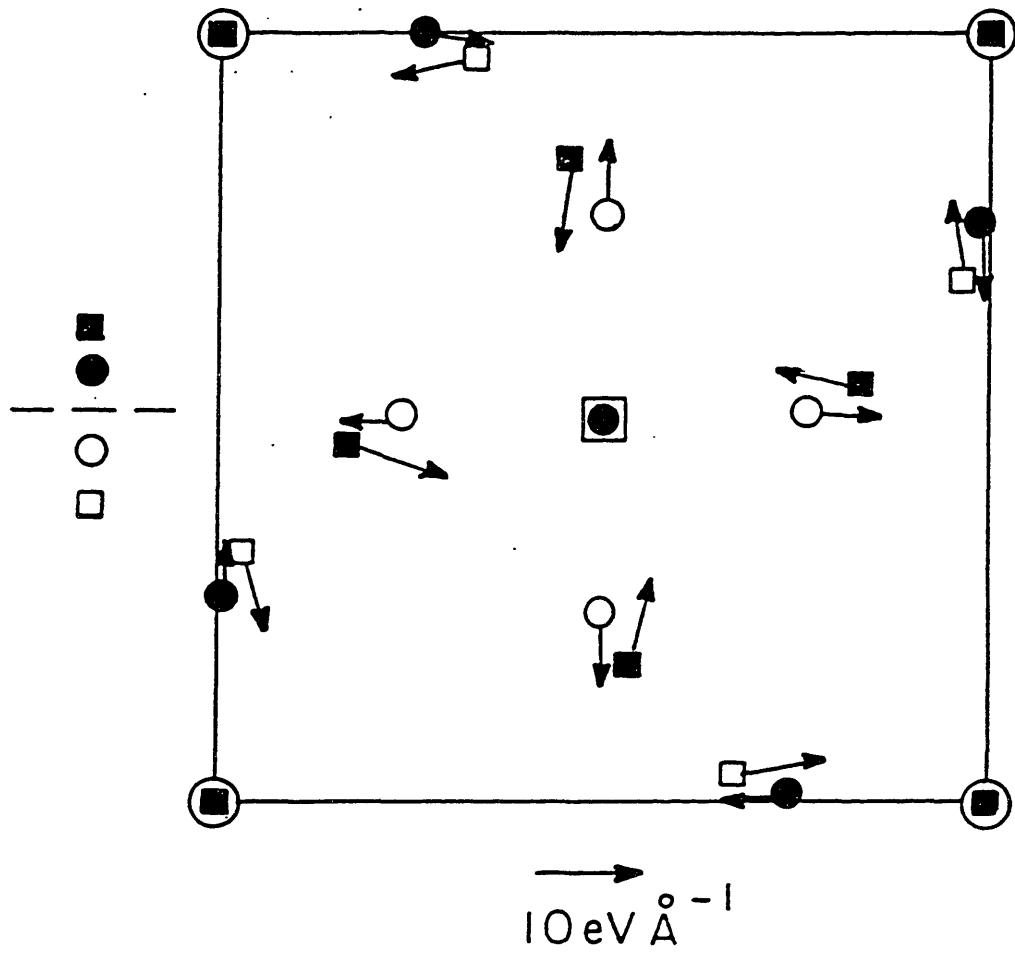


FIG. 2. Forces acting on atoms in FS model determined from the pseudopotential calculation. Only forces parallel to the boundary are shown and are of order  $10 \text{ eV}/\text{\AA}$ . Forces normal to the boundary are of order  $1 \text{ eV}/\text{\AA}$ .

		eV/Å <sup>2</sup>	ergs/cm <sup>2</sup>
Present Work	MBB Model	0.031	510
	FS Model	0.483	7,731
Embedded Atom Method	MBB Model	0.032	522
	FS Model	0.236	3,786

TABLE III. Computed twist boundary energies using *ab initio* method and EAM.

to that computed using the *ab initio* method. However, the FS model energy computed using EAM is smaller by a factor of two compared to the corresponding energy determined from the *ab initio* approach. This difference can be explained by the highly distorted nature of the FS model which results in large deviations in charge density. These deviations are accurately treated in the *ab initio* calculation but are poorly described using EAM. Nevertheless, the vector map showing the EAM forces on the atoms in the FS model is very similar (at least in terms of the directions of the vectors) to that shown in Figure 2. Since the FS model is known to relax back into the MBB model using EAM, it is reasonable to assume that such relaxations would also occur in the *ab initio* calculation if the atoms were free to move.

In Figure 3, we plot the planar charge density as a function of distance from the grain boundary. The charge density for half our unit cell is shown. The six atomic layers can be clearly identified as the six flat peaks. The grain boundary is at the middle of the plot and corresponds to the minimum in charge density. This minimum is broader than the other interlayer regions, as expected, because the interlayer distance is increased at the grain boundary to prevent the atoms in the two grains from being too close together. However, the minimum is also much deeper, being two thirds the average charge density in the unit

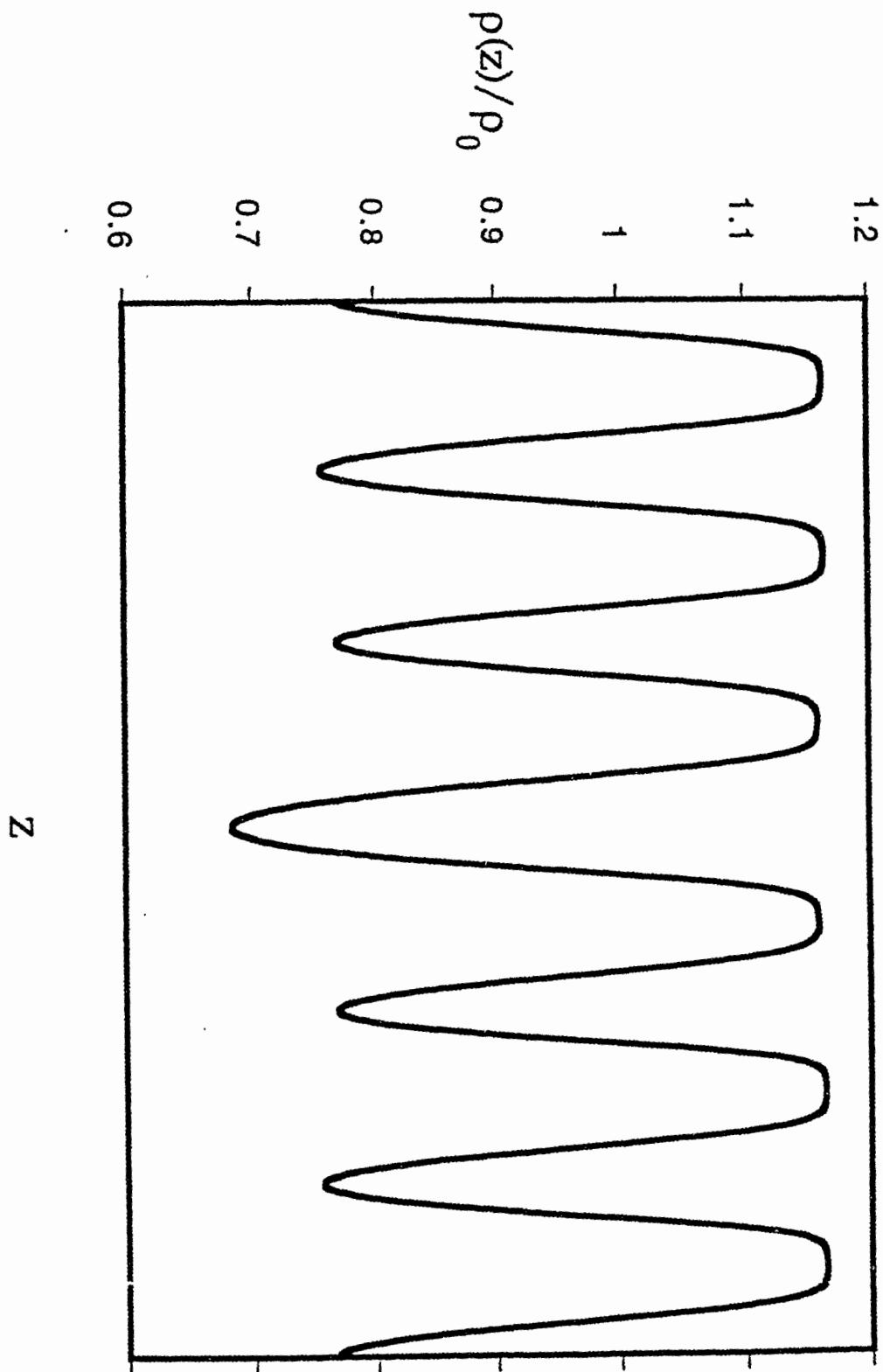


FIG. 3. Planar charge density, normalized to the average charge density in gold, as a function of distance from the grain boundary. The grain boundary is in the middle of the plot.

cell. Since there is a lack of fine structure in the charge density at the grain boundary, there are apparently no localized electronic states associated with it.

In conclusion, we have successfully performed an *ab initio* molecular dynamics calculation of the total energy of a grain boundary in the transition metal, gold. The calculation was made possible through the application of an optimized pseudopotential and a plane-wave basis set. In addition, we have confirmed the low energy and stability of the MBB model for the  $\Sigma = 5$  twist boundary in agreement with one of the previous studies.<sup>20</sup> We envision further applications of this technique to the study of other defects in transition metal systems.

This work was supported in part by the U. S. Air Force Office of Scientific Research Contract No. 87-0098 and by the Department of Energy Grant No. DE-FG02-87ER-45310. A. M. R. acknowledges the support of a Joint Services Electronics Program Fellowship. We wish to thank Dr. Sverre Froyen for the use of his relativistic atomic programs. Supercomputer time was provided by the National Science Foundation at the San Diego Supercomputer Center. Additional support was provided by the M. I. T. Supercomputer Facility.

**Chapter 5: Photonic Bound States in Periodic Dielectric Materials**  
**in collaboration with**  
**R. D. Meade and K. D. Brommer**

Recently, it has been suggested that a dielectric material with a three-dimensional periodicity may have a "photonic gap" in its frequency spectrum in which *propagating* electromagnetic modes are forbidden.<sup>1</sup> It has also been proposed that photonic states in this forbidden region should be localized in all three dimensions.<sup>2</sup> Calculations by Ho, Chan, and Soukoulis have shown that a diamond lattice of dielectric spheres can indeed have such a gap.<sup>3</sup> In this paper, we show that lattice imperfections can introduce exponentially localized states in the photonic band gap. We focus on the frequency spectrum of dielectric structures containing defects in an otherwise perfect diamond lattice. In our calculations, we consider impurities of two types, air spheres in the dielectric region and dielectric spheres in the air region. In both cases, localized photon modes are found to be introduced into the gap. Variation of the size of the impurity sphere leads to complete tunability of the frequency of this localized mode.

Although localized modes are common in metallic cavities, these photonic bound states are unique and physically interesting because they are localized in an entirely dielectric medium. Moreover, they have discrete frequency spectra and may be constructed from low-loss dielectric materials. Therefore, it may be possible to reduce the dissipation inherent in metallic cavities, due to skin heating of the metallic walls, and produce discrete modes with long lifetimes and correspondingly sharp frequency resonances. We propose that this effect can be exploited to produce high-quality resonant cavities and filters in the microwave and millimeter-wave regime.

In order to calculate the electromagnetic frequency spectrum of dielectric lattices, a number of authors<sup>3,4,5</sup> have suggested expanding the electromagnetic fields in a plane-wave basis,

---

<sup>1</sup>E. Yablonovitch and T. J. Gmitter, Phys. Rev. Lett. **63**, 1950 (1989). *J*

<sup>2</sup>E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987). *J*

<sup>3</sup>K. M. Ho, C. T. Chan, C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990). *J*

<sup>4</sup>K. M. Leung and Y. F. Liu, Phys. Rev. Lett. **65**, 2646 (1990). *J*

<sup>5</sup>Ze Zhang and Sashi Satpathy, Phys. Rev. Lett. **65**, 2650 (1990). *J*

$$\mathbf{H}(\mathbf{r}) = \sum_{\mathbf{G}} \sum_{\lambda=1,2} h_{\mathbf{G},\lambda} \hat{\mathbf{e}}_{\lambda} e^{i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}}, \quad (1)$$

where  $\mathbf{k}$  is in the Brillouin zone,  $\mathbf{G}$  is summed over the reciprocal lattice, and  $\hat{\mathbf{e}}_{\lambda}$  are orthogonal to  $(\mathbf{k}+\mathbf{G})$ . Maxwell's equations are then expressed as a simple eigenvalue equation,<sup>3</sup>

$$\sum_{\mathbf{G}',\lambda'} H_{\mathbf{G},\mathbf{G}'}^{\lambda,\lambda'} h_{\mathbf{G}',\lambda'} = \omega^2 h_{\mathbf{G},\lambda}, \quad (2)$$

where  $c=1$  and

$$H_{\mathbf{G},\mathbf{G}'} = |\mathbf{k} + \mathbf{G}| |\mathbf{k} + \mathbf{G}'| \epsilon_{\mathbf{G},\mathbf{G}'}^{-1} \begin{pmatrix} \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_1' \\ \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_1' \end{pmatrix} \quad (3)$$

and  $\epsilon_{\mathbf{G},\mathbf{G}'}^{-1}$  is the Fourier transform of the inverse of the dielectric function  $\epsilon(\mathbf{r})$ . This eigenvalue equation is then solved to yield the normal mode coefficients and frequencies of the electromagnetic modes. Having solved for  $\mathbf{H}(\mathbf{r})$  the other electromagnetic fields can be determined simply. This technique provides a simple and powerful method to solve problems in electrodynamics which takes full account of the vector nature of the electromagnetic radiation. The approximations, the finite size of the plane-wave basis and Fourier-transform grid, can both be improved systematically.

Of the three similar techniques discussed in Refs. 3-5, we chose to employ the methods of Ho, Chan, and Soukoulis.<sup>3</sup> First, they Fourier expand  $\mathbf{H}(\mathbf{r})$  rather than  $\mathbf{E}(\mathbf{r})$  or  $\mathbf{D}(\mathbf{r})$ . Since  $\mathbf{E}(\mathbf{r})$  and  $\mathbf{D}(\mathbf{r})$  are discontinuous at interfaces between media with differing dielectric constants, but  $\mathbf{H}(\mathbf{r})$  is continuous (with discontinuities in its first derivatives), we expect  $\mathbf{H}(\mathbf{r})$  to have faster plane-wave convergence. Second, Ho, Chan, and Soukoulis<sup>3</sup> expand  $\mathbf{H}(\mathbf{r})$  in a basis of transverse plane waves, rather than expanding in a complete basis, solving for the normal modes, and then projecting out the longitudinal solutions.

Although we would like to study the effect of a single defect in an otherwise perfect lattice, computationally it is desirable to choose a system with a finite unit-cell size. Therefore, we employ the supercell method in which we place one defect in a repeated cell of atoms. In our calculations, we used the 8-atom conventional cell of diamond. We have performed tests using larger supercells containing 16 and 32 atoms and found similar results. We also considered supercells of 2, 4, 8, and 16 atoms containing dense impurities in the (111) planes, in order to determine the dependence of bandwidth on impurity separation. We expand in plane waves up to a finite frequency, including  $\sim$ 130 plane waves per polarization, per primitive unit cell. Based on our calculations of a one-dimensional system of periodic dielectric slabs, for which an exact solution is available,<sup>6</sup> we expect that the frequencies are correct to within  $\sim$ 5% at this plane-wave cutoff. Although fully converged calculations in these one-dimensional systems agreed to 0.1% we found that convergence in absolute frequencies was relatively slow in comparison to conventional electronic-structure calculations. We attribute this to the discontinuity in the first derivative of  $H(r)$  at the dielectric surfaces. In order to calculate the density of states  $D(\omega)$ , we sampled the frequencies at 48  $k$  points in the irreducible Brillouin zone of the 8-atom unit cell, and then coarse grained the resulting frequencies. Although we have checked explicitly that  $\omega(k) = vk$  at long wavelengths and therefore  $D(\omega) \propto \omega^2$ , this is not accurately represented in  $D(\omega)$  because of our coarse sampling of the Brillouin zone.

Since the existence of a photonic gap is a prerequisite for the existence of photonic bound states, we chose to begin our search for localized modes considering a diamond structure with a large gap to midgap ratio. Following the results of Ho, Chan, and Soukoulis,<sup>3</sup> we considered the periodic arrangement of spheres of air in a dielectric medium. We considered a material with dielectric constant of 35, and chose air spheres of radius  $0.29a$ , where  $a$  is the conventional lattice constant of the diamond cell. We

---

<sup>6</sup>Hecht and Zajac, *Optics*, (Addison-Wesley, Menlo Park, CA, 1979), p.315.



considered impurities of two types, air spheres in the dielectric region, which were located at the hexagonal site,<sup>7</sup> and dielectric spheres in the air region, which were located in the bond-center site.<sup>7</sup> We also studied a diamond lattice of dielectric spheres in air, and also found evidence for photonic bound states when an air sphere was placed at the bond-center site.

Since the applicability of a band structure breaks down when an impurity destroys the translational symmetry of the dielectric lattice, we have calculated the density of states of the impurity system. The density of states of the bulk diamond lattice and of diamond with a single impurity are shown in Fig. 1. As in Ref. 3, we see that the perfect diamond crystal has a gap in its photonic band structure. Placing an air sphere at the hexagonal site introduces a single state in the gap, at  $\omega_b$ . Since there are no traveling modes in the diamond lattice at  $\omega_b$  this must be a localized mode. This is verified in Fig. 2, which shows that the field is localized about the defect. In fact, we find that bound states whose frequencies are in the center of the gap have decay lengths as small as one lattice constant. We expect that this length will diverge as the bound state frequency approaches the continuum of extended states.

In Fig. 1 it appears that the localized modes are spread over a modest range of frequencies. This is purely an artifact of our supercell technique. Because we are considering an array of defect states each of which is localized over some finite distance, there is tunneling between localized states on neighboring impurities. It is this hopping between defect states that introduces a nonzero width to the impurity band. For simplicity, we chose the maximum of the defect density of states to identify the actual position of the impurity state. In the inset of Fig. 1, we show calculations of the bandwidth for increasing impurity separations. This bandwidth decreases exponentially, as expected for exponentially bound states. Of course, in an experiment the lattice will be of finite size,

---

<sup>7</sup>M. Lannoo and J. Bourgoin, *Point Defects in Semiconductors I*, (Springer-Verlag, Berlin, 1981), p. 8. J

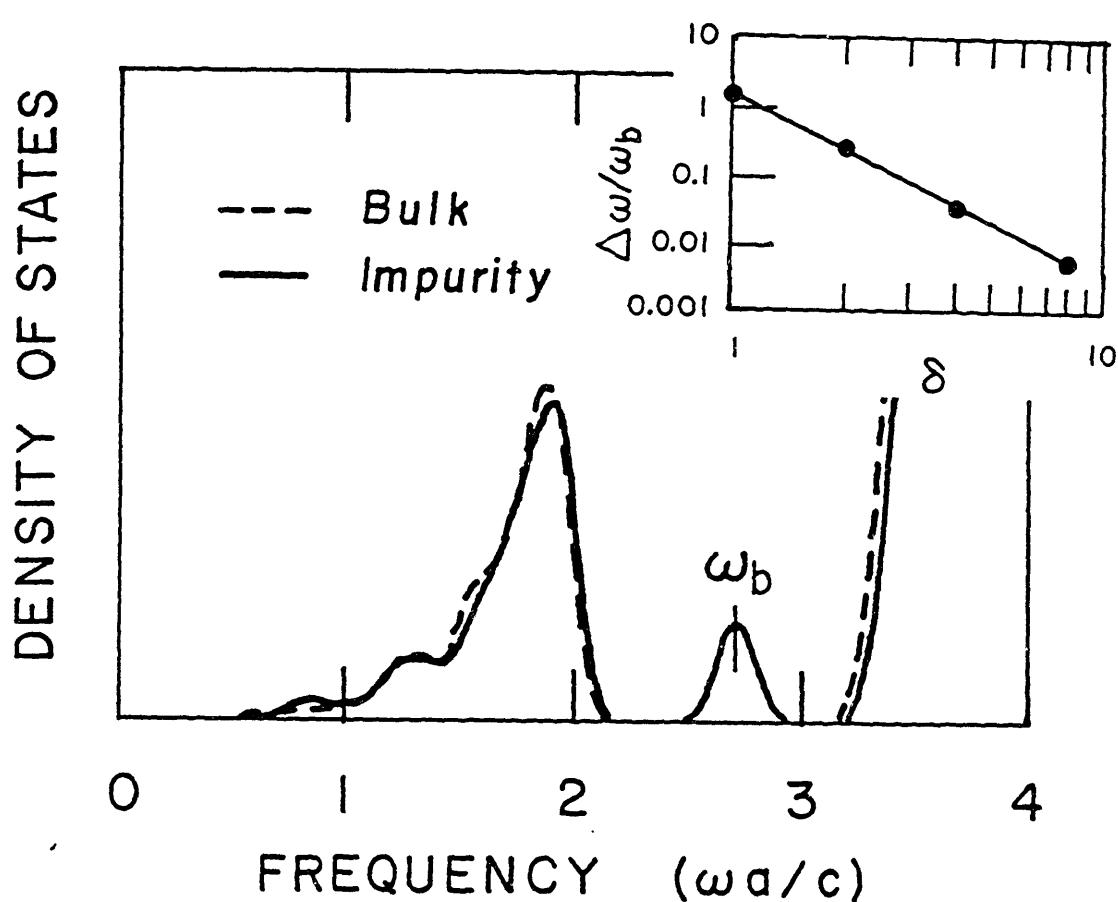


FIG. 1. Density of states in ideal diamond lattice (dashed line), and diamond lattice with defect of radius  $0.15a$  located at hexagonal site (solid line). Ideal diamond lattice has a gap in the density of states in which there are no propagating modes. The defect introduces a single localized state in this gap at  $\omega_b$ . Inset shows the bandwidth  $\Delta\omega$  normal to a plane of impurities as a function of the (111) impurity plane separation,  $\delta$ . Note that the bandwidth is exponential in the impurity separation, as expected for exponentially bound states. Units of frequency are  $2\pi/a$ , where  $a$  is lattice constant of the conventional unit cell, and  $c = 1$ .

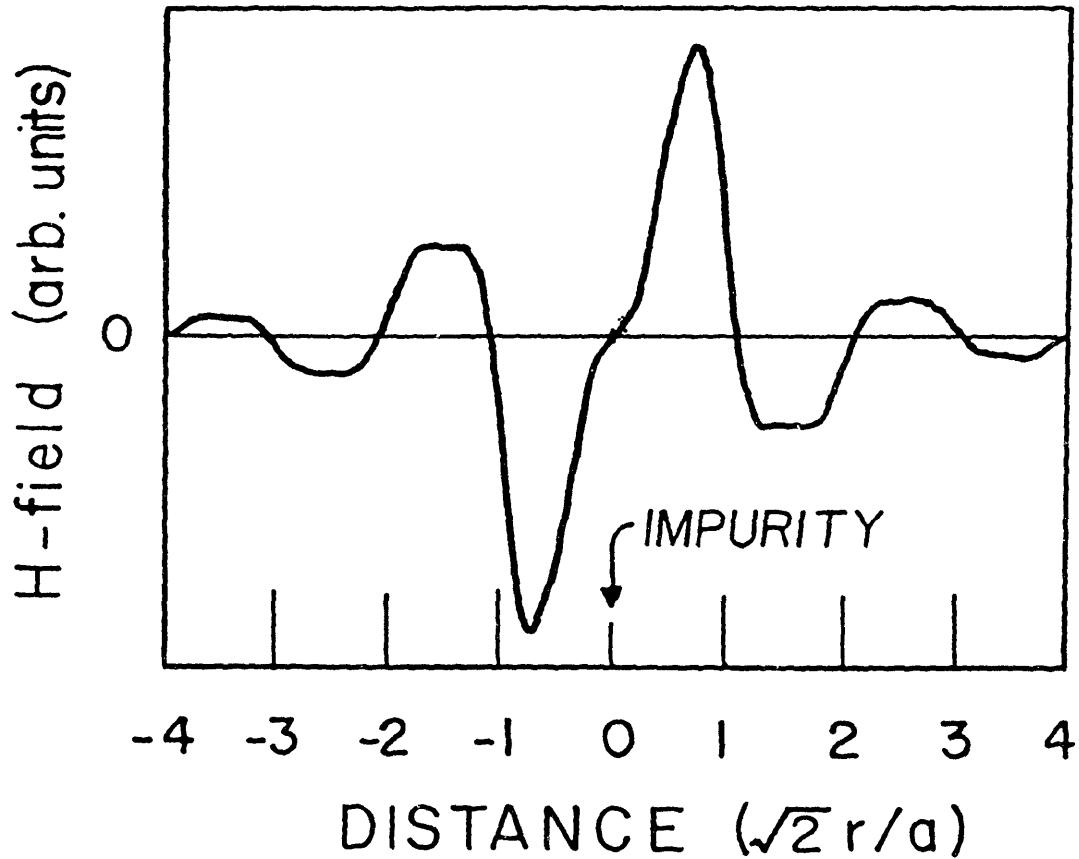


FIG. 2. H field associated with the state localized about a defect at the hexagonal site.  $H(r)$  is plotted along an axis which passes through the impurity at  $r = 0$ , and extends in the [110] direction. Because of a mirror symmetry through the (110) plane,  $H(r)$  lies in the [110] direction. Note that the localized state shows exponential decay with the sign alternating between neighboring cells.

and so the impurity mode will have some exponentially small amplitude at the walls of the lattice. This will allow the localized state to tunnel out, and will introduce a finite width to the frequency spectrum.

Qualitatively, we can understand this result by analogy with the more familiar case of impurities in a crystal with a band gap in its electronic structure. Since the wavelength of light is shorter in the dielectric, these regions are analogous to a region of deep potential in the crystal. Inserting a dielectric impurity adds an attractive potential. Sufficiently strong potentials can pull a state out of the photonic conduction band into the gap, and increasing the attractive character decreases the bound-state frequency, see Fig. 3. Similarly, adding an air sphere to the dielectric region is analogous to adding a repulsive potential which pushes a state out of the photonic valence band. In fact, the addition of an air sphere defect to the perfect diamond lattice decreased the number of valence states by one, and it created one state in the gap. We found one notable difference between air and dielectric impurities. While the presence of an air sphere creates a single, well-defined state in the gap, the presence of a dielectric sphere introduces a number of states in the gap, the lowest of which is doubly degenerate and is represented in Fig. 3.

Alternately, one can understand the localized mode as a three-dimensional Fabry-Perot interferometer.<sup>2</sup> Since there are no propagating modes in the dielectric material with frequencies in the gap, it behaves as a mirror to these frequencies. The defect, then, is surrounded by reflecting walls, and the localized state is analogous to the familiar resonances of a metallic cavity.

We have also considered potential microwave and millimeter-wave device applications for periodic dielectric materials as resonant cavities and filters. Materials with high dielectric constant are currently used to fabricate resonant cavities at microwave and millimeter-wave frequencies.<sup>8</sup> Dielectric cavities reduce the size and cost of

---

<sup>8</sup>M. Dydyk, in *Handbook of Microwave and Optical Components*, edited by Kai Chang (Wiley, New York, 1989), Vol. 1, p. 191. 

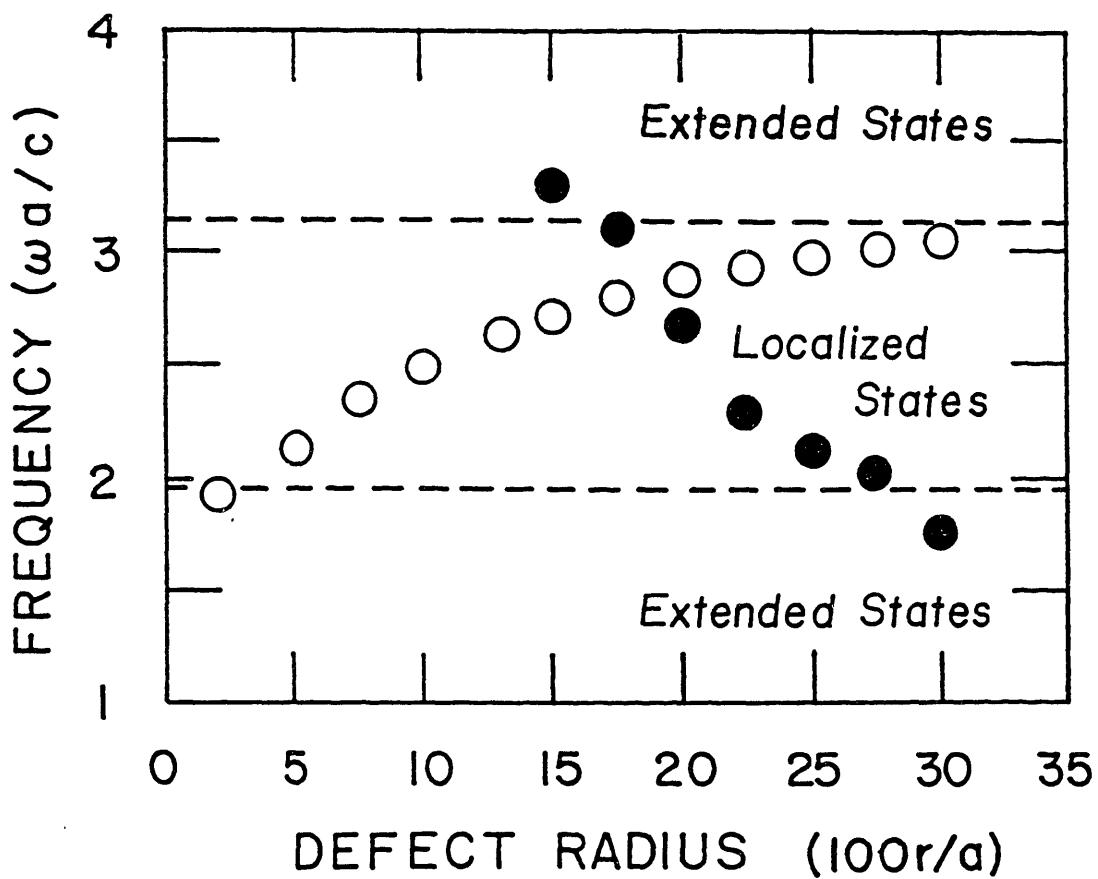


FIG. 3. Frequencies of localized modes due to lattice defects as a function of defect size. Open circles show frequencies of air spheres at the hexagonal site, and solid circles show frequencies of dielectric spheres at bond-centered sites. The dashed lines separate the regions of extended and localized states. The radius of the defect is in units of  $a/100$  and the frequencies are in units of  $2\pi/a$ , where  $a$  is the lattice constant of the conventional unit cell, and  $c = 1$ .

microwave and millimeter-wave circuits by replacing bulky discrete metallic waveguides and cavities with planar technology amenable to integrated circuit fabrication. However, all planar circuits require metallic shielding in order to reduce radiation losses and to control higher-order cavity modes. Since the fields attenuate only as  $1/r$  from a typical dielectric cavity, conduction losses due to the shielding can be appreciable. Embedding these dielectrics in a lattice will reduce these losses, since the field attenuates exponentially in the lattice region.

Because of the strong localization in these lattices, it seems likely that impurities in dielectric lattices will have less power dissipation than a well-polished metallic cavity. If the power dissipation indeed turns out to be less, then dielectric lattices would certainly be valuable as high-Q resonant cavities. Even if the power dissipation turns out to be greater, there are still advantages. As one attempts to work at higher wavelengths, the small size of a metallic cavity reduces the device power rating to levels that are too small for many applications. It may be possible to store larger electric fields in dielectric lattices without burning out the device.

A second potential application for periodic dielectric materials exploits the narrow line width of the forbidden state in a manner similar to a Fabry-Perot interferometer. Since a narrow transmission gap exists for all orientations, there are no collimation losses. Secondly, it may be possible to tune the filter by moving the impurity within the dielectric lattice.

In conclusion, we have performed calculations of the density of states of dielectric lattices containing defects. We find that defects can introduce localized modes, whose frequencies lie in the photonic band gap. We have studied size dependence of two types of defects, air spheres in dielectric and dielectric spheres in air. Finally, we have outlined potential device applications for localized modes in dielectric lattices.

We are grateful to Henry Mullaney for guidance on the microwave applications of local modes. We would also like to thank David Vanderbilt and Tomas Arias for helpful

discussions. Partial support for this work was provided by the Office of Naval Research Contract No. N00014-86-K-0158. Finally, one of us (A. M. R.) would like to acknowledge the support of the Joint Services Electronics Program.

**Chapter 6: Donor and Acceptor Modes in Photonic Band Structure  
in collaboration with**

**E. Yablonovitch, T. J. Gmitter, R. D. Meade, and K. D. Brommer**

There has been great progress recently in the creation of artificial three-dimensionally periodic dielectric structures which are to photon waves as semiconductor crystals are to electron waves. That is, these photonic crystals have a photonic band gap, a band of frequencies in which electromagnetic waves are forbidden<sup>1</sup>, irrespective of propagation direction in space. Both face-centered-cubic lattice<sup>2</sup> and diamond symmetry<sup>3</sup> dielectric structures have now been shown to produce a photonic band gap.

The photonic band gap is very interesting in its own right. It is an energy band in which optical modes, spontaneous emission, and zero-point fluctuations are all absent. Nevertheless, the photonic band gap might actually be at its most interesting when the perfect translational symmetry is disrupted in a controlled manner. For example, by introducing a known degree of disorder, mobility edges and Anderson localization transitions<sup>4</sup> can be investigated.

Lasers, perhaps the most interesting application, also require that the 3D translational symmetry should be broken. Even while spontaneous emission into all  $4\pi$  sr would be inhibited, a local electromagnetic mode is still necessary to accept the stimulated emission. In effect the local defect-induced structure resembles a Fabry-Perot cavity, except that it reflects radiation back upon itself in all  $4\pi$  spatial directions. Independently, Meade *et. al.*<sup>5</sup> have proposed that this could be accomplished by introducing a simple defect into the system.

The perfect three-dimensional translational symmetry of a dielectric structure can be lifted in either one of two ways: (1) Extra dielectric material may be added to one of the unit cells. We find that such a defect behaves very much like a donor atom in a semiconductor. It gives rise to donor modes which have their origin at the bottom of the

---

<sup>1</sup>E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987). *J*

<sup>2</sup>E. Yablonovitch, T. J. Gmitter, and K. M. Leung, Phys. Rev. Lett. **67**, 2295 (1991). *J*

<sup>3</sup>K. M. Ho, C. T. Chan, and C. M. Soukoulis, Phys. Rev. Lett. **65**, 3125 (1990). *J*

<sup>4</sup>S. John, Phys. Rev. Lett. **58**, 2486 (1987). *J*

<sup>5</sup>R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, Phys. Rev. B **44**, 13772 (1991). *J*

conduction band. (2) Conversely, translational symmetry can be broken by removing some dielectric material from one of the unit cells. Such defects resemble acceptor atoms in semiconductors. The associated acceptor modes have their origin at the top of the valence band. We will find that acceptor modes are particularly well-suited to act as laser microresonator cavities. Indeed it appears that photonic crystals made of sapphire or other low-loss dielectrics will make the highest-Q single-mode cavities (of volume  $\sim 1\lambda^3$ ) covering all electromagnetic frequencies above the useful working range of superconducting metallic cavities. The short-wavelength limit in the ultraviolet is set by the availability of optical materials with refractive index  $>2$ , the threshold index<sup>2,3</sup> for the existence of a photonic band gap.

For these experiments, we have chosen a face-centered-cubic (fcc) photonic crystal<sup>2</sup> employing non-spherical atoms. This fcc structure lends itself readily to microfabrication since it consists of intersecting drill holes<sup>2</sup> which can be made by reactive ion etching. While such microstructures have already been fabricated<sup>6</sup> in GaAs, we have chosen initially to investigate local defect modes in larger structures on the scale of 1 cm wavelengths. We selected a refractive index  $n \sim 3.6$  for the microwave material, matching that of the common semiconductors Si and GaAs. Experiment is supplemented by theoretical calculations of the photonic bound states.

Photonic crystals generally consist of a continuous three-dimensional web of dielectric material, made up of interconnecting ribs. The Wigner-Seitz unit cell of our photonic crystal<sup>2</sup> is the standard fcc rhombic dodecahedron<sup>2</sup> with "air atoms" created by drill holes centered on the top three faces, which exit through the bottom three faces. Figure 1 is a (1̄10) cross section of our photonic crystal cutting through the center of a unit cube. Shading represents dielectric material. The large dots are centered on the air atoms and the rectangular dashed line is a face-diagonal cross section of the unit cube. Such

---

<sup>6</sup>A. Scherer and B. P. van der Gaag (private communication).

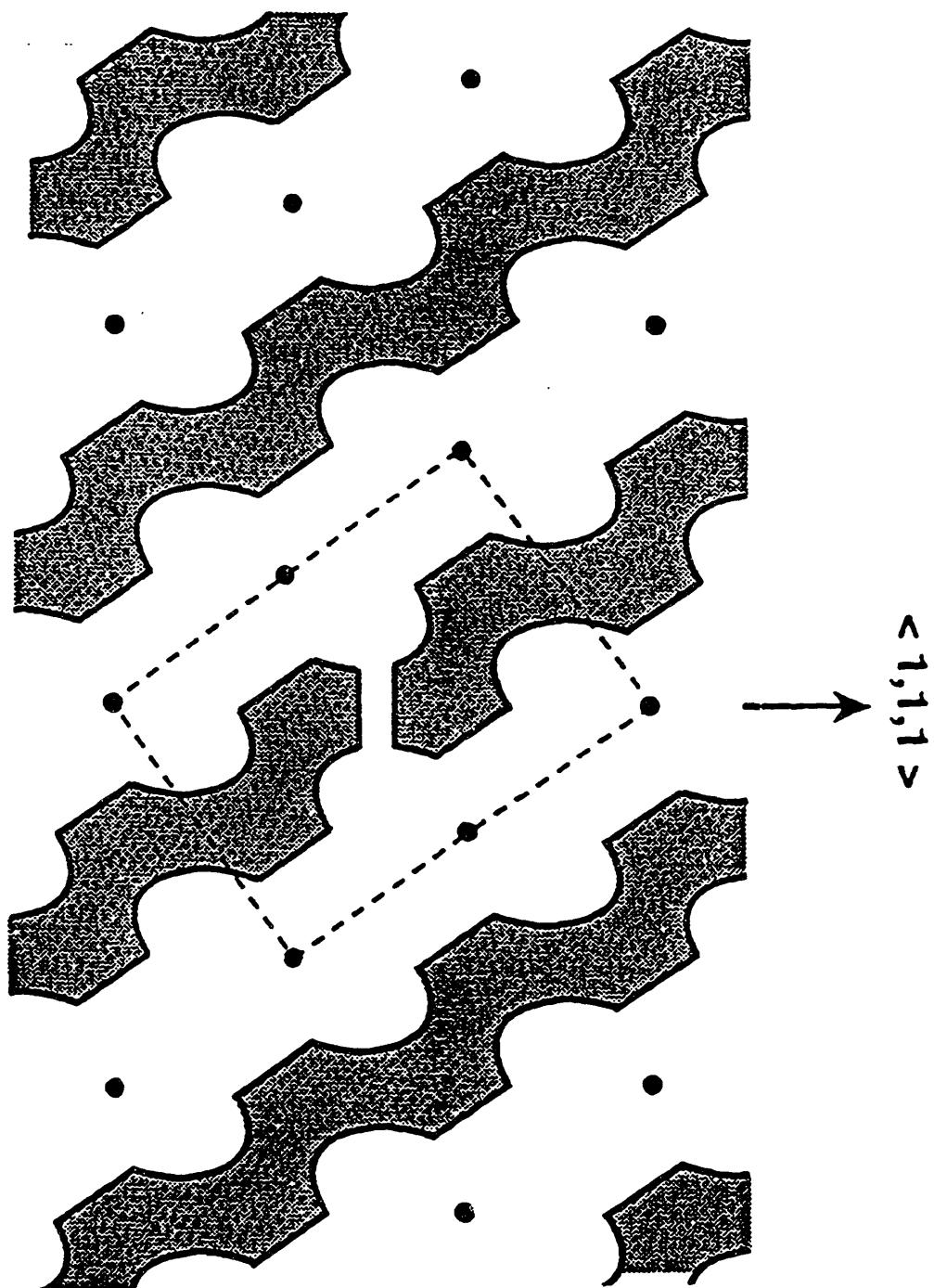


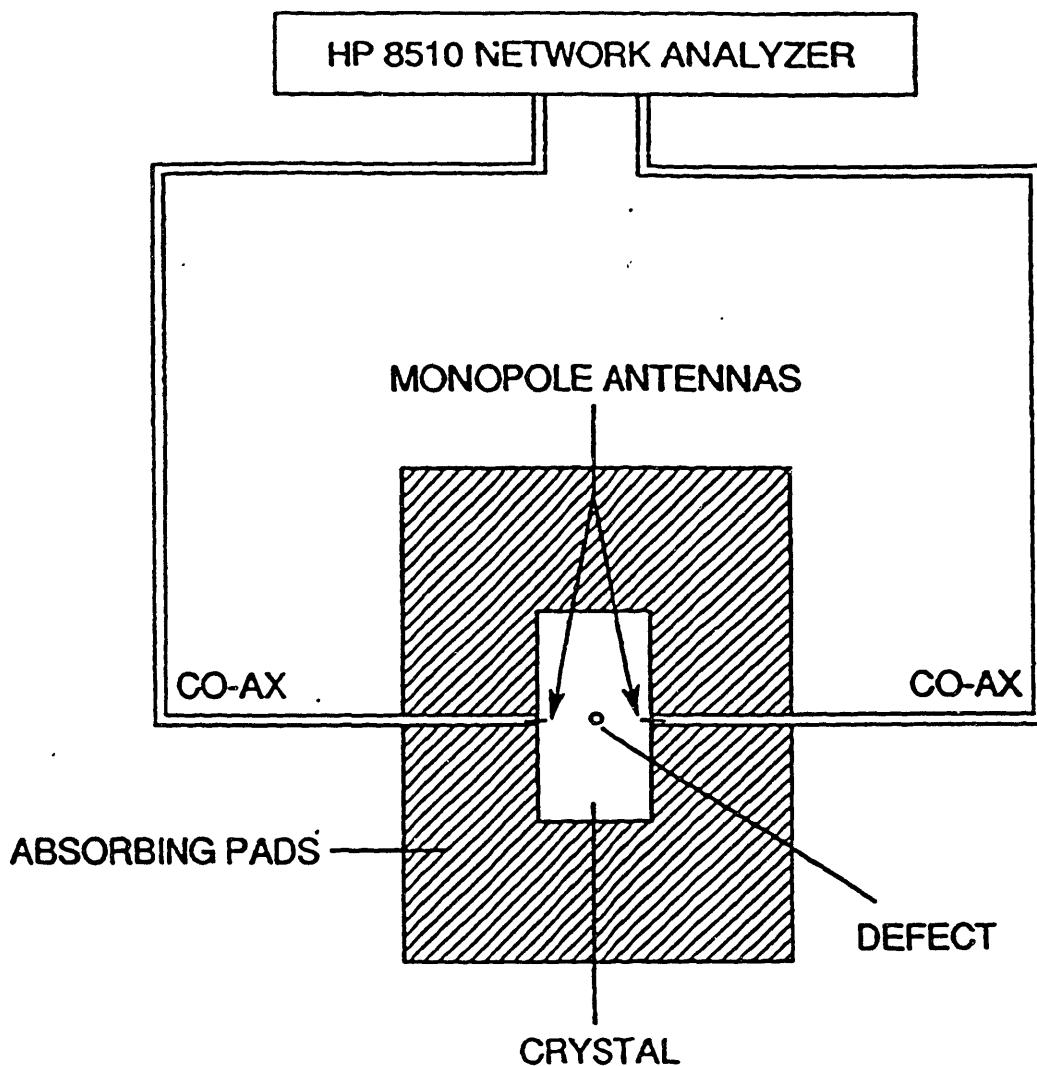
FIG. 1. A (110) cross-sectional view of our face-centered-cubic photonic crystal<sup>2</sup> consisting of nonspherical "air atoms" centered on the large dots. Dielectric material is represented by the shaded area. The rectangular dashed line is a face-diagonal cross section of the unit cube. Donor defects consisted of a dielectric sphere centered in an atom. We selected an acceptor defect as shown, centered in the unit cube. It consists of a missing horizontal slice in a single vertical rib.

structures are made simply by drilling three sets of holes  $35.26^\circ$  off vertical into the (111) top face.

Since we could design the structure at will, donor defects were chosen to consist of a single dielectric sphere centered in an air atom. Likewise, by breaking one of the interconnecting ribs, it is easy to create acceptor modes. We selected an acceptor defect as shown in Fig. 1, centered in the unit cube. It comprises a vertical rib which has a missing horizontal slice.

The heart of our experimental apparatus is a photonic crystal embedded in microwave absorbing pads as shown in Fig. 2. The photonic crystals were 8-10 atomic layers thick in the (111) direction. The cubic unit-cell length was  $a = 11$  mm and the hole diameter was 5.16 mm, leaving an empty volume fraction  $\sim 78\%$ . Monopole antennas, consisting of 6-mm pins, coupled radiation to the defect mode. The HP 8510 Network Analyzer was set up to measure transmission between the antennas. Figure 3(a) shows the transmission amplitude in the absence of a defect. There is very strong attenuation ( $\sim 10^{-5}$ ) between 13 and 16 GHz marking the valence- and conduction-band edges of the forbidden gap. This is a tribute to both the dynamic range of the network analyzer, and the sizable imaginary wave vector in the forbidden gap.

A transmission spectrum in the presence of an acceptor defect is shown in Fig. 3(b). Most of the spectrum is unaffected, except at the electromagnetic frequency marked "deep acceptor" within the forbidden gap. At that precise frequency, radiation "hops" from the transmitting antenna to the acceptor mode and then to the receiving antenna. The acceptor level frequency, within the forbidden gap, is dependent on the volume of material removed. Figure 4 shows the acceptor level frequency as a function of defect volume removed from one unit cell. When a relatively large volume of material is removed, the acceptor level is deep, as shown in Fig. 3(b). A smaller amount of material removed results in a shallow acceptor level, nearer the valence band. If the removed material volume



**FIG. 2.** Experimental configuration for the detection of local electromagnetic modes in the vicinity of a lattice defect. Transmission amplitude attenuation from one antenna to the other is measured. At the local mode frequency the signal hops by means of the local mode in the center of the photonic crystal producing a local transmission peak. The signal propagates in the (111) direction through 8-10 atomic layers.

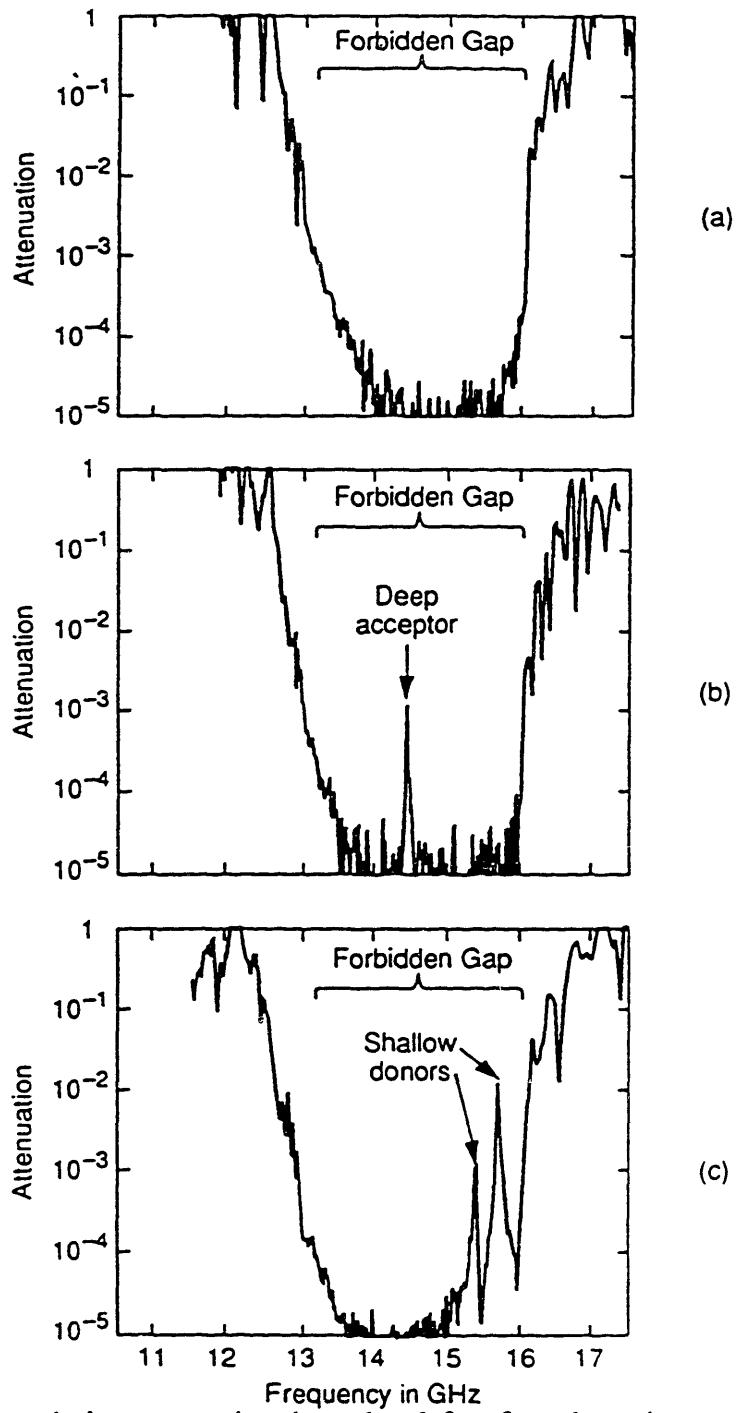


FIG. 3. (a) Transmission attenuation through a defect-free photonic crystal, as a function of microwave frequency. The forbidden gap falls between 13 and 16 GHz. (b) Attenuation through a photonic crystal with a single acceptor in the center. The large acceptor volume moved its frequency near midgap. The electromagnetic resonator  $Q$  was  $\sim 1000$ , limited only by the loss tangent of the dielectric material. (c) Attenuation through a photonic crystal with a single donor defect, an uncentered dielectric sphere, leading to two shallow donor modes.

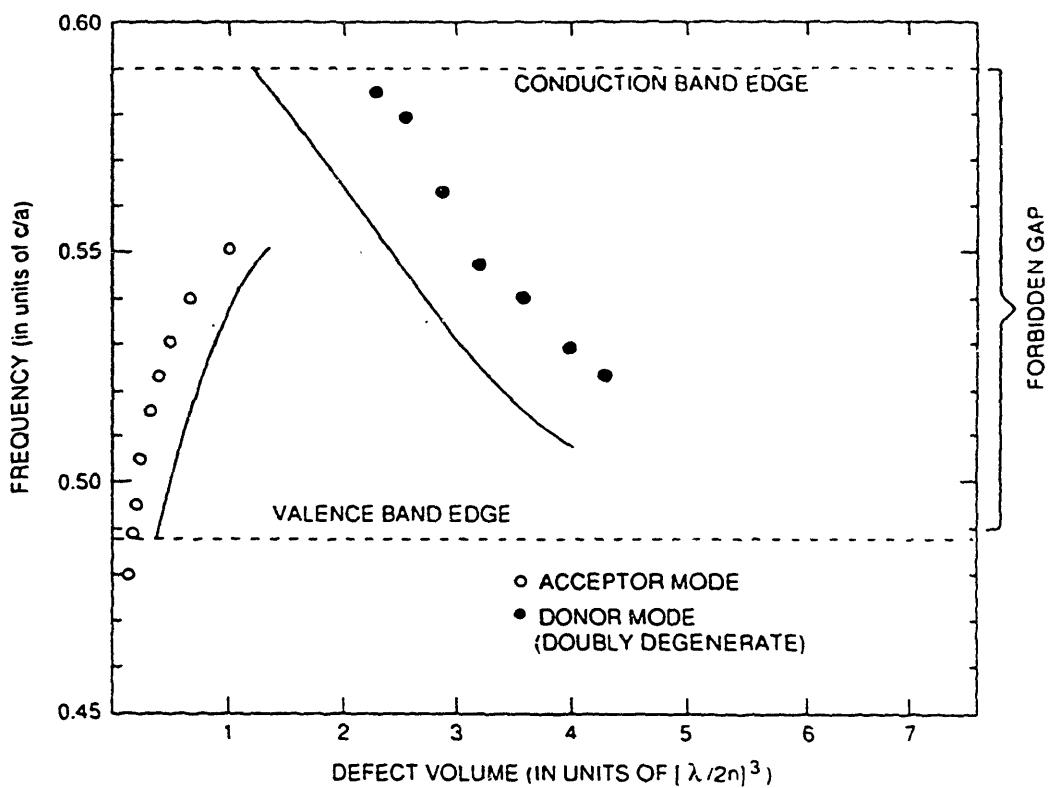


FIG. 4. Donor and acceptor mode frequencies as a function of normalized donor and acceptor defect volume. The points are experimental and the corresponding lines are calculated. Defect volume is normalized to  $(\lambda/2n)^3$ , where  $\lambda$  is the midgap vacuum wavelength and  $n$  is the refractive index. A finite defect volume is required to bind a mode in the forbidden gap.

falls below a threshold volume, the acceptor level falls within the continuum of levels below the top of the valence band, becoming metastable.

On an expanded frequency scale we can measure the resonator Q of the deep acceptor mode, which is  $Q \sim 1000$ , as limited by the loss tangent of the Emerson & Cuming Stycast material of which the photonic crystal was made.

The behavior of an off-center donor defect is shown in Fig. 3(c). In that case the donor volume was slightly above the required threshold for forming bound donor modes. Already two shallow donor modes can be seen in Fig. 3(c). When the donor is centered in the Wigner-Seitz unit cell, the two modes merge to form doubly degenerate donor modes as in Fig. 4. Single donor defects seem to produce multiple donor levels. Figure 4 gives the donor level frequency as a function of donor volume. As in the case of acceptors, there is a threshold defect volume required for the creation of bound modes below the conduction-band edge. However, the threshold volume for donor defects is almost 10 times larger than the acceptor threshold volume. Apparently this is due to the electric-field concentration in the dielectric ribs at the top of the valence band. Bloch wave functions at the top of the valence band are rather easily disrupted by the missing rib segment.

We have chosen in Fig. 4 to normalize the defect volume to a natural volume of the physical system,  $(\lambda/2n)^3$ , which is basically a cubic half wavelength in the dielectric medium. More specifically,  $\lambda$  is the vacuum wavelength at the midgap frequency, and  $n$  is the refractive index of the dielectric medium. Since we are measuring a dielectric volume, it makes sense to normalize to a half-wavelength cube as measured at the dielectric refractive index. Based on the reasonable scaling of Fig. 4, our choice of volume normalization would seem justified. (Experimentally, the odd-shaped defect volumes were measured by weighing the samples.)

It is interesting to compare our local modes to those of a one-dimensional Fabry-Perot resonator, constructed in the usual manner of quarter-wave multilayer dielectric mirrors. In such a resonator the mirrors face each other and are usually separated by an

integral number of half wavelengths. The net effect of the left and right mirrors facing each other is that they combine to form a monolithic 1D periodic dielectric structure, but with a quarter wavelength of phase slip introduced into the very center. This same quarter wavelength of phase slip is often employed in distributed feedback lasers<sup>7</sup>, effectively converting them into Fabry-Perot resonators. The standing-wave mode is sometimes<sup>8</sup> regarded as a bound state split off into the 1D band gap and localized to the quarter-wave defect in the periodic structure. In one dimension, however, nothing requires that the defect be one-quarter wavelength long. A phase slip, no matter how small, allows a bound mode to form in the forbidden gap, usually near a band edge. This is in distinction to Fig. 4 where a finite-sized defect volume is required to bind a mode.

This is similar to the comparison between 1D and 3D quantum mechanics. In one dimension even an infinitesimal quantum well will bind<sup>9</sup> a state. In three dimensions, a finite-depth potential well is required to produce<sup>9</sup> a bound state. Figure 4 is telling us that the same requirement for a finite defect volume applies to 3D confined photon modes. John and Wang<sup>10</sup> have shown that the requirement for a finite-volume-integrated polarizability can be satisfied even by a single resonant atom if undamped by nonradiative decay. Then an individual atom is capable of binding a local electromagnetic mode at its resonant transition frequency. They call this a photon-atom bound state. Inhibited spontaneous emission is accompanied with strong self-dressing of the atom by its own localized radiation fields, leading to anomalous Lamb shifts.

We have also performed calculations of the frequency spectrum of these dielectric systems. In order to solve for the electromagnetic modes, we expand Maxwell's equations in plane waves, employing the techniques of Ho, Chan, and Soukoulis<sup>3</sup>. We employ the

<sup>7</sup>H. Kogelnik and C. V. Shank, *J. Appl. Phys.* **43**, 2328 (1972). *J*

<sup>8</sup>H. A. Haus and C. V. Shank, *IEEE J. Quantum Electron.* **12**, 532 (1976); S. L McCall and P. M. Platzman, *IEEE J. Quantum Electron.* **21**, 1899 (1985). *J*

<sup>9</sup>L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1949), see pp. 37 and 77. *B*

<sup>10</sup>S. John and J. Wang, *Phys. Rev. Lett.* **64**, 2418 (1990). *J*

supercell method, placing one defect in a repeated cell of dielectric material. Our calculations were performed in supercells containing eight fcc Wigner-Seitz cells, and we have performed tests on larger supercells containing eight conventional fcc cubic unit cells (32 Wigner-Seitz cells). Because the distance between defects is relatively small, there is significant overlap between the localized modes on neighboring defects. This overlap manifests itself as a dispersion of the impurity band, and the frequency of the bound state is taken to be the band center. We expanded the  $\mathbf{H}(\mathbf{r})$  magnetic field in a basis of 4000 plane waves, and the lowest eigenvalues were obtained by the Natarajan and Vanderbilt<sup>11</sup> iterative diagonalization technique.

The solid lines in Fig. 4 show good qualitative agreement with the experimental points. The discrepancy between theory and experiment is larger for defects in supercells than for purely bulk systems, where better than 1% agreement has been achieved. There are two important reasons for this larger discrepancy. First, because of the larger unit-cell size, it is necessary to work at a lower plane-wave cutoff. Second is the small number of primitive unit cells used in the supercell approximations. As in Ref. 5, the dispersion of the impurity band is quite large in the eight-atom supercell, ~(20-80)% of the gap width for various states. Increasing the supercell to 32 atoms reduces the bandwidth by 80%.

The extension of these 3D dielectric resonators to laser wavelengths is marked by the following considerations. The refractive index we have chosen to work with ( $n = 3.6$ ) is a good match to the common semiconductors of which lasers are made. Furthermore, the fcc geometry we are using has begun<sup>6</sup> to be microfabricated in GaAs. Such microresonators will be particularly valuable for making tiny low-threshold lasers. But their value is greatly increased if the photonic band gap inhibits spontaneous emission. In a semiconductor laser, this would lead to near unity quantum efficiency into the lasing mode. Photon number state squeezing<sup>12</sup> into that mode would be greatly enhanced. Inhibited

---

<sup>11</sup>R. Natarajan and D. Vanderbilt, J. Comput. Phys. **81**, 218 (1989).

<sup>12</sup>Y. Yamamoto and S. Machida, Phys. Rev. A **35**, 5114 (1987).

spontaneous emission requires that the broad semiconductor luminescence band should be centered within the forbidden gap. In semiconductor lasers, population inversion and gain first appear at the red edge of the luminescence band, i.e., nearer to the valence band. This suggests that acceptor modes would be the appropriate type for lasers.

The other advantage for acceptor-mode laser cavities is associated with our acceptor defect geometry. The vertical rib with a missing horizontal slice, as in Fig. 1, can be readily microfabricated. It should be possible to create it in III-V materials by growing an aluminum-rich epitaxial layer and lithographically patterning it down to a single dot the size of one of the vertical ribs. After regrowth of the original III-V composition and reactive ion etching of the photonic crystal, HF acid etching, whose<sup>13</sup> selectivity  $\geq 10^8$ , will be used to remove the Al-rich horizontal slice from the one rib containing such a layer. The resonant frequency of the microcavity can be controlled by the thickness of the Al-rich sacrificial layer.

We would like to thank John Gural for his skillful machining of the photonic crystals. The work at M. I. T. was partially supported by ONR Contract No. N00014-86-K-0158 and U. S. JSEP Contract No. DAAL-03-86-K-0002. R. D. M. acknowledges valuable discussions with K. M. Ho.

---

<sup>13</sup>E. Yablonovitch, T. J. Gmitter, J. P Harbison, and R. Bhat, Appl. Phys. Lett. **51**, 2222 (1987). J

**Chapter 7: The Existence of a Photonic Band Gap in Two Dimensions**  
in collaboration with  
**R. D. Meade and K. D. Brommer**

Since the first suggestion that a dielectric material with three-dimensional periodicity could reflect light incident from any direction, there has been rapid progress in the newly emergent field of photonic band gap materials<sup>1,2,3,4,5,6,7,8,9</sup>. Several structures which exhibit a photonic band gap have been discovered<sup>5,6</sup> and a number of important properties of these materials, such as the localization of light at defects<sup>7,8</sup> and surfaces<sup>9</sup>, have been discussed. The related problem of dielectric crystals which have a two-dimensional periodicity in the  $\hat{x}$ - $\hat{y}$  plane and are continuous in the  $\hat{z}$  direction is also beginning to receive some attention<sup>10,11,12,13</sup>.

In previous work,<sup>10,11</sup> dielectric structures which had band gaps for individual polarizations were found. In this paper, we present a two-dimensional structure in which these band gaps overlap to form a complete gap which prevents the in-plane propagation of light of any polarization. We first employ theoretical calculations to discover such a system and then perform microwave transmission experiments to verify that light of any polarization cannot propagate in this structure. While this work concentrates on the propagation of light in the two-dimensional plane, we conclude this paper by discussing the out-of-plane photonic band structure.

In a two-dimensional photonic crystal, like its three-dimensional counterpart, the light propagates as Bloch waves,

<sup>1</sup>E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987).

<sup>2</sup>E. Yablonovitch and T. J. Gmitter, Phys. Rev. Lett. **63**, 1950 (1989).

<sup>3</sup>K. M. Leung and Y. F. Liu, Phys. Rev. Lett. **65**, 2646 (1990).

<sup>4</sup>Ze Zhang and Sashi Satpathy, Phys. Rev. Lett. **65**, 2650 (1990).

<sup>5</sup>K. M. Ho, C. T. Chan, and C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990).

<sup>6</sup>E. Yablonovitch, T. J. Gmitter, and K. M. Leung, Phys. Rev. Lett. **67**, 2295 (1991).

<sup>7</sup>R. D. Meade, K. D. Brommer, A. M. Rappe, J. D. Joannopoulos, Phys. Rev. B **44**, 13772 (1991).

<sup>8</sup>E. Yablonovitch, T. J. Gmitter, R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, Phys. Rev. Lett. **67** 3380 (1991).

<sup>9</sup>R. D. Meade, K. D. Brommer, A. M. Rappe, J. D. Joannopoulos, Phys. Rev. B **44**, 10961 (1991).

<sup>10</sup>M. Plihal, A. Shambrook, A. A. Maradudin, P. Sheng, Optics Communications **80**, 199 (1991).

<sup>11</sup>M. Plihal and A. A. Maradudin, Phys. Rev. B **44**, 8586 (1991).

<sup>12</sup>S. L. McCall, P. M. Platzman, R. Dalichaouch, D. Smith, and S. Schultz, Phys. Rev. Lett. **67**, 2017 (1991).

<sup>13</sup>W. Robertson, G. Arjavalingam, R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, Phys. Rev. Lett. **68**, 2023 (1992).

$$\mathbf{H}(\mathbf{r}) = \mathbf{u}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} . \quad (1)$$

To understand this propagation let us consider the reciprocal space of a two-dimensional system. Since the crystal has discrete translational symmetry in the  $\hat{x}$ - $\hat{y}$  plane, there is a finite in-plane Brillouin zone described by in-plane wave vectors  $\mathbf{k}_1$  and  $\mathbf{k}_2$  with  $-G_1/2 < k_1 < G_1/2$  and  $-G_2/2 < k_2 < G_2/2$ , where  $\mathbf{G}_1$  and  $\mathbf{G}_2$  are the reciprocal lattice vectors. In the  $\hat{z}$  direction, however, the system is continuous and so the wave vector can assume any value  $-\infty < k_z < \infty$ . Since the system has no index contrast in the  $\hat{z}$  direction, a wave traveling in this direction is not scattered. Thus, we argue on physical grounds that there cannot be a band gap for light propagating in the  $\hat{z}$  direction with  $k_1 = k_2 = 0$ . In the first part of this paper, we will restrict our attention to the in-plane band structure  $k_z = 0$ . With this assumption, we can take advantage of the mirror symmetry to restrict the fields as either odd or even under  $M_z$ . Thus we can catalogue the electromagnetic modes as one of two types, either  $(H_x, H_y, E_z)$  or  $(E_x, E_y, H_z)$  which we will refer to as  $\mathbf{E}$  perpendicular to the plane or  $\mathbf{E}$  parallel to the plane respectively.

In order to calculate the electromagnetic frequency spectrum of these dielectric lattices, we have employed the computational techniques described in references 3-5. Briefly, the macroscopic Maxwell's equations can be rearranged to yield the eigenvalue equation

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

The magnetic field  $\mathbf{H}(\mathbf{r})$  can be expressed as a sum of plane waves,

$$\mathbf{H}(\mathbf{r}) = \sum_{\mathbf{G}} \sum_{\lambda=1,2} h_{\mathbf{G},\lambda} \hat{\mathbf{e}}_{\lambda} e^{i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}} , \quad (3)$$

where  $\mathbf{k}$  is in the Brillouin zone,  $\mathbf{G}$  is summed over the reciprocal lattice, and  $\hat{\mathbf{e}}_\lambda$  are polarizations orthogonal to  $(\mathbf{k} + \mathbf{G})$ . By limiting the size of wave vector included in this plane-wave expansion, a finite sum is achieved. The eigenvalue equation is now be expressed in matrix form,

$$\sum_{\mathbf{G}', \lambda'} H_{\mathbf{G}, \mathbf{G}'}^{\lambda, \lambda'} h_{\mathbf{G}', \lambda'} = \omega^2 h_{\mathbf{G}, \lambda}, \quad (4)$$

where  $c = 1$  and

$$H_{\mathbf{G}, \mathbf{G}'} = |\mathbf{k} + \mathbf{G}| |\mathbf{k} + \mathbf{G}'| \epsilon_{\mathbf{G}, \mathbf{G}'}^{-1} \begin{pmatrix} \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_2' & \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_1' \\ \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2' & \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_1' \end{pmatrix} \quad (5)$$

and  $\epsilon_{\mathbf{G}, \mathbf{G}'}^{-1}$  is the Fourier transform of the dielectric function  $\epsilon(\mathbf{r})$ . This eigenvalue equation can now solved by standard numerical techniques yielding the normal mode coefficients and frequencies of the electromagnetic modes. Because of the two-dimensional nature of the problems considered, plane-wave convergence was achieved with a relatively small number of plane waves and we estimate that by including  $\sim 140$  plane waves per polarization the electromagnetic mode frequencies were obtained to better than 1%. This technique provides a simple and powerful method to solve problems in electrodynamics which takes full account of the vector nature of the electromagnetic radiation.

As in the three-dimensional case, our primary objective is to create a dielectric system which exhibits a complete photonic band gap, a set of frequencies in which light of any polarization cannot propagate in any in-plane direction. In order to find such a structure we searched all filling fractions of square, triangular and honeycomb lattices. We also considered dielectric rods in air, as well as air columns drilled in dielectric. At the dielectric contrast of GaAs ( $\epsilon = 13$ ), the only combination which was found to have a photonic band gap in both polarizations was the triangular lattice of air columns in dielectric (see Fig. 1). The maximum band gap we find is  $\omega_{\text{Gap}}/\omega_{\text{Midgap}} = 18.6\%$ , which occurs

### Photonic Band Structure of Triangular Lattice of Air Columns

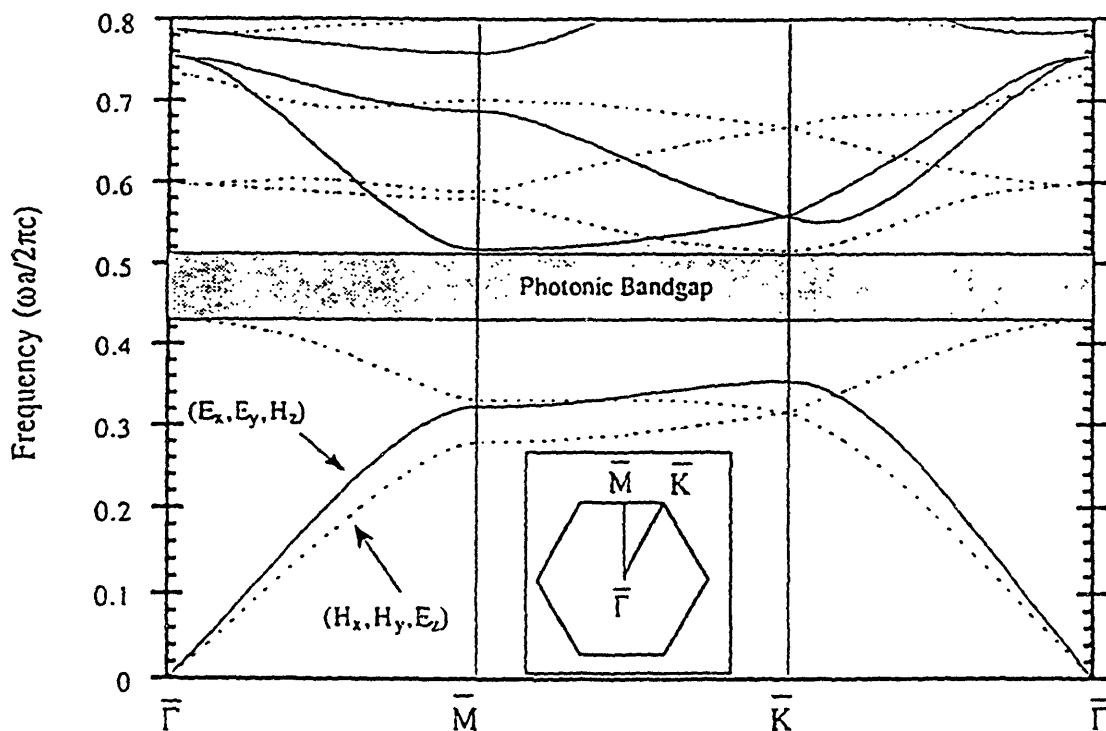


FIG. 1. Frequencies of the lowest photonic bands for a triangular lattice of air columns ( $\epsilon_{air} = 1$ ) drilled in dielectric ( $\epsilon_d = 13$ ). The band structure is plotted along special directions of the in-plane Brillouin zone ( $k_z = 0$ ), as shown in the lower inset. The radius of the air columns is  $r = 0.48 a$ , where  $a$  is the in-plane lattice constant. The solid(dashed) lines show the frequencies of bands which have the electric field parallel(perpendicular) to the plane. Notice the photonic band gap between the third and fourth bands.

when the radius of the air columns was  $r_{\text{Max}} = 0.48 a$ , where  $a$  is the in-plane lattice constant. A gap appears in this structure for dielectric contrast of 7.2 or greater (see Fig. 2). This photonic band gap is somewhat unusual because it occurs between the third and fourth bands, whereas the photonic band gaps in three-dimensional structures have always occurred between the second and third bands.

We verified our theoretical predictions by constructing a two-dimensional photonic band gap material from a dielectric block and measuring its transmissivity. We fabricated the sample from a piece of C-Stock 265 manufactured by Cuming Corp. This material consists of glass spheres and aluminum flake mixed in a binder to give an artificial dielectric with a constant  $\epsilon_a \sim 13-15$  (we assume  $\epsilon = 13$ ). This material was chosen because it was inexpensive and easy to machine. Using a numerically-controlled drill, the sample could be constructed in about four hours. An 8x14 array of air columns of diameter 0.992 cm were drilled in a triangular lattice with lattice constant 1.044 cm and so  $r/a = 0.475$ , close to the optimal theoretical ratio. Based on our calculations, we expected a gap in the range 12.3-14.6 GHz.

We performed free-space transmission measurements at an antenna test range. We mounted the photonic band-gap material inside a 7.5 by 15 cm aperture cut into a 90 by 90 cm sheet of radar absorbing material. A broad band horn antenna illuminated the sample at a range of 5.5 m. The 3 dB beam width of the radiated pattern was 60 cm at this range, ensuring flat, plane-wave illumination across the full extent of the sample. A notch antenna mounted directly behind the sample received the transmitted signal. An HP8510 network analyzer measured the power transmitted between the antennas. We used a chirped waveform along with digital range gating to eliminate extraneous reflections from objects near the sample. The sensitivity of our system was measured at 70 dB. This level of dynamic range enabled us to study detailed structure inside the photonic band gap, especially to compare the transmitted power of different photonic band gaps. Transmission amplitude was measured from 2 to 18 GHz at different sample orientations as the sample

### Maximum Gap Size and Optimal Radius

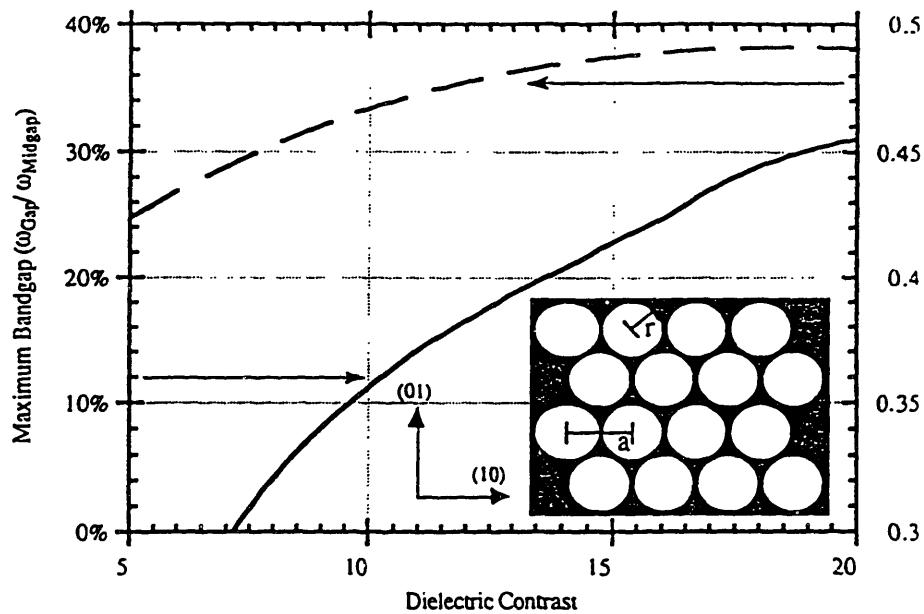


FIG. 2. The maximum photonic band gap achievable for different dielectric contrast (solid line), and the optimal radius for which the maximum band gap is achieved (dashed line). The band gap is expressed as a percentage of the midgap frequency, and the radius in units of the lattice constant. Note that the band gap first appears at a dielectric contrast of 7.2.

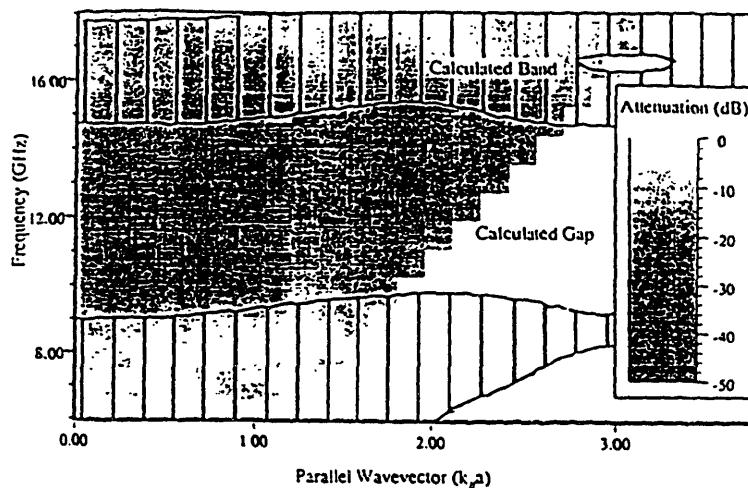
was rotated on an antenna pedestal. Both polarizations were studied by rotating the antennas. Two separate pairs of antennas were required to cover the entire 2 to 18 GHz range.

In Fig. 3, we display the experimental transmission spectrum  $T(\omega, k_{\parallel})$  and the calculated band structure of our two-dimensional structure. Note that there is no transmission for either polarization in the electromagnetic band gap, between 13-15.5 GHz. We anticipate if there are electromagnetic modes at  $(\omega, k_{\parallel})$ , then the light will travel through the sample and the transmission  $T(\omega, k_{\parallel})$  will be large. On the other hand, if the density of states is zero at  $(\omega, k_{\parallel})$ , then the light will decay into the sample and  $T(\omega, k_{\parallel})$  will be greatly reduced. These expectations are fulfilled in Fig. 3(a), which shows the transmission spectrum for light which is polarized with electric field normal to the plane. As we can see in this figure, the transmission is large (light regions) in the area where there are traveling modes (shaded by vertical lines). In the gap, however, the fields are strongly attenuated (dark regions) with transmission dropping by five orders of magnitude.

Figure 3(b) shows the transmission spectrum for light which is polarized with electric field perpendicular the plane. This polarization has two gaps, a complete gap at 13 GHz, and a mini gap at 8.5 GHz. The attenuation is small for the mini gap, with the transmission dropping by only 20 dB. The complete gap begins at 12.5 GHz, as we expect theoretically, and the attenuation remains large (~50 dB) until 18 GHz, which is the limit of our experimental resolution. However, we expect theoretically that the material should transmit light via the band centered at 16 GHz. Although this seems puzzling, it is quite similar to the results of Robertson *et al.*<sup>13</sup> These authors found that symmetry constraints prevented certain bands from transmitting light which was incident as a plane wave.

Until this point, we have considered only the case in which light propagates in the two-dimensional plane ( $k_z = 0$ ). However, for some applications it is important to understand the out-of-plane band structure, shown in Fig. 4. The inset to this figure

(a) Transmission Spectrum for E-Field // to 2d Plane



(b) Transmission Spectrum for E-Field  $\perp$  to 2d plane

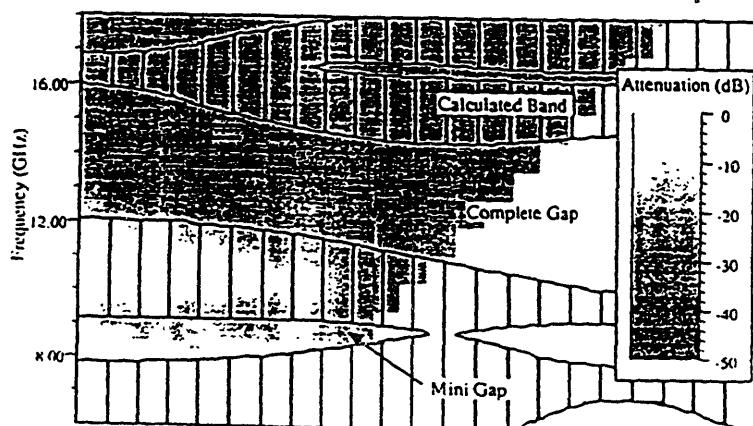


FIG. 3. The transmission spectrum of our two-dimensional structure for light polarized with electric field parallel to the 2D plane (a) and perpendicular to the 2D plane (b). The transmission amplitude  $T(\omega, k_{\parallel})$  is plotted for different frequencies and parallel wave vectors. The attenuation is displayed in grayscale, with the light regions representing large transmission and the dark regions representing large attenuation. Variations in the attenuation are due to reflection, not absorption. Note that there is no transmission for either polarization in the electromagnetic band gap, between 13-16 GHz. Because the sample was illuminated along the (01) face (see Fig. 2, inset), the parallel wave vector refers to variation in the (10) direction. The vertical lines in the figure designate the regions of  $(\omega, k_{\parallel})$  in which calculations find electromagnetic modes. These modes may have any wave vector along the (01) direction, but have  $k_z = 0$ .

shows the frequency dependence of the lowest band as  $k_z$  is varied. When  $k_z = 0$ , this lowest band spans a broad range of frequencies (see Fig. 1). However, in the inset to Fig. 4 we see that as  $k_z$  increases the lowest band becomes flat and so the in-plane bandwidth tends to zero. Figure 4 also plots the frequency of the lowest six bands as  $k_z$  is varied. It shows both  $\omega(\bar{\Gamma}, k_z)$  and  $\omega(\bar{K}, k_z)$ , corresponding to the frequency change of the bands which start at  $\bar{\Gamma}$  and  $\bar{K}$  respectively. Clearly, as  $k_z$  increases the bandwidth of each of the photonic bands tends towards zero.

Thus, the primary feature of the out-of-plane band structure is that as  $k_z$  increases, the in-plane bandwidth tends to zero. We can understand this simply. For large  $k_z$ , the light is totally internally reflected inside the dielectric regions, just as in an optical waveguide. Because there is very little overlap between light trapped in neighboring dielectric waveguides, the modes are decoupled and so is the bandwidth tends to zero. This is especially true for modes which have  $\omega \ll c k_z$ . Recall that in the vacuum region, light obeys the wave equation

$$\nabla^2 \mathbf{H} = -\frac{\omega^2}{c^2} \mathbf{H} . \quad (6)$$

If the fields oscillate along the z axis as

$$\mathbf{H}(\mathbf{r}) = \mathbf{u}(x, y) e^{ikz} , \quad (7)$$

then the wave equation reduces to

$$\nabla_{\parallel}^2 \mathbf{u} = (k_z^2 - \frac{\omega^2}{c^2}) \mathbf{u} . \quad (8)$$

Thus, if  $\omega \ll c k_z$  then the fields are strongly exponentially decaying in the vacuum region, and so the overlap between modes in neighboring regions of dielectric vanishes,

### The Out-of-Plane Band Structure

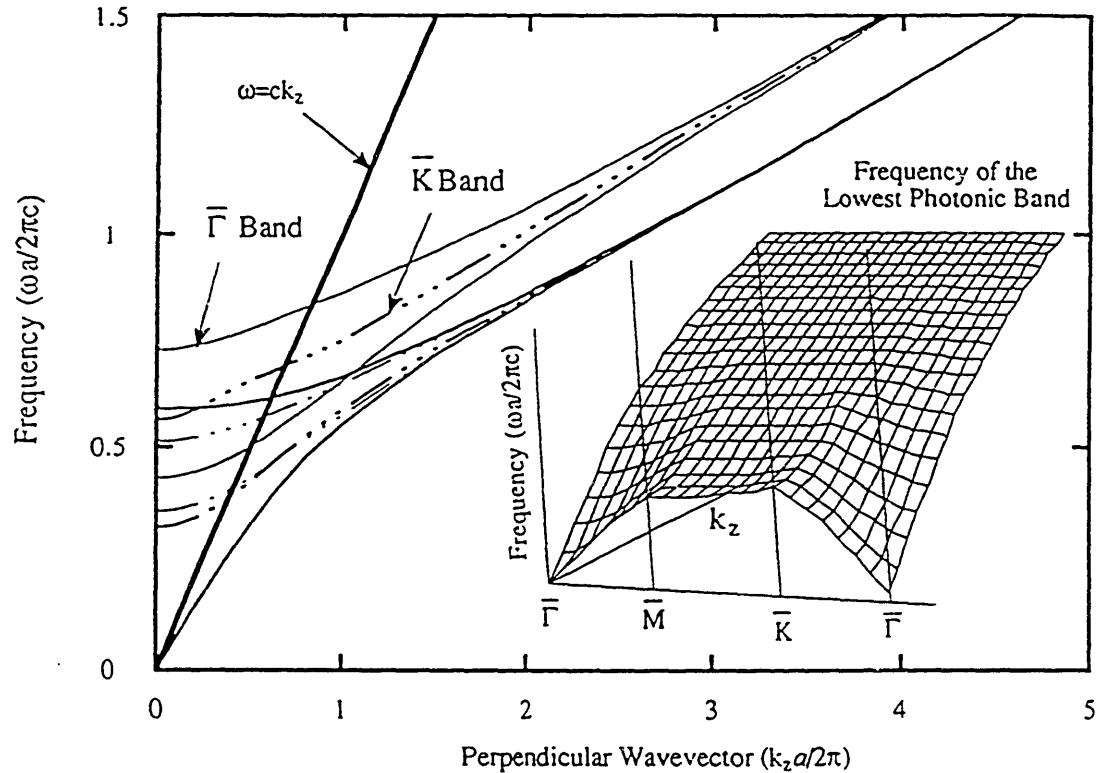


FIG. 4. The out-of-plane band structure of the triangular lattice of air columns. The frequencies of the lowest six bands are shown versus  $k_z$ . The frequencies of the bands which start at  $\bar{\Gamma}$ ,  $\omega(\bar{\Gamma}, k_z)$ , are plotted using solid lines, whereas the frequencies of the bands which start at  $\bar{K}$ ,  $\omega(\bar{K}, k_z)$ , are plotted using dashed-dotted lines. The thick(thin) lines correspond to doubly(singly) degenerate modes. The straight line  $\omega = c k_z$  separates the regions in which the modes are oscillatory ( $\omega > c k_z$ ) in the air regions from those which are evanescent ( $\omega < c k_z$ ) in the air regions. The inset to this figure shows the frequency dependence of the lowest band as  $k_z$  is varied. Note that as  $k_z$  increases the lowest band becomes flat.

and the bandwidth goes to zero. The bands in Fig. 4 display this behavior. Above the light cone  $\omega > c k_z$ , the bands have a large dispersion, while for  $\omega \ll c k_z$  this bandwidth goes to zero.

Experiments were funded by Lockheed Sanders. We thank R. Gilbert and H. Mullaney for helpful discussion and D. Larochelle for assisting in designing the experiment. Partial support for this work was provided by the Office of Naval Research Contract No. N00014-86-K-0158. Finally, one of us (A. M. R.) would like to acknowledge the support of the Joint Services Electronics Program.

Addendum: Since submission, another group<sup>14</sup> has reported the existence of a band gap common to both polarizations for a square lattice of air columns in dielectric. This band gap is appreciable at large index contrast and vanishes at  $\epsilon = 12.3$ .

---

<sup>14</sup>P. Villeneuve and M. Piche, submitted to Phys. Rev. Lett. 3

**Chapter 8: Measurement of Photonic Band Structure in a Two-Dimensional Periodic Dielectric Array**

in collaboration with

**W. M. Robertson, G. Arjavalingam, R. D. Meade, and K. D. Brommer**

The propagation of electromagnetic radiation in periodic dielectric structures has received much recent experimental and theoretical attention. The existence in such structures of photonic band gaps--frequency intervals in which no photon modes are allowed--leads to a variety of phenomena of both fundamental<sup>1,2</sup> and practical interest<sup>3,4</sup>. Theoretical calculations of the dispersion relations for propagation in both three-dimensional<sup>5,6,7</sup> and two-dimensional<sup>8,9</sup> dielectric structures using the plane-wave expansion have become relatively sophisticated. Experimentally, investigations of photonic band-structure phenomena have been confined to measurements at microwave frequencies because of difficulties in fabricating ordered dielectric arrays of optical length scales. Although microwave experiments have been used to determine the frequencies which define photonic band gaps<sup>10,11</sup>, and to perform elegant explorations of localized defect modes<sup>11,12</sup>, these techniques have not been used to measure the dispersion of radiation at frequencies away from the gaps. Measurement of the photonic dispersion relation over a broad frequency range would provide a direct test of the theoretical formalisms, and would also elucidate the transition in propagation from long wavelengths, where the dielectric array behaves as an effective medium, to shorter wavelengths where strong scattering leads to dispersion and the opening of photonic band gaps.

In this letter we investigate the dispersion relation for electromagnetic wave propagation in a periodic dielectric array using the recently developed coherent microwave

<sup>1</sup>G. Kurizki and A. Z. Genack, Phys. Rev. Lett. **61**, 2269 (1988). *J*

<sup>2</sup>S. John and J. Wang, Phys. Rev. Lett. **64**, 2418 (1990). *J*

<sup>3</sup>E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987). *J*

<sup>4</sup>R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, Phys. Rev. B **44**, 13772 (1991). *J*

<sup>5</sup>K. M. Leung and Y. F. Liu, Phys. Rev. Lett. **65**, 2646 (1990). *J*

<sup>6</sup>Ze Zhang and Sashi Satpathy, Phys. Rev. Lett. **65**, 2650 (1990). *J*

<sup>7</sup>K. M. Ho, C. T. Chan, and C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990). *J*

<sup>8</sup>R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos (to be published).

<sup>9</sup>M. Plihal, A. Shambrook, A. A. Maradudin, and P. Sheng, Opt. Commun. **80**, 199 (1991). *J*

<sup>10</sup>E. Yablonovitch, T. J. Gmitter, and K. M. Leung, Phys. Rev. Lett. **67**, 2295 (1991). *J*

<sup>11</sup>S. L. McCall, P. M. Platzman, R. Dalichaouch, D. Smith, and S. Schultz, Phys. Rev. Lett. **67**, 2017 (1991). *J*

<sup>12</sup>E. Yablonovitch, T. J. Gmitter, R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, Phys. Rev. Lett. **67**, 3380 (1991). *J*

transient spectroscopy (COMITS) technique<sup>13,14</sup>. COMITS is based on the radiation and detection of picosecond-duration electromagnetic transients with optoelectronically pulsed antennas. It is capable of free-space microwave and millimeter wave measurements over a broad frequency range (15-140 GHz) with good (60:1) polarization sensitivity<sup>15,16</sup>. In particular, the phase sensitivity of the COMITS technique is used to directly measure the dispersion relation in a two-dimensional photonic crystal across the fundamental and higher band gaps. The experimental results are compared to theoretical predictions obtained with the plane-wave expansion method, which has been described elsewhere<sup>7</sup>. In brief, the macroscopic Maxwell's equations are expressed as an eigenvalue equation in reciprocal space and the resulting matrix is then diagonalized to yield the mode frequencies and field patterns. This technique is a simple and powerful method to solve problems in electrodynamics which takes full account of the vector nature of the electromagnetic radiation.

Although there is much interest in three-dimensional photonic crystals, i.e., structures for which there is a photonic band gap in all propagation directions, in these experiments we chose to study a two-dimensional dielectric structure consisting of alumina-ceramic cylinders arranged in a square array. The primary reason for this choice is that the two-dimensional structures are easier to fabricate. However, the COMITS technique is capable of measuring the dispersion relation in three-dimensional systems as well. The results of reference 9 were used to determine the optimal choices for cylinder diameter and lattice spacing so that photonic band gaps appear in the experimentally accessible frequency range (15-140 GHz). For the measurements presented here,  $0.74 \pm 0.03$ -mm-diameter 100-mm-long alumina-ceramic rods were arranged with a 1.87-mm lattice constant in a

---

<sup>13</sup>Y. Pastol, G. Arjavalingam, J.-M. Halbout, and G. V. Kopcsay, Electron. Lett. **25**, 523 (1989).

<sup>14</sup>G. Arjavalingam, Y. Pastol, J.-M. Halbout, and G. V. Kopcsay, IEEE Trans. Microwave Theory Tech. **38**, 615 (1990).

<sup>15</sup>Y. Pastol, G. Arjavalingam, G. V. Kopcsay, and J.-M. Halbout, Appl. Phys. Lett. **55**, 2277 (1989).

<sup>16</sup>G. Arjavalingam, N. Theophilou, Y. Pastol, G. V. Kopcsay, and M. Angelopoulos, J. Chem. Phys. **93**, 6 (1990).

square array. Alumina was chosen because it has a large dielectric constant and low dielectric loss over the microwave and millimeter-wave spectrum<sup>13</sup>. The COMITS experimental setup is shown schematically in Fig. 1. The transmitting and receiving elements are exponentially tapered, coplanar strip antennas fabricated photolithographically on silicon-on-sapphire.<sup>14</sup> The silicon epilayer is subsequently ion implanted to reduce the carrier lifetime to less than 1 ps. Ultrafast optical pulses from a mode-locked, pulse-compressed, and frequency-doubled Nd-doped yttrium-lithium-fluoride laser are arranged in a conventional pump-probe configuration. The 527-nm wavelength pulses are ~1.5 ps wide and have a repetition rate of 240 MHz. A short current pulse is generated on the dc-biased transmitter by the pump beam. This pulse propagates down the coplanar strip line, spreading in time to about 7 ps, and is radiated into free space by the exponentially tapered flare. Hemispherical-fused silica lenses are used to collimate the transient radiation from the transmitter and to focus it, after passing through the sample, onto the receiver<sup>14</sup>. The voltage induced on the receiver is measured, by photoconductive sampling, as a function of the time delay between the pump and probe pulses<sup>14</sup>. Signal averaging is performed by adding ~1000 scans with a rapid-scan delay line.

To measure the transmission properties of a dielectric array, the sample is placed in the beam path so that the transient radiation propagates in a plane perpendicular to the axis of the cylinders. The cylinders forming the array were fixed in a machined holder so that the orientation of the sample, with respect to the direction of propagation, could be adjusted. Also, the number of rows in the array can be changed easily. The transverse dimensions of the array were always made larger than the collimated beam size (~3 cm) so that there were no end effects<sup>14</sup>. By suitably rotating the sample we could make measurements with polarization either parallel or perpendicular to the axis of the cylinders<sup>15,16</sup>. By symmetry, waves incident on the rods with polarization either parallel or perpendicular to the rod axis will preserve this polarization.

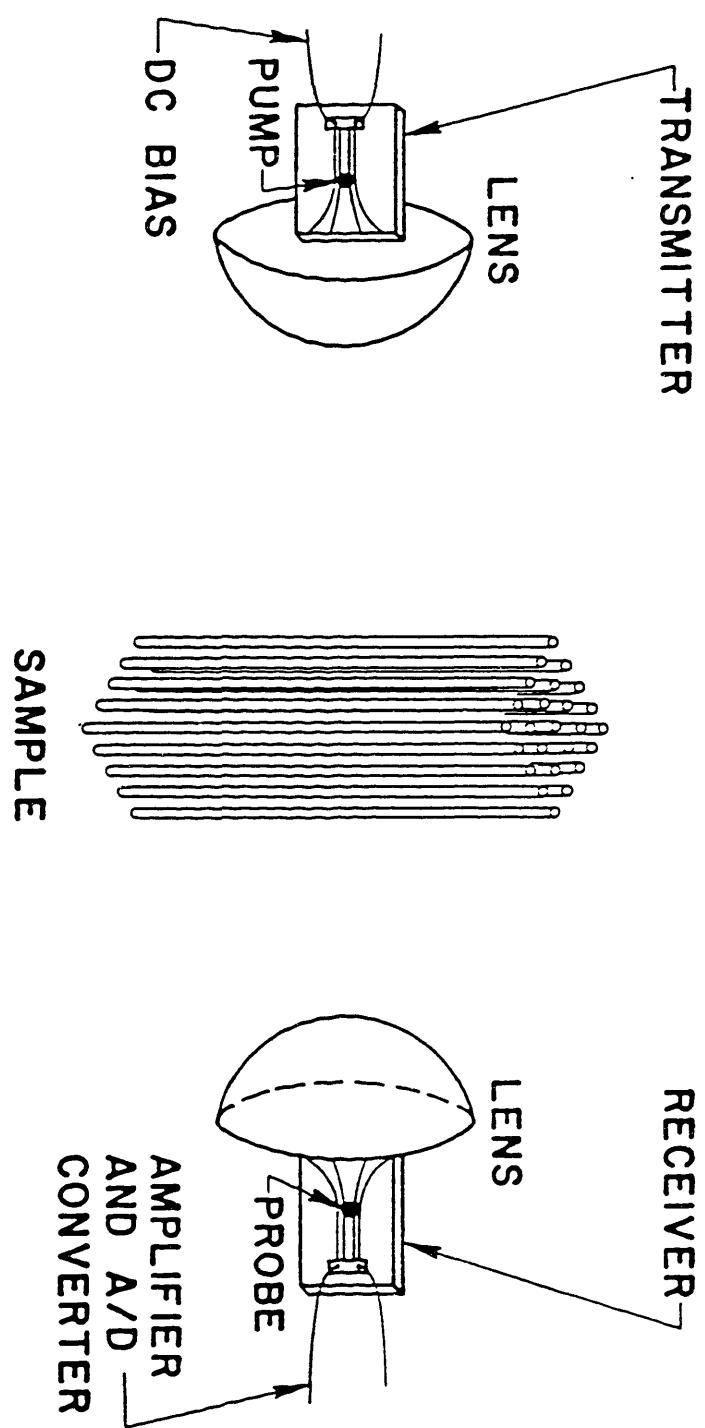


FIG. 1. Experimental setup for coherent microwave transient spectroscopy measurements. The two-dimensional photonic crystal is made of 0.74-mm-diam 100-mm-long alumina-ceramic rods arranged in a square array with a 1.87-mm lattice constant.

Time-domain wave forms are recorded first without and then with the sample in the beam path. The amplitude spectrum of the former, obtained by a numerical Fourier transform, is shown as the dashed lines in Fig. 2. Clearly, the pulse contains usable frequency components in the 15-140 GHz frequency range. Because the measured signal is proportional to the time-dependent voltage induced on the receiving antenna, phase information (not shown) is preserved. As we describe below, this phase sensitivity is crucial in determining the dispersion relation for electromagnetic wave propagation in the photonic crystal. The amplitude spectrum of the measured wave form, with the E field parallel to the axis of the ceramic rods is shown as the solid line in Fig. 2(a). The corresponding spectrum with E field perpendicular to the axis of the rods is depicted in Fig. 2(b). For both measurements the sample was 7 rows of rods deep in the direction of propagation and 25 columns wide transverse to it. Propagation was along the (10) axis of the lattice. Figure 2(a) clearly indicates the existence of a band gap between 45 and 70 GHz, and the suggestion of other gaps at about 100 and 125 GHz. Very narrow band gaps are not clearly resolved because our frequency resolution is limited to 5 GHz by the 200-ps window of the temporal data<sup>14</sup>. The results presented above are consistent with previous experimental measurements on two-dimensional photonic crystals<sup>11</sup>.

Although the amplitude spectrum alone gives an indication of the gaps, it is much more instructive to analyze the phase data to determine the effective refractive index of the photonic crystal and, thus, to determine the full dispersion relation. The complex transmission function of the dielectric array is calculated by dividing the complex Fourier transform of the wave form with the sample by the complex Fourier transform of the reference wave form without a sample. This transmission function contains both the amplitude and phase information, and represents the electromagnetic propagation properties of the photonic crystal. Using the known thickness of the dielectric array L and the net phase difference  $\phi$ , the effective microwave refractive index, n(f) can be calculated at each frequency<sup>14</sup>:

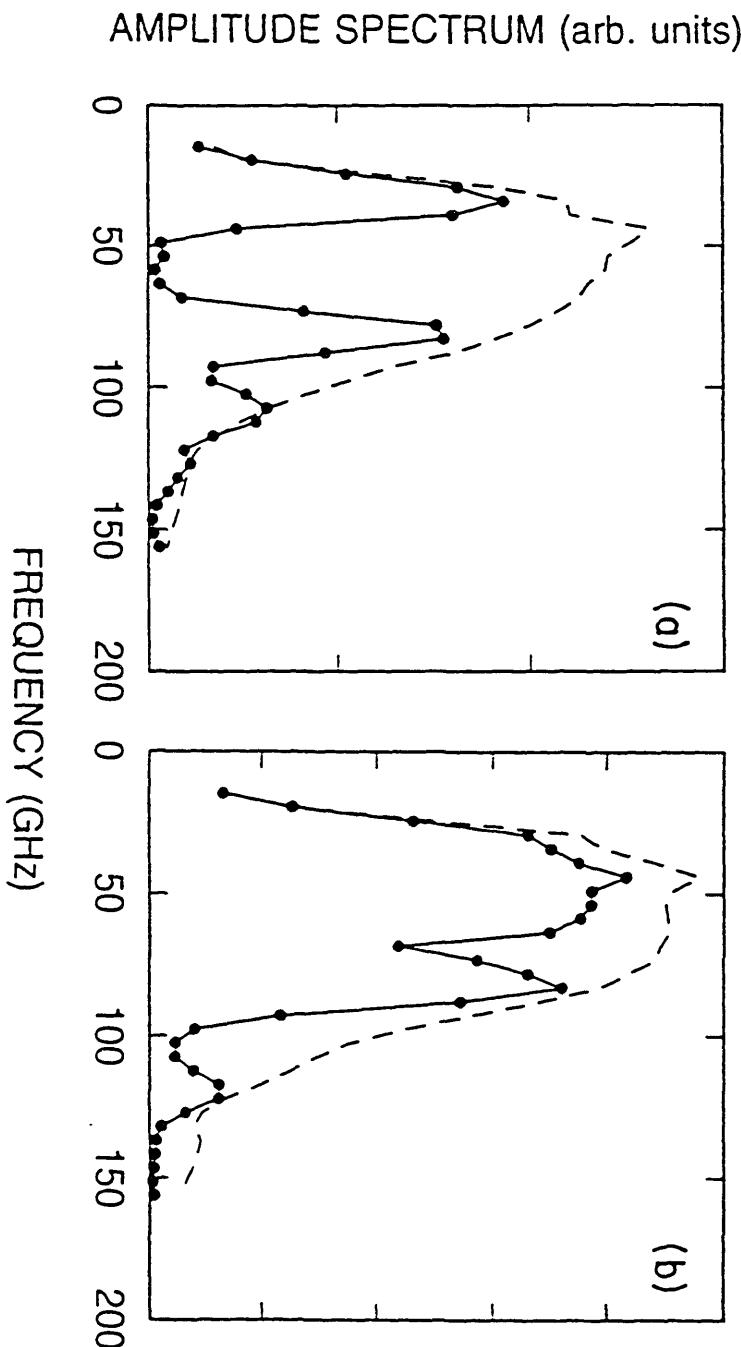


FIG. 2. Amplitude spectra obtained by a numerical Fourier transform of experimentally measured time-domain data. The dashed line represents the reference amplitude spectra obtained with no sample in the beam. Solid circles represent the amplitude spectra transmitted through seven rows of rods with propagation along the (10) direction with  $\mathbf{E}$  field of the transient radiation polarized (a) parallel and (b) perpendicular to the axis of the rods.

$$n(f) = (\phi c / 2\pi L f) + 1, \quad (1)$$

where  $c$  is the velocity of light. Using the effective index values, the dispersion relation,  $f$  vs.  $k$ , can be calculated directly with

$$k = (2\pi f/c) n(f). \quad (2)$$

The measured dispersion relation for propagation along the (10) direction with the  $\mathbf{E}$  field polarized parallel to the rods is plotted as the solid circles in Fig. 3(a). Overlaid on the measured data are the theoretical predictions (dashed lines) calculated using the plane-wave expansion technique. The agreement between the measured data and the theoretical calculation is generally excellent. In Fig. 3(b) we show the dispersion relation for propagation along the (10) direction but with the polarization rotated such that the  $\mathbf{E}$  field is perpendicular to the axis of the cylinders. The solid squares in the figure are the experimentally measured values and the dashed lines predictions of theory. Again, the agreement between theory and experiment is excellent. For both polarizations photonic band gaps appear, as expected, at the Brillouin zone boundaries. Although the theory predicts that small gaps open at higher frequencies, as mentioned above, the resolution of the COMITS technique is insufficient to fully resolve them. However, in the amplitude spectra of Fig. 2 there are dips at these points reflecting a strong suppression in the density of states. It should be noted in Fig. 3(b) that there is a large jump in wave vector (i.e., a large phase change) between band 2 and band 3 resulting in an ambiguity in the correct value for the phase, and hence in the correct form of the dispersion relation. This ambiguity can generally be resolved by determining the transmission function due to a single row of rods, accomplished by comparing the transmission through two photonic crystals whose thicknesses differ by one row of rods.

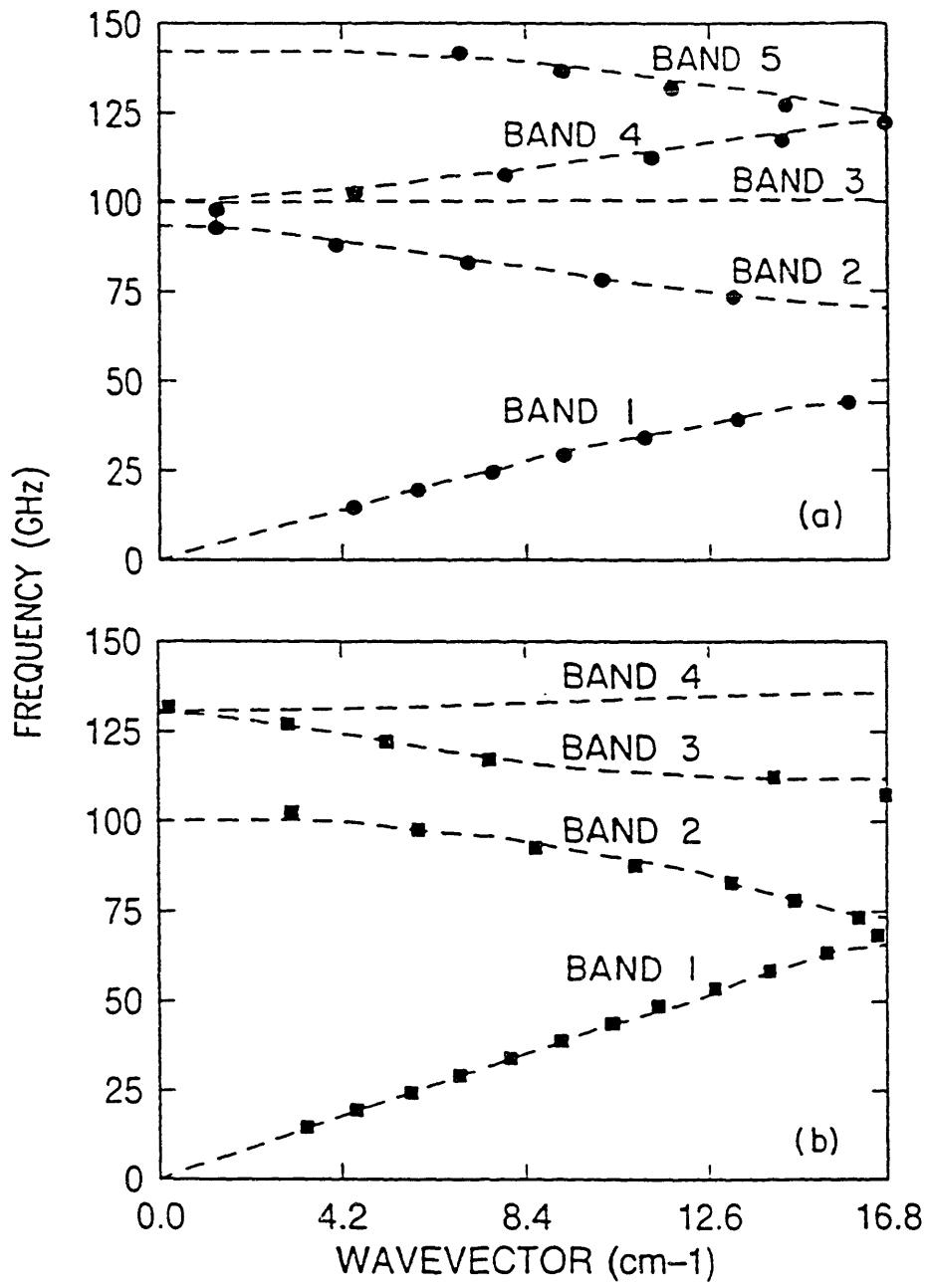


FIG. 3. Dispersion relation for propagation of electromagnetic waves along the (10) direction of a two-dimensional photonic crystal with polarization (a) parallel and (b) perpendicular to the rod axis. The solid symbols are the experimentally determined values and the dashed lines are the predictions of theory.

The theoretical calculations also predict the existence of band which is not seen experimentally, band 3 in Fig. 3(a). Considering the excellent agreement between theory and experiment for the other modes, this result is rather surprising. However, an examination of the field patterns for these bands reveals that a planar wave front (as in our experiment) cannot excite these modes. The reason is as follows. In our experiment we measure the band structure along the (10) direction, which is a special direction of the Brillouin zone. Physically, this means that the fields must be either even or odd upon reflection through the mirror plane shown in Fig. 4. As this figure shows, the fields for band 3 are *odd* under reflection through this mirror plane. However, an incoming plane wave with electric field oriented along the rod axis is *even* with respect to this reflection, and so it cannot excite the mode of band 3. In fact, all of the modes which are seen experimentally are even under this reflection, and all the modes which are predicted but not found are odd. Symmetry plays an analogous role for radiation of the other polarization. Our calculations show that the lowest three bands in Fig. 3(b) are symmetry allowed and we find experimentally that all three are observed. Theoretically, we predict that band 4 of Fig. 3(b) is symmetry forbidden and we do not observe this band in our experiment. It is possible that we do not observe this band because of the high-frequency limitations of our experiment. Thus, we believe that the missing modes in Fig. 3 do exist in the crystal, but that they do not transmit electromagnetic radiation in this specific experiment.

At long wavelengths, i.e., for frequencies well below the fundamental gap, the dielectric response of the two-dimensional array is well described by effective-medium theory. The orthogonal orientations of the polarization with respect to the two-dimensional array of cylinders correspond precisely to the configurations considered in effective-medium theory for the limiting cases of maximum ( $\mathbf{E}$  field parallel to all interfaces) and

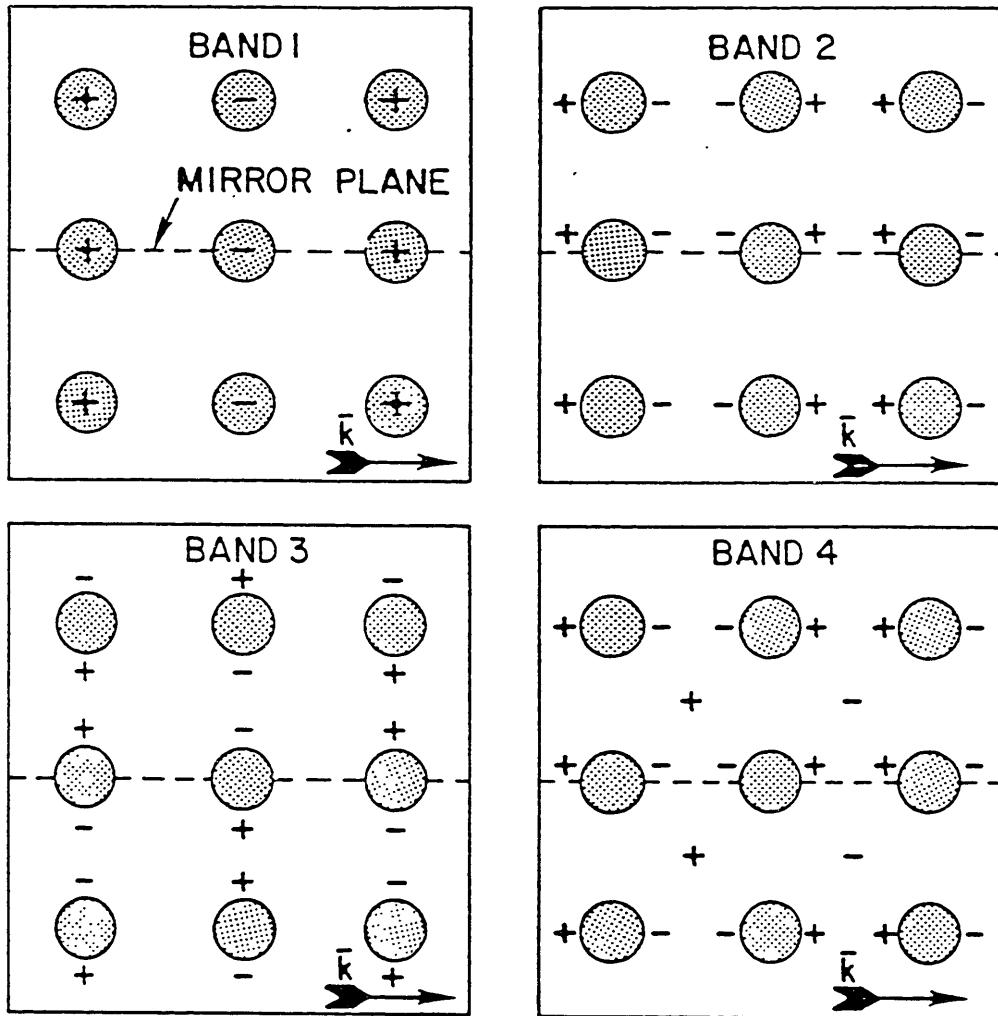


FIG. 4. The symmetry of the electric fields associated with the lowest four photonic bands. For these modes, the electric-field lines run parallel to the rod axis, and so the +(−) signs indicate regions in which the electric field is oriented into (out of) the page. The states depicted lie at the Brillouin zone edge and so the fields alternate sign in neighboring lateral unit cells. The shaded circles (not shown to scale) indicate the square lattice of dielectric rods. The modes along the (10) direction must be even (bands 1,2, and 4) or odd (band 3) with respect to reflection in the mirror plane, shown as the dashed line.

minimum ( $\mathbf{E}$  field perpendicular to all interfaces) screening<sup>17</sup>. The dielectric constants for these two situations are given by

$$\epsilon = f_1\epsilon_1 + f_2\epsilon_2 \quad (3)$$

$$1/\epsilon = f_1/\epsilon_1 + f_2/\epsilon_2, \quad (4)$$

where  $f_1$  and  $f_2$  are the volume filling fractions occupied by the two media with dielectric constants  $\epsilon_1$  and  $\epsilon_2$ , respectively. For our two-dimensional array of alumina cylinders ( $\epsilon_1 = 8.9$ ) these expressions predict microwave refractive indices of 1.41 and 1.06 for  $\mathbf{E}$  field parallel and perpendicular to the cylinders respectively. Experimentally the lowest-frequency (long-wavelength) refractive index values are 1.47 and 1.11.

In summary, we have investigated the photonic band structure of a two-dimensional dielectric array using the coherent microwave transient spectroscopy technique. The experimentally determined dispersion relations for electromagnetic wave propagation, for  $\mathbf{E}$  field polarization both parallel and perpendicular to the rod axis, were in good agreement with theoretical predictions calculated using the plane-wave expansion technique. Certain modes which were predicted experimentally were not observed because coupling between these modes and the incident plane-wave radiation was forbidden by symmetry. Although results for propagation along the (10) direction were presented, similar results were also observed for propagation along the (11) direction. Furthermore, the COMITS technique can also be used to explore the photonic band structure of three-dimensional photonic crystals.

The authors thank F. Stern and J. Tsang of the IBM Research Division for their critical reading of the manuscript.

---

<sup>17</sup>D. E. Aspnes, Am. J. Phys. **50**, 704 (1982). <sup>15</sup>

**Chapter 9: Electromagnetic Bloch Waves at the Surface of a Photonic  
Crystal**

in collaboration with

**R. D. Meade and K. D. Brommer**

The recent discovery<sup>1,2</sup> that periodic structures of dielectric materials (photonic crystals) can be made to have an optical band gap in which propagating electromagnetic modes are forbidden has been greeted with enthusiasm. These materials possess many unique physical properties. For instance, defects in photonic crystals have been predicted<sup>3</sup> and observed<sup>4</sup> to localize electromagnetic modes. Moreover, the absence of radiating modes will inhibit the decay of excited states whose energies lie in the band gap, which may lead to unusual electronic<sup>2</sup> and chemical properties of atoms or molecules embedded in the photonic crystal. The absence of zero-point fluctuations in the photonic band gap also has physical consequences, such as an anomalous Lamb shift.<sup>5</sup> In addition, interest in these materials has also been generated by their potential applications. In fact, the prediction that photonic crystals would lead to efficient semiconducting lasers has been a strong motivation for their development.<sup>6</sup>

In previous work, authors have considered properties of bands,<sup>1,2,7,8,9</sup> defects,<sup>3</sup> and atoms<sup>5</sup> in *infinite* photonic crystals. In this paper, we will determine the effect of terminating the crystal by examining the surface band structure of an interface between air and a photonic crystal. We will show that photonic crystals have electromagnetic surface modes, in which light is exponentially localized to the surface plane. However, unlike the familiar case of metal surfaces,<sup>10</sup> in which the localization of electromagnetic surface waves is a result of a negative dielectric constant in the metal, surface Bloch waves are localized because of interference effects in the photonic crystal. Thus these surface waves are

---

<sup>1</sup>K. M. Ho, C. T. Chan, and C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990). *J*

<sup>2</sup>E. Yablonovitch, T. J. Gmitter, and K. M. Leung, Phys. Rev. Lett. **67**, 2295 (1991). *J*

<sup>3</sup>R. D. Meade, K. D. **Brommer**, A. M. Rappe, and J. D. Joannopoulos, Phys. Rev. B **44**, 13772 (1991). *J*

<sup>4</sup>E. Yablonovitch, T. J. Gmitter, R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, Phys. Rev. Lett. **67**, 3380 (1991). *J*

<sup>5</sup>S. John and J. Wang, Phys. Rev. Lett. **64**, 2418 (1990). *J*

<sup>6</sup>E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987). *J*

<sup>7</sup>Ze Zhang and Sashi Satpathy, Phys. Rev. Lett. **65**, 2650 (1990). *J*

<sup>8</sup>K. M. Leung and Y. F. Liu, Phys. Rev. Lett. **65**, 2646 (1990). *J*

<sup>9</sup>E. Yablonovitch and T. J. Gmitter, Phys. Rev. Lett. **63**, 1950 (1989). *J*

<sup>10</sup>R. F. Wallis and G. I. Stegeman, *Electromagnetic Surface Excitations*, (Springer-Verlag, Berlin, 1986), p. 2. *J*

analogous to the electromagnetic Bloch waves occurring on the surface of multilayer films.<sup>11</sup> We will also discuss the impact of these results on the design of semiconducting lasers. Understanding the surface band structure is of particular importance to these devices, because by choosing a proper termination of the photonic crystal one can eliminate the losses associated with radiation into surface modes.

In order to study the surface of a photonic crystal we have performed a series of calculations of interface states. We have found that the surface states can be cataloged as one of four types: extended both in the crystal and the air (EE), decaying in the air but extended in the dielectric (DE), extended in the air and decaying in the dielectric (ED), and decaying in both the dielectric and the air (DD). In the final case, which we will discuss in detail, the light is exponentially localized at the surface. We will find that the surface electromagnetic states are of two general types, having primarily transverse electric (TE) or transverse magnetic (TM) fields. We will also present general arguments that surface modes must always exist for some termination of any surface of a photonic crystal.

It has been noted by several authors<sup>1,9</sup> that there is a similarity between electronic states in a periodic potential and electromagnetic states in a periodic dielectric. We can take advantage of this similarity by applying the methods developed for electronic-structure calculations, such as plane-wave expansions, in order to solve for the eigenstates of an electromagnetic system. These methods, which have been described in detail by a number of authors,<sup>1,7,8</sup> provide a simple and powerful way to solve problems in electrodynamics which takes full account of the vector nature of the electromagnetic radiation.

Although we would like to study the surface of a semi-infinite photonic crystal, computationally it is desirable to choose a system with a finite unit-cell size. Therefore, we employ the supercell method in which slabs of dielectric material alternate with slabs of vacuum. We employ a supercell with six double layers of photonic crystal and six double

---

<sup>11</sup>P. Yeh, *Optical Waves in Layered Media* (Wiley, New York, 1988), p.337. 

layers of vacuum. Because the surface states are strongly localized to the uppermost dielectric layer, the error introduced by the finite size of the slab is negligible, less than 0.1% as judged by the splitting between states on opposite surfaces. We expand in plane waves up to a finite frequency, including  $\sim 130$  plane waves per polarization, per primitive unit cell. We found that this was a sufficient plane-wave cutoff to describe the surface states accurately, and increasing the plane-wave cutoff to 250 plane waves per polarization, per primitive unit cell changed the frequencies by less than 0.1%.

Because it is the structure most likely to be employed in the production of a semiconductor laser system, we have chosen to study the (111) surface of the photonic crystal proposed by Yablonovitch, Gmitter, and Leung.<sup>2</sup> This structure is described in detail in Ref. 2. Qualitatively, this structure can be thought of as a set of dielectric columns connecting the sites of a diamond lattice (in the same way that the bonds connect atoms in a silicon crystal). These columns, then, point along the [111], [ $\bar{1}\bar{1}1$ ], [ $\bar{1}1\bar{1}$ ], and [1 $\bar{1}\bar{1}$ ] directions. The three columns along the [ $\bar{1}\bar{1}1$ ], [ $\bar{1}1\bar{1}$ ], and [1 $\bar{1}\bar{1}$ ] directions are identical, but the fourth column in the [111] direction has a larger diameter. A (110) cross-section of this surface is shown in the inset to Fig. 1. The columns along the [111] direction point vertically in this picture, and the diagonal rib connecting these [111] columns is a [ $\bar{1}\bar{1}1$ ] column. This surface contains a  $C_{3v}$  symmetry, with the rotational axis normal to the surface and passing through the center of the [111] column. Although we have chosen to look at the (111) surface of the Yablonovitch structure, we expect that many of the features of the electromagnetic surface states will be common to all photonic crystals.

Before we consider the band structure of the interface, let us first consider the projected band structure of the air and photonic crystals separately. The union of regions EE and ED shown in Fig. 1 is the projection of the free photon modes into the surface Brillouin zone. For a given  $\mathbf{k} = (k_x, k_y)$ , there are light modes at all frequencies  $\omega \geq ck$ . Along the line  $\omega = ck$ , the light travels parallel to the surface, and increasing  $\omega$  represents an increasing component of  $k_z$ . Similarly, the union of regions EE and DE represents the

## Electromagnetic Surface Band Structure

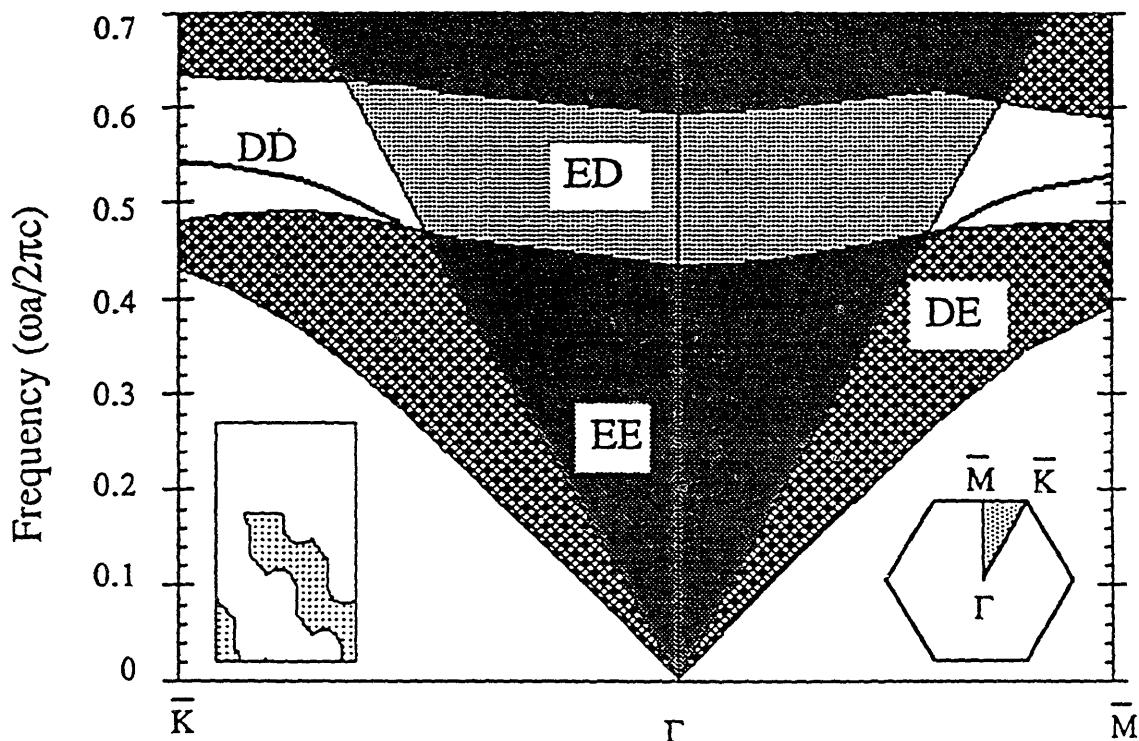


FIG. 1. The band structure of the (111) surface of the Yablonovitch crystal along special directions in the surface Brillouin zone. The shading denotes regions in which light is transmitted (EE), internally reflected (DE), and externally reflected (ED). The lines in the gap correspond to surface bands in which light is exponentially localized to the surface (DD). The surface Brillouin zone is shown in the lower right, with the irreducible BZ shaded. The  $\bar{M}$  point is directed along  $(1, \frac{2}{3}, 1)$  a distance of  $\sqrt{2}/3 (2\pi/a)$  from  $\bar{\Gamma}$ . The  $\bar{K}$  point is directed along  $(0, 1, 1)$ , a distance of  $(\sqrt{8}/3)(2\pi/a)$  from  $\bar{\Gamma}$ . This surface band structure corresponds to a termination of  $\tau = 0.75$ , shown in the inset with air above and photonic crystal below (dielectric regions are shaded, and air regions are unshaded). Frequencies are in units of  $(2\pi c/a)$ , where  $a$  is the conventional lattice constant.

projected band structure of the photonic crystal. Note that the photonic crystal contains a gap  $0.49 < \omega a/(2\pi c) < 0.59$  in which there are no allowed states.

It is now straightforward to understand the four types of surface states, transmitted (EE), internally reflected (DE), externally reflected (ED), and surface modes (DD). In the region of  $(k, \omega)$  marked EE, the modes are extended in both the air and in the dielectric and so it is possible to have light transmitted through the crystal. In the DE region, there are modes in the dielectric, but they are beneath the band edge for the air states. Thus the light is extended in the dielectric but exponentially decaying into the vacuum, which constitutes total internal reflection. In the ED region, the situation is reversed; and these states are extended states in the air but they are in the gap of the dielectric. In this case, the modes are extended in the vacuum region but exponentially decaying into the dielectric and so incident light is reflected. Finally, we find there can be states in the region marked DD. In this regime, the states are below the band edge of the light in the air, as well as in the gap of the dielectric material. Thus the light is exponentially decaying in both directions, and so it is localized to the surface plane.

The fields associated with the zone edge ( $\bar{M}$ ) surface mode are shown in Fig. 2(a), which is a (110) cross section through the surface plane. Because the crystal contains a mirror symmetry through this plane, the **D** field lies in the plane, while the **H** field is everywhere normal to the plane, and is displayed in contours. Note that the fields are strongly localized in the plane of the surface, with most of their character in the uppermost dielectric layer. As is the case for the bulk modes and defect states,<sup>3,4</sup> most of the power of the **D** field is concentrated in the dielectric regions. Fig. 2(b) displays the fields associated with the  $\bar{M}$  surface mode in a (111) plane passing through the top layer of the surface. Although this is not a mirror plane, the **H** field is primarily in the plane, and is shown as vectors, and the **D** field is primarily normal to the plane, and is shown in contours.

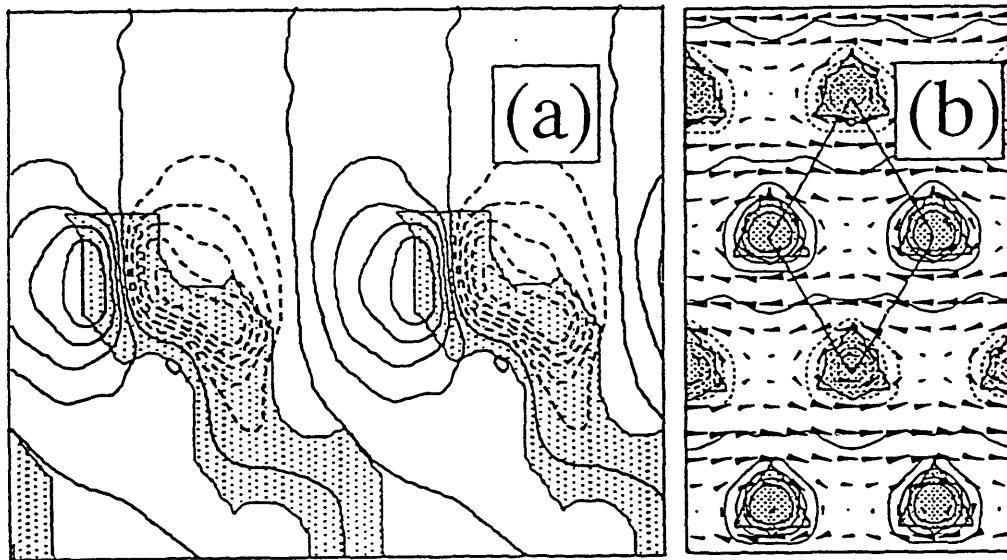


FIG. 2. The fields associated with the zone-edge ( $\tilde{M}$ ) surface mode for a surface termination of  $\tau = 0.75$ . The regions of dielectric are shown shaded, while the regions of air are left unshaded. (a) The fields shown in the  $(110)$  plane. The contours indicate lines of constant  $H$ . (Dashed lines indicate contours of negative value). Because  $D$  is the curl of  $H$ ,  $D$  is always parallel to the equipotential lines of  $H$ , and is large only where the gradient of  $H$  is large (where the equipotential lines are closely spaced). In this cross section,  $k$  lies in the plane of the figure. (b) The fields associated with the  $\tilde{M}$  surface mode, in a plane passing through the top layer of the surface. The  $H$  field is primarily in the plane, and is shown as vectors, and the  $D$  field is primarily normal to the plane, and is shown as contours. The surface unit cell is circumscribed by the parallelogram, and the fields change sign in alternate unit cells, consistent with a zone-edge state. In this figure,  $k$  is directed vertically.

One interesting feature of the fields shown in Fig. 2(b) is that the magnetic field is primarily in the transverse direction, perpendicular to the surface and to the  $\mathbf{k}$  vector. For this reason, we describe the mode as TM-like. However, the surface modes of a photonic crystal are not rigorously TE or TM. To understand why this is so, let us consider the simpler case of the surface of a multilayer film. In this case the distinction between modes which are TE and TM is exact. This can be seen from symmetry arguments. The surface of a multilayer film has a continuous translational symmetry parallel to the surface plane, so we can label the states by in-plane  $\mathbf{k}$  vector. Consider the mirror plane defined by the  $\mathbf{k}$  vector and the surface normal. Fields must be either even or odd with respect to this mirror plane. Fields which are odd are normal to the plane, while fields that are even lie in the mirror plane. If the  $\mathbf{H}(\mathbf{D})$  field is normal to the plane, than the mode is TM(TE). However, at the surface of a photonic crystal, the symmetry is lower and so there is no rigorous distinction between TE and TM modes. Even though the rigorous distinction between TE and TM breaks down at the surface of a photonic crystal, the fields displayed in Fig. 2(b) lead us to describe the mode qualitatively as TM-like.

Until this point, we have considered the band structure of the one particular termination of the (111) surface, shown in the inset to Fig. 1. However, the surface can be terminated at many levels [see the inset to Figs. 3(a) and (b)]. By specifying the Miller indices of a surface we determine the surface normal, and for periodic arrays of atoms this uniquely describes the physical system whose surface properties we wish to investigate. However, a photonic crystal has a lattice constant of macroscopic dimensions, and so we need a nomenclature to determine not just the surface normal, but also the termination of the surface inside a unit cell. To do so, let us introduce a termination parameter  $\tau$  ( $0 \leq \tau < 1$ ), which describes the termination of any surface. Let  $\tau = 0$  when the surface is terminated through the bond center, as shown in the inset to Fig. 3(b), and increase linearly as the height of the surface is raised, normalized so that the surface terminates at  $\tau = 1$  when the same physical surface as  $\tau = 0$  is achieved.

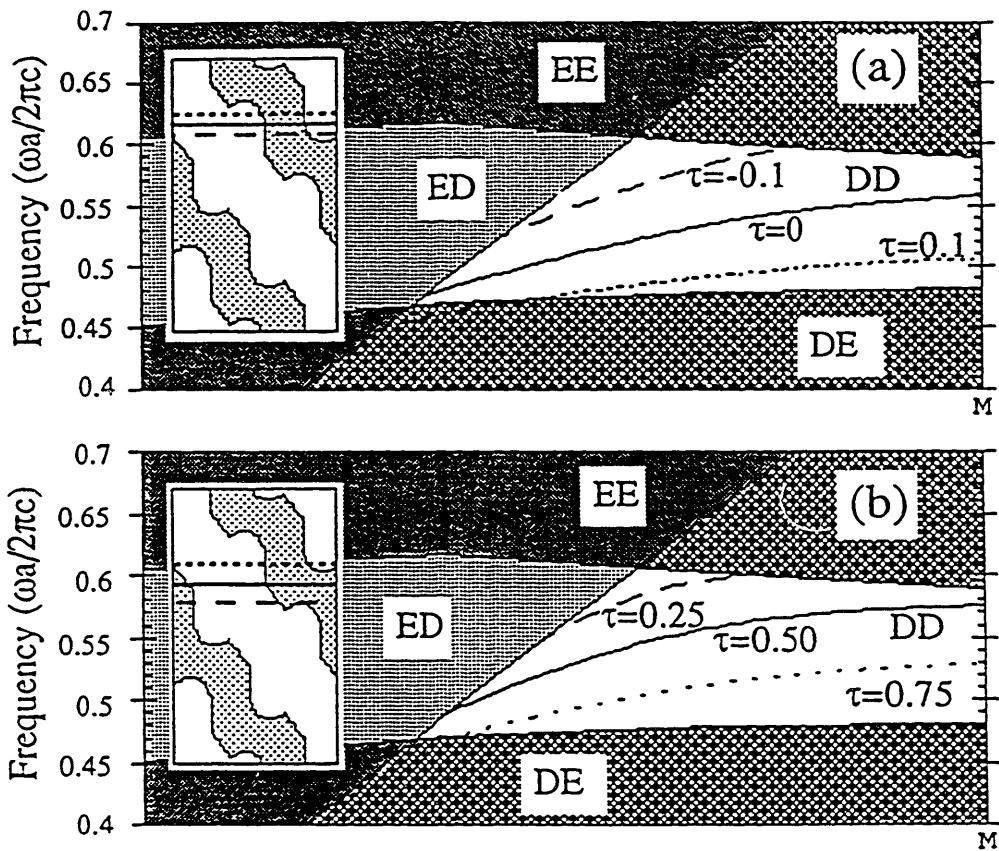


FIG. 3. Surface band structure of (a) TM-like and (b) TE-like modes. This figure shows the gap region of Fig. 1 on an expanded scale, with the same labeling of the four regions EE, ED, DE, DD. The insets indicate the surface terminations corresponding to the three values of  $\tau$  for which the surface bands have been plotted (see Fig. 1, inset). Larger values of  $\tau$  correspond to vertically higher terminations on the inset, which contain more dielectric on the surface. Note that higher  $\tau$  yields lower values of surface band frequency.

In fact, we find that the surface band structure varies in an interesting manner as the termination of the surface is changed, as is shown in Figs. 3(a) and 3(b). As the termination of the surface is increased, more dielectric is added, and the frequency of the surface mode is lowered. This can be understood in an intuitive way, recalling that the dielectric is analogous to a region of low potential. As more dielectric is added to the surface, the potential which the surface state experiences decreases, and so its frequency decreases. Thus, as  $\tau$  is increased, bands sweep down from the conduction band to the valence band. These states may be either TM-like, as shown in Fig. 2(a), or TE-like with the  $\mathbf{D}$  field primarily in the transverse direction. Let us consider the band structure as  $\tau$  is increased from  $\tau = 0$  to  $\tau = 1$ . The surface bands in the final termination are identical to the surface bands in the initial termination, since the two correspond to equivalent surfaces. As the termination is increased, exactly two states are swept from the conduction band to the valence band, one of which is TM-like and the other is TE-like. Thus, by increasing the termination of the surface from  $\tau = 0$  to  $\tau = 1$  we have added one bulk unit cell (per surface unit cell), and increased the total number of states in the valence band by two (at each  $\mathbf{k}$  point).

This suggests a general argument, that any  $\{nlm\}$  surface (with  $n,l,m$  integers) of any photonic crystal with a gap will always have surface states for some termination. If the surface has integer indices, then it has a finite unit cell size, and a finite surface Brillouin zone. Since the crystal as a whole has a gap, the surface Brillouin zone must also have a gap. Suppose that as the surface termination is varied between any  $\tau$  and  $\tau + 1$ , there are  $b$  bulk unit cells introduced per surface unit cell. There must then be  $2b$  states transferred from the conduction band to valence band. As the frequencies of these states are lowered from the bottom of the conduction band to the top of the valence band, they must sweep through the gap and become surface bands. This argument for the existence of surface states also applies to the case of multilayer films which lack a gap in their full three-dimensional Brillouin zone, but have a gap in the direction normal to the surface.

Finally, we discuss the importance of these results on the design of semiconducting lasers. The possibility of employing photonic crystals to produce efficient semiconducting lasers was first suggested by Yablonovitch.<sup>2</sup> His proposal has several elements, which we will review briefly. As a first ingredient, one uses microfabrication techniques to produce a photonic band-gap structure from a semiconducting material. By choosing the proper structure, one insures that the electronic band-gap energy  $E_{\text{gap}}$  falls within the photonic band gap. The next ingredient is to create a lasing mode at  $\omega_d$  in the electromagnetic band gap. This can be done by introducing a defect into the photonic crystal which creates a localized electromagnetic mode. If the defect size is chosen so that  $h\omega_d = 2\pi E_{\text{gap}}$ , then the system can radiate into the local mode, which acts as a laser cavity. Because there is only a single mode in the gap, there is no possibility for spontaneous emission into other frequencies, which enhances the efficiency of this laser system.

The final step is to extract the light from the local mode. Even if the defect is located several lattice constants below the surface, there will be some overlap between the exponential tail of the surface states, penetrating into the crystal, and the exponential tail of the defect state embedded within the crystal. Because of this overlap, the light confined in the defect mode can escape by tunneling. But, depending on the surface, there are two channels into which the light can escape, into the air states (ED) or into the surface states (DD). Only the radiation which escapes into air is useful. The radiation which is confined to the surface cannot be collected, and so this is a mechanism for loss. Fortunately, by choosing a proper termination for the photonic crystal, one can insure that there are no surface states at the lasing frequency. For instance, at a termination of  $\tau = 0.75$  (see Fig. 1), there are no surface states in the region  $0.54 < \omega_a/(2\pi c) < 0.59$ , and so this range would be the optimal choice for the lasing frequency.

In conclusion, electromagnetic states can be localized at the surface of a photonic crystal. We have found that these modes are of two general types, TE-like and TM-like,

and have presented general arguments that surface modes must always exist for some termination of any surface of a photonic crystal.

We wish to acknowledge E. Yablonovitch for helpful conversations regarding the semiconducting laser system, and for pointing out the similarity between surface states in photonic crystals and multilayer films. Partial support for this work was provided by the Office of Naval Research Contract No. N00014-86-K-0158. Finally, one of us (A. M. R.) would like to acknowledge the support of the Joint Services Electronics Program.