Photonic bands: Convergence problems with the plane-wave method

H. S. Sözüer* and J. W. Haus†

Department of Physics, Rensselaer Polytechnic Institute, Troy, New York 12180

R. Inguva*

NASA Marshall Space Flight Center, EB23, Optical Systems Branch, Huntsville, Alabama 35812

(Received 4 February 1991; revised manuscript received 7 October 1991)

The problems associated with the poor convergence of the Fourier transform of the hard-sphere dielectric function are discussed. A significant band gap between the eighth and ninth levels has been found for air spheres in fcc. We also consider a periodic array of Gaussian spheres, which converges well and allows a consistent and reliable determination of the general features of photonic bands. It is found that when $\epsilon(\mathbf{r})$ is sharply peaked, photonic levels become almost degenerate throughout the Brillouin zone, corresponding to standing waves.

I. INTRODUCTION

The idea of photonic bands in periodic dielectric structures was suggested[1,2] some time ago. It has also been investigated experimentally[3] and theoretically for scalar waves,[4-6] and more recently for electromagnetic (EM) waves[7-9] using the plane-wave method. Some of the consequences of forbidden band gaps in such structures, such as the modification of molecular interactions and of the quantum electrodynamics of an atom, have also been discussed.[10] Band gaps were shown to exist for scalar waves for the face-centered-cubic (fcc) lattice; however, for EM waves no gap was found for spherical atoms in the fcc lattice for the lower-lying frequencies.

In this paper we will formulate the problem in a slightly different fashion which we believe exposes its overall features more effectively and focus on those features of photonic levels that depend on the gross features of the dielectric function $\epsilon(\mathbf{r})$. We start with the wave equations in a lossless periodic dielectric structure, i.e. $\mu(\mathbf{r}) = 1$, $\sigma(\mathbf{r}) = 0$,

$$\nabla^2 \psi(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \epsilon(\mathbf{r}) \psi(\mathbf{r}, t) = 0 \,,$$

$$\nabla \times \nabla \times \mathcal{E}(\mathbf{r}, t) + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \sum_{i,j=1}^{3} \epsilon^{ij}(\mathbf{r}) \mathcal{E}^j(\mathbf{r}, t) \hat{\mathbf{e}}^i = 0 \,. \tag{1}$$

For the EM wave equation $\mathcal{E}$ must also satisfy the auxiliary condition $\nabla \cdot \mathcal{D} = 0$. It is worth noting that for $\epsilon(\mathbf{r}) = 1$ and $\mu(\mathbf{r})$ periodic, one gets the identical eigenvalue problem for $\mathcal{H}$ including the auxiliary condition $\nabla \cdot \mathcal{B} = 0$. Therefore *all* the properties of periodic dielectric structures to be discussed in this paper also hold for periodic paramagnetic structures.

Here we consider a linear, isotropic, and positive-definite dielectric medium periodic in some lattice with lattice vectors $\mathbf{R}$,

$$\epsilon^{ij}(\mathbf{r}) = \delta^{ij} \epsilon(\mathbf{r}) \,,$$

$$\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} + \mathbf{R}) > 0 \,. \tag{2}$$

It is also convenient to assume

$$\epsilon(\mathbf{r}) = \epsilon_b + \sum_{\mathbf{R}} \epsilon_0(\mathbf{r} - \mathbf{R}) \,, \tag{3}$$

where ϵ_b is the dielectric constant of the background and $\epsilon_0(\mathbf{r} - \mathbf{R})$ is the contribution to the dielectric function at $\mathbf{r}$ from the "atom" at $\mathbf{R}$. We assume that $\epsilon_0(\mathbf{r})$ is square integrable and not necessarily confined to a single primitive cell. Its Fourier transform is defined by

$$\epsilon(\mathbf{q}) = \frac{1}{(2\pi)^3} \int d\mathbf{r} \, e^{-i\mathbf{q}\cdot\mathbf{r}} \epsilon(\mathbf{r}) = \sum_{\mathbf{G}} \delta(\mathbf{q} - \mathbf{G}) \, \epsilon(\mathbf{G}) \,, \tag{4}$$

where the second equality follows from (2) and

$$\epsilon(\mathbf{G}) = \frac{1}{V_{\text{cell}}} \int_{\text{WS cell}} d\mathbf{r} \, e^{-i\mathbf{G}\cdot\mathbf{r}} \epsilon(\mathbf{r}) = \frac{1}{V_{\text{cell}}} \int_{\text{WS cell}} d\mathbf{r} \, e^{-i\mathbf{G}\cdot\mathbf{r}} \left[\epsilon_b + \sum_{\mathbf{R}} \epsilon_0(\mathbf{r} - \mathbf{R}) \right]$$

$$= \epsilon_b \delta_{\mathbf{G}0} + \frac{1}{V_{\text{cell}}} \sum_{\mathbf{R}} \int_{\text{WS cell}} d\mathbf{r} \, e^{-i\mathbf{G}\cdot(\mathbf{r}-\mathbf{R})} \epsilon_0(\mathbf{r} - \mathbf{R})$$

$$= \epsilon_b \delta_{\mathbf{G}0} + \frac{1}{V_{\text{cell}}} \int_{\text{all } \mathbf{r}} d\mathbf{r} \, e^{-i\mathbf{G}\cdot\mathbf{r}} \epsilon_0(\mathbf{r})$$

$$= \epsilon_b \delta_{\mathbf{G}0} + \epsilon_0(\mathbf{G}) \,, \tag{5}$$

$$\epsilon_0(\mathbf{G}) \equiv \frac{1}{V_{\text{cell}}} \int_{\text{all } \mathbf{r}} d\mathbf{r}\, e^{-i\mathbf{G}\cdot\mathbf{r}}\, \epsilon_0(\mathbf{r})\,. \qquad (6)$$

V_{cell} is the volume of the primitive cell of the lattice, which we take to be the Wigner-Seitz (WS) cell for convenience and $\mathbf{G}$ is a reciprocal-lattice vector. For a lattice with n basis vectors $\mathbf{b}_j$, $j = 1, \ldots, n$ with identical "atoms" at each site,

$$\epsilon(\mathbf{r}) = \epsilon_b + \sum_{\mathbf{R}} \sum_{j=1}^{n} \epsilon_0(\mathbf{r} - \mathbf{b}_j - \mathbf{R})\,, \qquad (7)$$

$$\epsilon(\mathbf{G}) = \epsilon_b \delta_{\mathbf{G}0} + \left[\sum_{j=1}^{n} e^{-i\mathbf{G}\cdot\mathbf{b}_j}\right] \epsilon_0(\mathbf{G})\,. \qquad (8)$$

For $n = 2$, a lattice with a two-point basis such as the diamond structure, the basis vectors can always be chosen equal and opposite, $\mathbf{b}_{1,2} = \pm\mathbf{b}$. Hence

$$\epsilon(\mathbf{G}) = \epsilon_b \delta_{\mathbf{G}0} + 2\cos(\mathbf{G}\cdot\mathbf{b})\,\epsilon_0(\mathbf{G})\,.$$

Inverting $\epsilon(\mathbf{G})$, one gets

$$\epsilon(\mathbf{r}) = \sum_{\mathbf{G}} \epsilon(\mathbf{G})\, e^{i\mathbf{G}\cdot\mathbf{r}}\,. \qquad (9)$$

If $\epsilon_0(\mathbf{r})$ is spherically symmetric, then

$$\epsilon_0(\mathbf{G}) = \epsilon_0(G) = \frac{4\pi}{G\,V_{\text{cell}}} \int_0^{\infty} dr\, r\, \sin Gr\, \epsilon_0(r)\,. \qquad (10)$$

To solve (1), we write

$$\psi(\mathbf{r},t) = \int d\mathbf{q}\, e^{i\mathbf{q}\cdot\mathbf{r}} \int d\omega\, e^{-i\omega t} \psi(\mathbf{q},\omega)\,,$$

$$\mathcal{E}(\mathbf{r},t) = \int d\mathbf{q}\, e^{i\mathbf{q}\cdot\mathbf{r}} \int d\omega\, e^{-i\omega t} \mathcal{E}(\mathbf{q},\omega)\,,$$

where $\mathbf{q}$ varies over all space. We also note that, for any function $f(\mathbf{q})$,

$$\int d\mathbf{q}\, f(\mathbf{q}) = \sum_{\mathbf{G}} \int_{\text{BZ}} d\mathbf{k}\, f(\mathbf{k}+\mathbf{G})\,,$$

and $\mathbf{k}$ varies over the first Brillouin zone only. In anticipation of discrete frequencies we also write

$$\psi(\mathbf{k}+\mathbf{G},\omega) = \sum_n \delta(\omega - \omega_{n\mathbf{k}})\phi_{n\mathbf{k}}(\mathbf{G})\,,$$

$$\mathcal{E}(\mathbf{k}+\mathbf{G},\omega) = \sum_n \delta(\omega - \omega_{n\mathbf{k}})\mathbf{E}_{n\mathbf{k}}(\mathbf{G})\,. \qquad (11)$$

Using the standard arguments of orthogonality of the Fourier basis the wave equations (1) can now be put into the form

$$|\mathbf{k}+\mathbf{G}|^2\phi(\mathbf{G}) - \frac{\omega^2}{c^2}\sum_{\mathbf{G}'}\epsilon(\mathbf{G}-\mathbf{G}')\phi(\mathbf{G}') = 0\,,$$
$$\qquad (12)$$

$$(\mathbf{k}+\mathbf{G})\times(\mathbf{k}+\mathbf{G})\times\mathbf{E}(\mathbf{G})$$

$$+\frac{\omega^2}{c^2}\sum_{\mathbf{G}'}\epsilon(\mathbf{G}-\mathbf{G}')\mathbf{E}(\mathbf{G}') = 0\,;$$

we have dropped the subscripts $n\mathbf{k}$ for simplicity. These equations define an infinite-dimensional generalized eigenvalue problem of the form

$$\mathsf{A}x = \lambda\mathsf{B}x\,,$$

where for the scalar case,

$$\mathsf{A}_{\mathbf{G}\mathbf{G}'} \leftrightarrow |\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{G}\mathbf{G}'}\,,$$
$$\mathsf{B}_{\mathbf{G}\mathbf{G}'} \leftrightarrow \epsilon(\mathbf{G}-\mathbf{G}')\,,$$
$$\mathsf{x}_{\mathbf{G}} \leftrightarrow \phi(\mathbf{G})\,,$$

and for vector waves,

$$\mathsf{A}^{ij}_{\mathbf{G}\mathbf{G}'} \leftrightarrow [|\mathbf{k}+\mathbf{G}|^2\delta^{ij} - (\mathbf{k}+\mathbf{G})^i(\mathbf{k}+\mathbf{G})^j]\delta_{\mathbf{G}\mathbf{G}'}\,,$$
$$\mathsf{B}^{ij}_{\mathbf{G}\mathbf{G}'} \leftrightarrow \epsilon(\mathbf{G}-\mathbf{G}')\delta^{ij}\,,$$
$$\mathsf{x}^i_{\mathbf{G}} \leftrightarrow \mathbf{E}^i(\mathbf{G})\,, \quad i,j = x,y,z$$

and

$$\lambda \leftrightarrow \frac{\omega^2}{c^2}\,.$$

We note that A and B are Hermitian matrices, and in addition B is positive definite. B is Hermitian because $\epsilon(\mathbf{r})$ is real, i.e., $\epsilon(\mathbf{G}) = \epsilon^*(-\mathbf{G})$, and it is positive definite because $\epsilon(\mathbf{r}) > 0$. The eigenvalue problem stated in this form is a generalization of the Hermitian eigenvalue problem. Using the Hermiticity of A and B, and the positive definiteness of B, one has, for a given $\mathbf{k}$,

$$x^{\dagger}\mathsf{B}x = C_{n\mathbf{k}}\delta_{nn'}\,,$$

where $C_{n\mathbf{k}}$ is a real positive but otherwise arbitrary normalization constant with the dimension of energy density. This effectively replaces the usual inner product $\langle\,|\,\rangle$ with the generalized inner product $\langle\,|\mathsf{B}|\,\rangle$. It is now straightforward to show that the Bloch functions $e^{-i\mathbf{k}\cdot\mathbf{r}}\mathbf{E}_{n\mathbf{k}}(\mathbf{r})$ are orthogonal with the weight function $\epsilon(\mathbf{r})$,

$$\int_{\text{all } \mathbf{r}} d\mathbf{r}\, e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}\, \epsilon(\mathbf{r})\, \mathbf{E}^*_{n'\mathbf{k}'}(\mathbf{r}) \cdot \mathbf{E}_{n\mathbf{k}}(\mathbf{r})$$

$$= C_{n\mathbf{k}}\delta_{nn'}\delta(\mathbf{k}-\mathbf{k}')\,.$$

To further simplify the problem, we notice first that A

is always real symmetric for both scalar and EM waves. For the special case when $\epsilon(\mathbf{r})$ has inversion symmetry, the origin can be chosen such that $\epsilon(\mathbf{r}) = \epsilon(-\mathbf{r})$ and $\epsilon(\mathbf{G})$ becomes real, with the result that B is now real symmetric and positive definite. This allows the use of standard library routines and substantially reduces computational effort.

Casting the problem into a form such as $\mathbf{B}^{-1}\mathbf{A} \cdot \mathbf{x} = \lambda \mathbf{x}$ obscures its underlying simple structure. Since $\mathbf{B}^{-1}\mathbf{A}$ is no longer Hermitian, this also adds unnecessary computational complexity to the problem; symmetric eigenvalue problems can be handled much more efficiently than unsymmetric ones. Clearly, solving $\det|\mathbf{A} - \lambda \mathbf{B}| = 0$ for λ is inefficient and unnecessary *unless* ϵ is a function of ω. For EM waves, $\omega \neq 0$ implies $\nabla \cdot \mathbf{D} = 0$. For $\mathbf{k} \neq 0$ this can be used to discard the N redundant $\omega = 0$ eigenvalues.

We also find that eliminating one Cartesian component of E not only destroys the symmetry of the problem, but also results in poorer accuracy due to an additional truncation error during the inversion involved. Finally, we feel that calculating $\epsilon(\mathbf{G})$ numerically rather than analytically should be avoided if possible. EM fields, time dependent or not, tend to be singular near sharp edges and corners.[11] The discrete Fourier transform, in effect, rounds these and may cause additional inaccuracies if the sampling mesh is not fine enough.

An alternate way of formulating the EM eigenvalue problem is to start with the equation satisfied by the magnetic field[9]

$$\nabla \times \eta(\mathbf{r})\nabla \times \mathcal{H}(\mathbf{r}, t) + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathcal{H}(\mathbf{r}, t) = 0, \qquad (13)$$

which, in Fourier space, takes the form

$$(\mathbf{k} + \mathbf{G}) \times \sum_{\mathbf{G}'} \eta_{\mathbf{G}\mathbf{G}'}(\mathbf{k} + \mathbf{G}') \times \mathbf{H}_{\mathbf{G}'} + \frac{\omega^2}{c^2}\mathbf{H}_{\mathbf{G}} = 0,$$

$$(14)$$

where $\eta(\mathbf{r}) \equiv 1/\epsilon(\mathbf{r})$ and $\eta_{\mathbf{G}\mathbf{G}'} \equiv \eta(\mathbf{G} - \mathbf{G}')$. This is an ordinary Hermitian eigenvalue problem. We note that

$$\epsilon(\mathbf{r})\eta(\mathbf{r}) = 1 \ \Rightarrow \ \sum_{\mathbf{G}''} \epsilon_{\mathbf{G}\mathbf{G}''}\eta_{\mathbf{G}''\mathbf{G}'} = \delta_{\mathbf{G}\mathbf{G}'}, \qquad (15)$$

i.e., $\eta_{\mathbf{G}\mathbf{G}'}$ is the inverse of the $\infty \times \infty$ matrix $\epsilon_{\mathbf{G}\mathbf{G}'}$. This ensures that the spectrum of both formulations would be identical, as they should be, when the problem is solved with infinite-dimensional matrices. However, when the matrix $\epsilon_{\mathbf{G}\mathbf{G}'}$ is first truncated at $N \times N$, and then inverted, the matrix obtained, $\bar{\epsilon}_{\mathbf{G}\mathbf{G}'}^{-1}$ can be, and often is if $\epsilon(\mathbf{r})$ is discontinuous, quite different from $\eta_{\mathbf{G}\mathbf{G}'}$. Thus the spectrum obtained from the truncated forms of Eqs. (12) and (14) will not be the same in general. We will call the finite forms of Eqs. (12) and (14) the E method and the H method, respectively.

Using $\nabla \cdot \nabla \times \mathbf{E} = 0$, the $3N \times 3N$ matrix equations (12) and (14) can be cast into $2N \times 2N$ ordinary Hermitian forms which are computationally more efficient. Hence the E method is identical to the method of Ref. 9 while

the H method uses the same matrix equation with $\bar{\epsilon}_{\mathbf{G}\mathbf{G}'}^{-1}$ replaced by $\eta_{\mathbf{G}\mathbf{G}'}$.

Yet another formulation of the problem[8] is to start with the equation satisfied by D,

$$\nabla \times \nabla \times \eta(\mathbf{r})\mathcal{D}(\mathbf{r}, t) + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathcal{D}(\mathbf{r}, t) = 0,$$

which is written in Fourier space as

$$(\mathbf{k} + \mathbf{G}) \times (\mathbf{k} + \mathbf{G}) \times \sum_{\mathbf{G}'} \eta_{\mathbf{G}\mathbf{G}'}\mathbf{D}_{\mathbf{G}'} + \frac{\omega^2}{c^2}\mathbf{D}_{\mathbf{G}} = 0.$$

This is a non-Hermitian form, but its truncated form is completely equivalent to the H method. Defining $Q_{\mathbf{G}\mathbf{G}'}^{ij} \equiv i\varepsilon^{ijk}(\mathbf{k} + \mathbf{G}) \cdot \hat{\mathbf{e}}^k \delta_{\mathbf{G}\mathbf{G}'}$, where ε^{ijk} is the Levi-Civita tensor, it can be written as

$$\mathbf{Q}\mathbf{Q}\eta\mathbf{D} + \frac{\omega^2}{c^2}\mathbf{D} = 0.$$

Multiplying on the left by $\mathbf{Q}\eta$ and regrouping, one obtains

$$\mathbf{Q}\eta\mathbf{Q}\,[\mathbf{Q}\eta\mathbf{D}] + \frac{\omega^2}{c^2}\,[\mathbf{Q}\eta\mathbf{D}] = 0,$$

which is the same eigenproblem in Eq. (14).

Finally, when both ϵ and μ are position dependent, one has to work with the $3N \times 3N$ generalized Hermitian form. One will then have four different "methods" depending on whether ϵ or $1/\epsilon$ and μ or $1/\mu$ is expanded in Fourier series.

II. TRUNCATION

One must truncate the matrix equation at N terms to solve it. We consider the $\mathbf{G} = 0$ term of the scalar wave equation to illustrate the issues involved.

$$|\mathbf{k} + 0|^2 \phi(\mathbf{G} = 0) = \frac{\omega^2}{c^2}\left[\sum_{|\mathbf{G}'| \leq G_{\max}} \epsilon(0 - \mathbf{G}')\phi(\mathbf{G}') \right.$$
$$\left. + \sum_{|\mathbf{G}'| > G_{\max}} \epsilon(0 - \mathbf{G}')\phi(\mathbf{G}')\right]$$

$$(16)$$

and the truncation is to be done so as to include all G points inside a sphere of radius $G_{\max}$. In order for the truncated equation to yield accurate solutions of the original problem, the second sum, which is left out, must be small compared to the first sum, which is retained. Thus we must require either

$$\alpha \equiv \frac{\|\epsilon - \epsilon_{\text{trunc}}\|}{\|\epsilon\|} \ll 1, \qquad (17)$$

where

$$\epsilon_{\text{trunc}}(\mathbf{r}) \equiv \sum_{|\mathbf{G}| \leq G_{\max}} \epsilon(\mathbf{G})\,e^{i\mathbf{G}\cdot\mathbf{r}}$$

and

$$\|\epsilon\|^2 \equiv \frac{1}{V_{\text{cell}}} \int_{\text{WS cell}} d\mathbf{r} \, |\epsilon(\mathbf{r})|^2 = \sum_{\mathbf{G}} |\epsilon(\mathbf{G})|^2 \,,$$

or

$$\frac{\|\phi - \phi_{\text{trunc}}\|}{\|\phi\|} \ll 1 \,, \tag{18}$$

where $\phi_{\text{trunc}}(\mathbf{r})$ is defined similarly.

We emphasize that *at least one* of the above must be satisfied in order to guarantee acceptable accuracy in the calculated frequencies. There is a third way for the second term to be negligible, namely when ϕ turns out to be orthogonal to ϵ within the truncated Fourier subspace spanned by $\{|\mathbf{G}\rangle \mid G > G_{\text{max}}\}$. This is clearly a special case and will not be considered here.

If *both* $\epsilon(\mathbf{G})$ and $\phi(\mathbf{G})$ are large for $|\mathbf{G}| > G_{\text{max}}$, then convergence of the eigenvalues is not guaranteed. The arguments so far apply for both scalar and vector waves. When $\epsilon(\mathbf{r})$ has a jump discontinuity, due to Gibbs phenomenon,[12] its Fourier transform will have a long tail. Hence the validity of truncation depends largely on the behavior of $\phi(\mathbf{G})$—$\mathbf{E}(\mathbf{G})$ for the EM problem. The scalar wave function $\phi(\mathbf{r})$ and its first derivatives are continuous; if $\epsilon(\mathbf{r})$ is discontinuous, then the second derivative(s) of $\phi(\mathbf{r})$ will be discontinuous: a relatively mild pathology. EM fields, $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, or $\mathbf{H}$, on the other hand, will, in general, themselves be discontinuous across the discontinuity in $\epsilon(\mathbf{r})$. Therefore they will have nonvanishing Fourier components even for very large values of $|\mathbf{G}|$. The physical origin of this is the polarization current/charge density which has a δ-function singularity at the surface of discontinuity, for which there is no analog in the scalar problem.

The discontinuity presents acute convergence problems for the vector problem even in the long-wavelength limit. Setting $\omega = 0$ in Eqs. (1), the scalar wave problem reduces to Laplace's equation, the only solution of which is $\phi(\mathbf{r}) = \phi_0$ =const. As a result, the effective long-wavelength dielectric constant is always $\bar{\epsilon}$ for scalar waves and the inequality (18) is automatically satisfied in the long-wavelength limit. For the vector wave equation, we obtain $\nabla \times \nabla \times \mathbf{E} = 0$. It may seem at first sight that $\mathbf{E}(\mathbf{r}) = \mathbf{E}_0$ is the solution in the long-wavelength limit. However, the solutions of the vector equation must also satisfy the auxiliary condition $\nabla \cdot \mathbf{D}(\mathbf{r}) = \mathbf{E}(\mathbf{r}) \cdot \nabla \epsilon(\mathbf{r}) + \epsilon(\mathbf{r}) \nabla \cdot \mathbf{E}(\mathbf{r}) = 0$. Hence $\mathbf{E}(\mathbf{r}) = \mathbf{E}_0$ cannot be a solution unless $\mathbf{E}_0 \cdot \nabla \epsilon(\mathbf{r}) = 0$ for all $\mathbf{r}$, which is realized only when $\epsilon(\mathbf{r}) = \epsilon(z)$. For spheres in any spatial arrangement, the solutions in this regime are very close to those of the familiar "dielectric sphere in a uniform $\mathbf{E}$ field" problem, provided that the separation of the spheres is large compared with their radii, i.e., for low filling fractions. For dielectric spheres, the $\mathbf{E}$ field is "attracted" toward the spheres and hence is largely perpendicular to the surface of the spheres and is very weak inside the spheres. From a convergence standpoint this implies that, in general, neither (17) nor (18) will be satisfied in the long-wavelength regime when the $\mathbf{E}$ field is

expanded in series, since it is highly discontinuous. The $\mathbf{D}$ field, on the other hand, is much better behaved than the $\mathbf{E}$ field. Hence a $\mathbf{D}$ field expansion, or equivalently the H method, is expected to work better for this case. For air spheres, the $\mathbf{E}$ field is "repelled" from the spheres and therefore is largely parallel to the spherical surface. Hence, it is by and large continuous across the interface. Therefore the E method would work much better. These considerations apply, to a large extent, for nonspherical atoms as well.

For more complicated geometries, one has no *a priori* reason to believe that one method would work better than the other, and hence both must be used and a sufficient number of plane waves must be included until reasonable consistency is obtained. Since the necessary condition for convergence involves the eigenfunctions, which are unknown, the sufficient condition is $\alpha \ll 1$. With the choice of the Euclidean norm, we have

$$\alpha = \sqrt{1 - \frac{\|\epsilon_{\text{trunc}}\|^2}{\|\epsilon\|^2}} \,. \tag{19}$$

α is the truncation error of the $\mathbf{G} = 0$ row of the matrix $\epsilon(\mathbf{G} - \mathbf{G}')$. The $\mathbf{G} \neq 0$ rows of $\epsilon(\mathbf{G} - \mathbf{G}')$ will have errors due not only to truncation, but *lopsided* truncation. The row elements of $\epsilon(\mathbf{G} - \mathbf{G}')$ are the Fourier coefficients within a sphere of radius G_{max}, centered at $\mathbf{G}$, *not* 0. The additional error introduced by the asymmetric truncation gets progressively worse as $|\mathbf{G}| \to G_{\text{max}}$. Therefore the above estimate for the truncation error would be meaningful only for the lowest-lying levels.

Some additional insight into the problem may be gained if we write

$$\epsilon(\mathbf{r}) = \bar{\epsilon} + \epsilon_{\text{fluc}}(\mathbf{r}) \equiv \bar{\epsilon} \left[1 + \epsilon_r(\mathbf{r}) \right] \,,$$

where $\bar{\epsilon}$ is the space average,

$$\bar{\epsilon} \equiv \frac{1}{V_{\text{cell}}} \int_{\text{WS cell}} d\mathbf{r} \, \epsilon(\mathbf{r}) \,,$$

and

$$\epsilon_r(\mathbf{r}) \equiv \frac{\epsilon_{\text{fluc}}(\mathbf{r})}{\bar{\epsilon}}$$

is the relative ripple. The significance of decomposing $\epsilon(\mathbf{r})$ into a space average part and a fluctuating part can be seen more clearly when one notes that the Fourier transform of $\bar{\epsilon}$ yields only the diagonal part of $\epsilon(\mathbf{G} - \mathbf{G}')$, while that of $\epsilon_{\text{fluc}}(\mathbf{r})$ yields only the off-diagonal part, which is solely responsible for the deviation from the free photon problem—since it mixes the different $\mathbf{q}$ components of $\psi(\mathbf{q}, \omega)$. Hence $\bar{\epsilon}$ is responsible only for "forward" scattering, or *no* scattering, while $\epsilon_{\text{fluc}}(\mathbf{r})$ is responsible for the off-axis scattering. $\|\epsilon_r\|$ is then, in a crude sense, the ratio scattering/no-scattering strength of the medium. Since $\bar{\epsilon}$ can be absorbed into $\bar{c} \equiv c/\sqrt{\bar{\epsilon}}$, we are left with $\|\epsilon_r\|$ as *the* measure of deviation from the free photon problem. Hence the problem naturally falls into three interesting categories:

1. "free" photons: $\|\epsilon_r\| \ll 1$,
2. "valence" photons: $\|\epsilon_r\| \approx 1$,
3. "bound" photons: $\|\epsilon_r\| \gg 1$;

we use the jargon for the electronic bands to emphasize qualitative similarities in the band structure. We demonstrate in the next sections that $\|\epsilon_r\| \approx 1$ is indeed the point where photonic gaps, or pseudogaps if gaps are prohibited by symmetry, begin to emerge.

We may now define a more strict measure of accuracy,

$$\alpha_r \equiv \frac{\|\epsilon_r - \epsilon_{r-\text{trunc}}\|}{\|\epsilon_r\|} ,$$

which is a measure of how accurately the *deviations* from the free photon case are represented in the truncated Fourier series.

III. HARD SPHERES

A. The Fourier transform

For a periodic array of dielectric spheres, each of radius R_s and dielectric constant ϵ_a, embedded in a host medium with ϵ_b,

$$\epsilon(\mathbf{r}) = \epsilon_b + (\epsilon_a - \epsilon_b)\sum_{\mathbf{R}} \theta(R_s - |\mathbf{r} - \mathbf{R}|) .$$

Its Fourier transform is

$$\epsilon(\mathbf{G}) = \epsilon(G) = \epsilon_b \delta_{\mathbf{G0}} + 3\beta(\epsilon_a - \epsilon_b)\frac{\sin x - x\cos x}{x^3} ,$$

where $x \equiv GR_s$ and β is the volume packing fraction of the ϵ_a material. Other quantities of interest are

$$\bar{\epsilon} = \beta\epsilon_a + (1 - \beta)\epsilon_b,$$

$$\|\epsilon\|^2 = \beta\epsilon_a^2 + (1 - \beta)\epsilon_b^2, \qquad (20)$$

$$\|\epsilon_r\|^2 = \frac{\|\epsilon\|^2}{\bar{\epsilon}^2} - 1 = \frac{\beta\epsilon_a^2 + (1 - \beta)\epsilon_b^2}{[\beta\epsilon_a + (1 - \beta)\epsilon_b]^2} - 1 .$$

It is worth noting that the expression for $\|\epsilon_r\|^2$ applies to any geometry or topology for a two-component dielectric medium, so it may also be relevant for the proposed photon localization in random dielectric structures. It also does not distinguish between "host" and "guest" materials and can easily be generalized to dielectric structures with more than two components.

Given ϵ_a and ϵ_b, the value of β that maximizes $\|\epsilon_r\|$ is

$$\beta_0 = \frac{\epsilon_b}{\epsilon_a + \epsilon_b} , \qquad (21)$$

and for this case

$$\|\epsilon_r\| = \frac{|\epsilon_a - \epsilon_b|}{2\sqrt{\epsilon_a \epsilon_b}} .$$

In light of the above, Yablonovitch and Gmitter[3] might have had a slightly better chance of seeing gaps (they used $\epsilon_b = 3.5^2$, $\epsilon_a = 1$) had they tried $\beta = 0.92$ rather

than $\beta = 0.86$.

For given ϵ_b and β,

$$\lim_{\epsilon_a \to \infty} \|\epsilon_r\|^2 = \frac{1 - \beta}{\beta} ,$$

consistent with the observations that increasing ϵ_a beyond a certain value has a negligible effect on scaled quantities such as $\Delta\omega/\omega$. In Fig. 1, contours of $\|\epsilon_r\|$ are plotted against (ϵ_a, β) for $\epsilon_b = 1$. The dashed line is the optimal value of β given by Eq. (21). Qualitatively there is good agreement with the computational results obtained in Refs. 6–9. For example, the computed relative gap $\Delta\omega/\omega$ as a function of ϵ_a and/or β peaks roughly where $\|\epsilon_r\|$ does. Clearly, the precise form of $\epsilon(\mathbf{r})$, and the scalar/vector character of the wave would ultimately determine the exact photonic levels. Despite its phenomenological nature, however, $\|\epsilon_r\|$ seems to explain some of the outstanding features of the problem fairly adequately. To investigate the convergence properties of the inverse transform of $\epsilon(G)$, we convert Eq. (9) to an integral over the G space. After performing angle integrations and dropping irrelevant constant factors, we obtain

$$\epsilon(\mathbf{r}) \sim \frac{1}{r}\int_0^\infty dx \sin\left(\frac{r}{R_s}x\right)\frac{\sin x - x\cos x}{x^2} .$$

Clearly, the integral of the absolute value of the integrand diverges logarithmically. What keeps the integral itself from diverging is the cosine factor which oscillates with almost a constant envelope $1/x$ as $x \to \infty$. This oscillatory convergence is very slow since $\Delta GR_s = 2\pi$, $\Delta G \sim \pi$ and the number of modes within a shell of thickness ΔG is $\Delta N \approx \frac{\Delta N}{\Delta^3 G}4\pi G^2\Delta G = \pi^2 G^2$. Hence the number of extra terms one needs to include in the summation in order to get the next cycle of the cosine term grows as G^2.
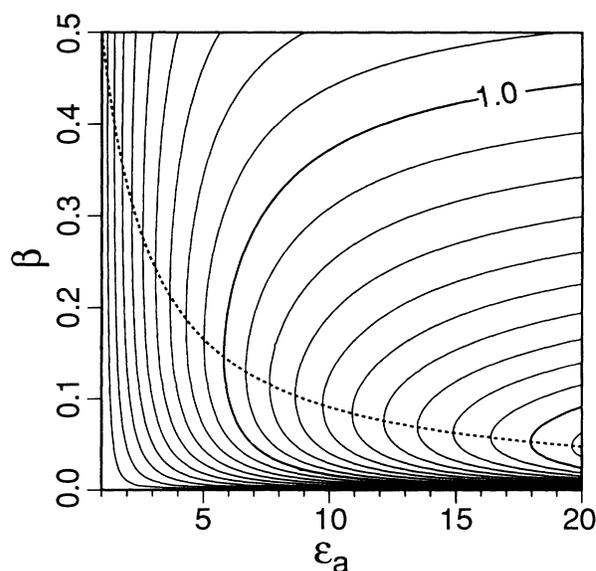

FIG. 1. Contours of $\|\epsilon_r\|$ vs ϵ_a and β for $\epsilon_b = 1$. The dashed line is the optimal value of β for a given ϵ_a given in Eq. (21).

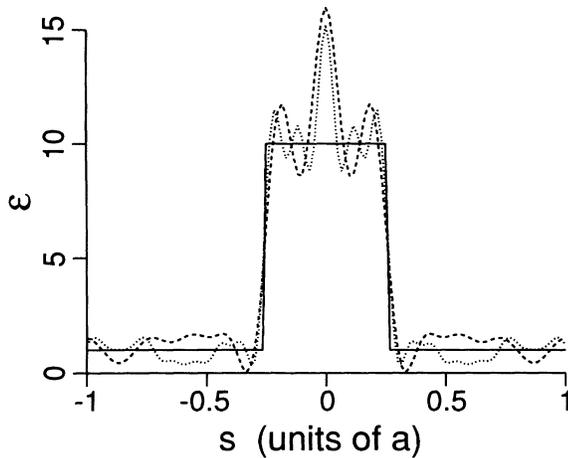FIG. 2. $\epsilon(\mathbf{r})$ (solid) and $\epsilon_{\mathrm{trunc}}(\mathbf{r})$ with $N=331$ (dashed) and $N=1139$ (dotted) for dielectric spheres in a fcc lattice along $x = y = z$ plotted against the path length s. $\epsilon_b = 1$, $\epsilon_a = 10$, and $\beta = 0.3$.

In Fig. 2 the partial sums of the Fourier series for the case $\epsilon_b = 1, \epsilon_a = 10$, and $\beta = 0.3$ is plotted for $N=331$, 1139, and ∞. It is clear that, just because increasing N does not produce visible differences in the resulting band structure, one has not necessarily converged to the "true" values. In this case, it is merely an indication of the slow convergence of the Fourier series.

B. The long-wavelength limit

Since the sufficiency condition for convergence is not satisfied for hard spheres, there remains the question of how accurate the calculated spectrum actually is for a particular case. For this we turn to an area that has been intensely investigated, the effective long-wavelength dielectric constant of a heterogeneous medium, defined by

$$\epsilon_{\mathrm{eff}} \equiv \lim_{k \to 0} \frac{c^2}{\left(\frac{d\omega}{dk}\right)^2} . \tag{22}$$

An experimental definition of ϵ_{eff} for a heterogeneous medium can be given by considering the capacitance of two large parallel plates sandwiching the medium,

$$\epsilon_{\mathrm{eff}} \equiv \lim_{A,d \to \infty} \frac{d}{A} C(A, d), \qquad A \gg d^2. \tag{23}$$

A reasonably accurate ϵ_{eff} is of paramount importance in band-structure calculations because it is related to the slope of the lowest two bands and hence to the accuracy of the first gap.[13] Previously ϵ_{eff} was calculated using the average value of the lowest four frequencies at the X point.[3,7,8] We prefer to calculate ϵ_{eff} using the lowest frequency at $\mathbf{k} \approx (2\pi/a)(0.05, 0, 0)$. The slight curvature of the bands here affects only the fourth digit of ϵ_{eff} and is therefore irrelevant.

We plot in Fig. 3 the coefficients α and $\epsilon_{\mathrm{eff}}^E$, calculated
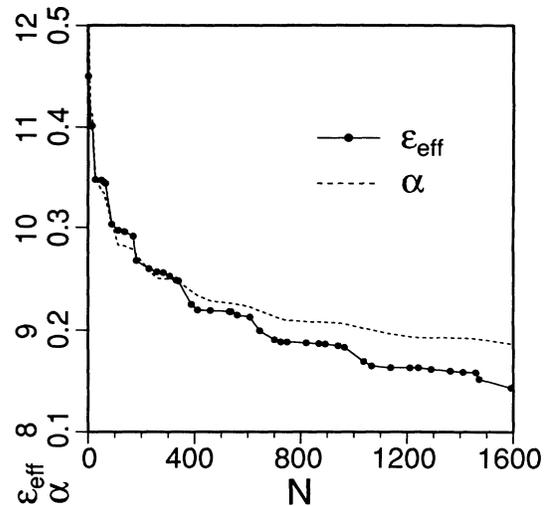

FIG. 3. α (dashed) and $\epsilon_{\mathrm{eff}}^E$ (solid) for dielectric spheres in a fcc lattice, plotted as a function of N. $\epsilon_b = 1$, $\epsilon_a = 16$, and $\beta = 0.7$. The staircase appearance is a reflection of oscillatory convergence. The change in ϵ_{eff} in each step has not significantly diminished even for $N=1600$.

with the E method, for the case of dielectric spheres in fcc with $\epsilon_b = 1$, $\epsilon_a = 16$, and $\beta = 0.7$ with N up to 1600. The behavior of ϵ_{eff} and α is very similar in appearance. The extrema of $\epsilon(x)$ correspond to regions where α and ϵ_{eff} change rapidly, and wide plateaus are observed when $x_{\max}$ is close to a zero of $\epsilon(x)$. It is these plateaus that grow wider with $N^{2/3}$ that give the false impression that convergence has been achieved. However, this "seasonal" variation should be disregarded, and we should really be looking at the "trend" as the true test of convergence. For this, one needs to increment $x_{\max} \equiv R_s G_{\max}$ in intervals of π. Since the truncation volume in G space is spherical, for large N, $N^{-1/3} \propto 1/G_{\max}$. When $x_{\max}$ is large, the sine term in $\epsilon(x)$ can be ignored and we have,

$$\alpha^2 \propto \alpha_r^2 \propto \sum_{|\mathbf{G}|>G_{\max}} |\epsilon(\mathbf{G})|^2 \approx \int_{x_{\max}}^{\infty} 4\pi x^2 \frac{\cos^2 x}{x^4} dx$$

$$\sim \int_{x_{\max}}^{\infty} \frac{dx}{x^2} = \frac{1}{x_{\max}} = \frac{1}{R_s G_{\max}} \sim N^{-1/3}, \tag{24}$$

where the seasonal variation has effectively been smeared out by ignoring the $\cos^2 x$ factor. In Fig. 4 we plot $\epsilon_{\mathrm{eff}}^H$ and $\epsilon_{\mathrm{eff}}^E$ as a function of $N^{-1/3}$ for dielectric and air spheres packed at the maximum filling ratio of $\pi/6$ in a simple cubic lattice. These results agree well with the results of Ref. 13. It is seen readily that the E method is reasonably accurate for air spheres but fails badly for dielectric spheres, and the opposite holds, to a large extent, for the H method. We found similar convergence characteristics for the two methods in other periodic arrangements we considered, such as body centered cubic, diamond structure, and fcc.

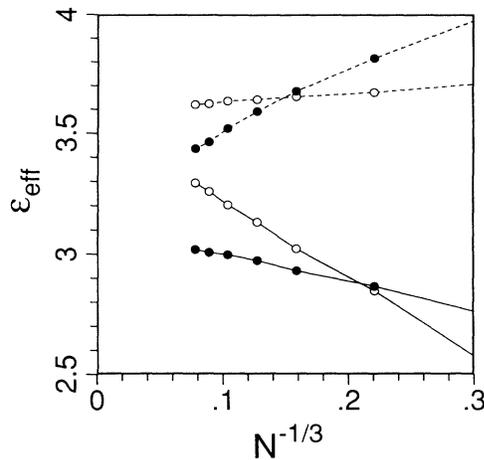The difference between the results of the two methods for finite N is to be expected. Using the lowest-order

FIG. 4. ϵ_{eff}^E (dashed curve) and ϵ_{eff}^H (solid curve) vs $N^{-1/3}$ for air (hollow circles) and dielectric (filled circles) spheres in the simple cubic lattice. ϵ_b, $\epsilon_a = 1, 8$, and $\beta = \pi/6$.

Fourier expansion, i.e., $N = 1$, one obtains $\epsilon_{\text{eff}}^E = \bar{\epsilon}$ with the E method, whereas with the H method, $\epsilon_{\text{eff}}^H = 1/\bar{\eta}$. It is easy to verify, using parallel and series combinations of capacitors, that these are the two possible extreme values that ϵ_{eff} can assume for *any* dielectric function $\epsilon(\mathbf{r})$, and indeed both are realized for a one-dimensional (periodic or nonperiodic) dielectric medium, $\epsilon(\mathbf{r}) = \epsilon(z)$: $\bar{\epsilon}$ is the ordinary ϵ_{eff}, whereas $1/\bar{\eta}$ is the extraordinary. Hence, for a given medium, depending on whether the actual ϵ_{eff} is closer to $\bar{\epsilon}$ or $1/\bar{\eta}$, one method can be expected to be more accurate than the other. Considering the **D** field lines for air and dielectric atoms described earlier, one readily sees that series and parallel combinations would be good first approximations to ϵ_{eff} for dielectric and air atoms, respectively. Indeed, for cubic atoms in the simple cubic lattice, a somewhat more refined three-capacitor model (connected differently for air and dielectric atoms) reproduces the Maxwell-Garnett results with remarkable accuracy.

In Fig. 5 we plot ϵ_{eff} vs β for air and dielectric spheres in fcc obtained by a simple linear extrapolation. The discrepancy between the *extrapolated* values from each method was 1–5 %. Up to 2200 plane waves were used for dielectric spheres. The agreement with the Maxwell-Garnett theory is good for the lower volume fractions and qualitatively it compares well with the findings of a study of the simple cubic lattice by Lamb *et al.*,[13] who use the Korringa-Kohn-Rostoker (KKR) method. The effective-medium result, in which both types of materials appear on an equal footing, is particularly inaccurate for this topology[13] except at very low volume fractions. We found, for the topologies where both a- and b-type materials are connected, that ϵ_{eff} is well modeled by the effective-medium theory. For dielectric spheres at high volume fractions, the deviation from the Maxwell-Garnett results has been attributed to proximity effects[13,14] which induce higher-order multipoles. From a Fourier expansion point of view, close packing means that the scale of change of $\epsilon(\mathbf{r})$ and that of the
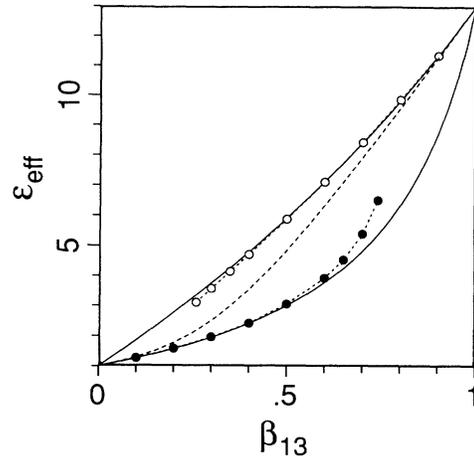

FIG. 5. ϵ_{eff} vs β_{13}, the volume fraction of $\epsilon = 13$ material for dielectric spheres (filled circles) and air spheres (hollow circles) in fcc calculated by the plane-wave method. The solid curves are the corresponding results from Maxwell-Garnett theory. The dashed curve is the prediction of the effective-medium theory. See Ref. 13 for a detailed discussion.

fields is small, hence the need for very large **G** components.

It is important that the E and the H methods yield consistent results in the limit $N^{-1/3} \to 0$. As an example of a case where both methods fail badly, we plot in Fig. 6 ϵ_{eff}^E and ϵ_{eff}^H vs $N^{-1/3}$ for cubic dielectric atoms in a simple cubic lattice with $\epsilon_b = 1$, $\epsilon_a = 13$, and $\beta = 0.9$. A cubic truncation in **G** space was employed. The two results differ by $\sim 35\%$ for N as large as 2200. Furthermore, even an extrapolation to $N^{-1/3} \to 0$ seems hopeless.

A word about extrapolation. No one likes to make extrapolations and we are no exception. However, when one uses a method that employs an infinite series expansion,
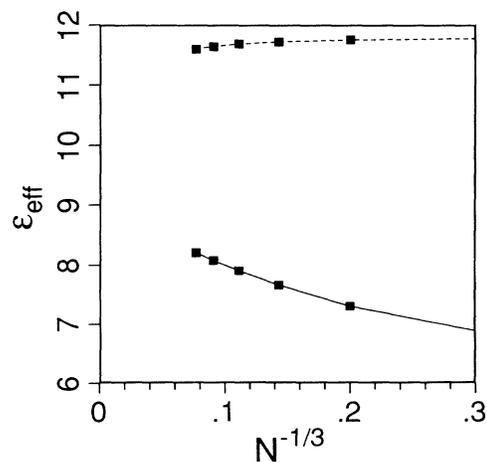

FIG. 6. ϵ_{eff}^E (dashed) and ϵ_{eff}^H (solid) vs $N^{-1/3}$ for dielectric cubic atoms in the simple cubic lattice. $\epsilon_b = 1$, $\epsilon_a = 13$, and $\beta = 0.9$.
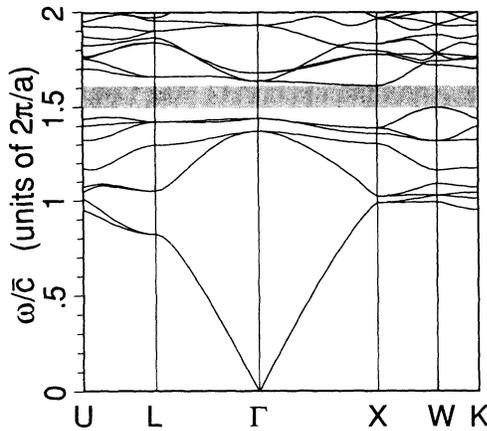
FIG. 7. Band structure for EM waves for air spheres in a fcc lattice. $\epsilon_b = 16$, $\epsilon_a = 1$, $\beta = 0.74$, $\|\epsilon_r\| = 1.34$, $N = 1211$.

one *has to* make an extrapolation since one is actually interested in the exact solution which involves an infinite series, and not in the result obtained with a finite number of terms, *per se*. Hence, *not* making an extrapolation is, quite possibly, making the *worst* extrapolation. It is equivalent to saying that the results obtained with the finite number of terms used is *the* solution.

C. The band structure

With dielectric spheres in fcc we were unable to find any gaps using either method. For air spheres, we find a gap of significant size between the eighth and ninth levels, and a direct gap between the fifth and sixth. We say that there is a direct gap between the nth and $(n+1)$th levels when $\omega_{n+1}(\mathbf{k}) > \omega_n(\mathbf{k})$ for all $\mathbf{k}$, whereas a gap is said to exist when $\omega_{n+1}(\mathbf{k}) > \omega_n(\mathbf{k}')$ for all $\mathbf{k}, \mathbf{k}'$.

In Fig. 7 we plot the band structure calculated with the E method which converges better for this case. We spanned the parameter space with ϵ_a fixed at 1 and ϵ_b varied between 1 to 100 and β from 0 to 0.74, and found that the 8-9 gap opens for $\|\epsilon_r\| > 1$. In Fig. 8 we
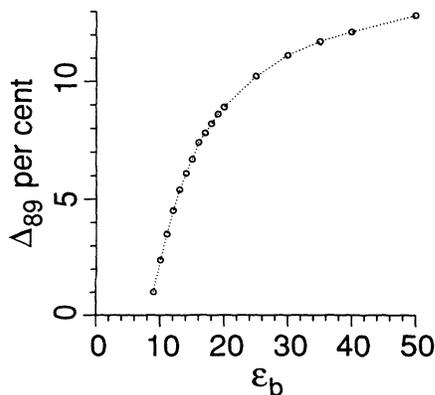
plot the relative 8-9 gap $\Delta_{89} \equiv 2(\omega_{9\mathrm{min}} - \omega_{8\mathrm{max}})/(\omega_{9\mathrm{min}} +\omega_{8\mathrm{max}})$ vs ϵ_b with $\beta = 0.74$ and $\epsilon_a = 1$. We computed the gap with the E method using $N \approx 110, 330, 750, 1200$, and 1600 and extrapolated to $N^{-1/3} \to 0$. For these values of N, we obtain, for $\epsilon_b=16$, $\Delta_{89}^E=6.2\%$, 7.1%, 7.3%, 7.35%, and 7.4%, respectively. The convergence of the H method is much worse, as would be expected, and we obtain $\Delta_{89}^H=0\%$, 1.4%, 7.1%, 8.3%, and 8.45%, respectively. It is because both methods yield a gap of increasing width as N increases that we are able to say, with a good degree of confidence, that this gap is not an artifact of poor convergence.

IV. GAUSSIAN SPHERES

The limitations of the plane-wave method for dielectric hard spheres led us to consider the Gaussian form for the dielectric function,

$$\epsilon(\mathbf{r}) = \epsilon_b + (\epsilon_a - \epsilon_b) \sum_{\mathbf{R}} \exp\left[-\frac{|\mathbf{r} - \mathbf{R}|^2}{2\sigma^2}\right] , \quad (25)$$

$$\epsilon(\mathbf{G}) = \epsilon(G)$$
$$= \epsilon_b \delta_{\mathbf{G}0} + (\epsilon_a - \epsilon_b)\frac{(2\pi\sigma^2)^{3/2}}{V_{\mathrm{cell}}} \exp\left[-\frac{G^2\sigma^2}{2}\right] . \quad (26)$$

This choice of $\epsilon(\mathbf{r})$ is artificial and practically impossible to realize from an engineering point of view, and we certainly are not proposing it as a solution to the hard-sphere problem. However, it retains one fundamental nature of the problem, namely, periodicity with large variations. The major physical difference between hard spheres and Gaussian spheres is that, whereas with hard spheres the polarization charge/current density is confined to the surface of the discontinuity and hence to within each primitive cell, with Gaussian spheres polarization charges are distributed and are free to move throughout the entire crystal. Therefore the model of Gaussian spheres could be interesting in its own right;

FIG. 8. The relative 8-9 gap Δ_{89} (see the text) vs ϵ_b with $\epsilon_a = 1$ and $\beta = 0.74$ for air spheres in fcc.
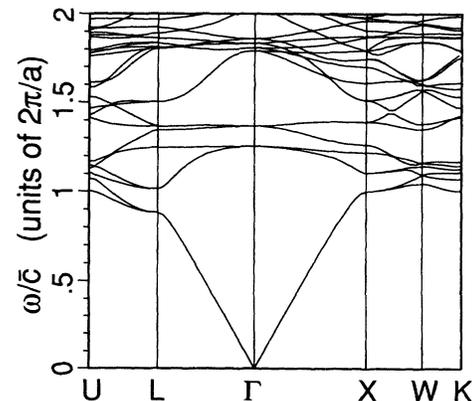

FIG. 9. Band structure for EM waves for Gaussian spheres in the fcc lattice. $\epsilon_b = 1$, $\epsilon_a = 25$, $\sigma = 0.5$, $\|\epsilon_r\| = 1.38$, $N = 307$.
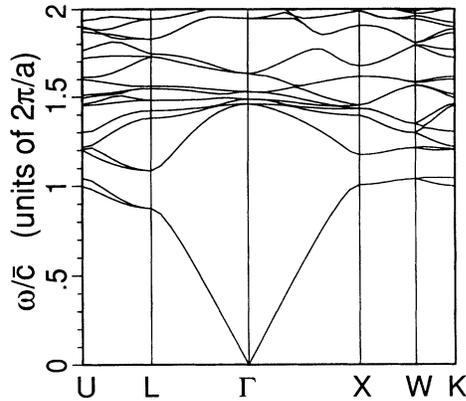
FIG. 10. Band structure for EM waves for Gaussian spheres in the diamond structure. $\epsilon_b = 1$, $\epsilon_a = 25$, $\sigma = 0.5$, $\|\epsilon_r\| = 1.29$, $N = 307$.

FIG. 12. Band structure for EM waves for Gaussian spheres in the fcc lattice. $\epsilon_b = 1$, $\epsilon_a = 10^6$, $\sigma = 0.5$, $\|\epsilon_r\| = 3.18$, $N = 1243$.

and because its Fourier transform converges *much* faster, it provides a good, consistent "textbook case" to see the general behavior of photonic bands in a wide range of ω. On the other hand, since the discontinuity of hard spheres is poorly represented by a Fourier series, the results obtained with the plane-wave method should be used with extreme caution and from a "dynamic" perspective, i.e., as $G_{\max} \rightarrow \infty$, since a lot of physics may, and does, critically depend on the discontinuity itself.

We were able to achieve a very high degree of accuracy, as measured by α_r, in computing the bands up to $\tilde{\omega} \equiv \frac{\omega}{c}\sqrt{\bar{\epsilon}}\frac{a}{2\pi} \approx 2$. For $\sigma = 0.5$, with 307 Fourier terms one has $\alpha_r < 0.01$. A plot similar to Fig. 2 for Gaussian spheres shows that $\epsilon_{\text{trunc}}(\mathbf{r})$ and $\epsilon(\mathbf{r})$ are practically indistinguishable. For EM waves in a fcc structure, we did not see any gaps for the lower-lying levels. We also tried Gaussian-type atoms in a diamond structure and clearly there are gaps. However, the lowest gap seems to be very sensitive to the overlap of $\epsilon_0(\mathbf{r})$ from adjacent atoms and vanishes as $\sigma \rightarrow 0$. Therefore we computed

the band structure for hard spheres in diamond near the close-packing arrangement with both the E and the H methods and found that for $\beta < 0.34$, the gap disappears as more and more plane waves are included.[15] When a finite number of Fourier terms are retained, the partial sums of these are "extended" atoms. As N grows larger, the truncated dielectric function slowly approaches the actual discontinuous function for which the polarization charges/currents are localized—and the gap disappears.

Results for the Gaussian sphere model band structure are displayed in Figs. 9 and 10. The parameters in both cases are $\epsilon_b = 1$, $\epsilon_a = 25$, and $\sigma = 0.5$. These show a structure similar to that found in the hard-sphere calculations. Our value of the peak dielectric constant is larger than those in the hard-sphere calculations, but it is comparable to the peak of $\epsilon_{\text{trunc}}(\mathbf{r})$ in hard-sphere models.

We also tried the "limit" $\epsilon_a \rightarrow \infty$ at $\epsilon_a = 10^6$, $\sigma = 0.5$. In Figs. 11–14 we plot the band structure for this case for scalar/vector and fcc/diamond combinations. Prob-
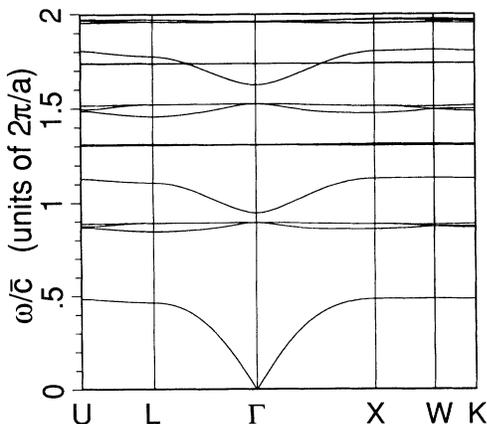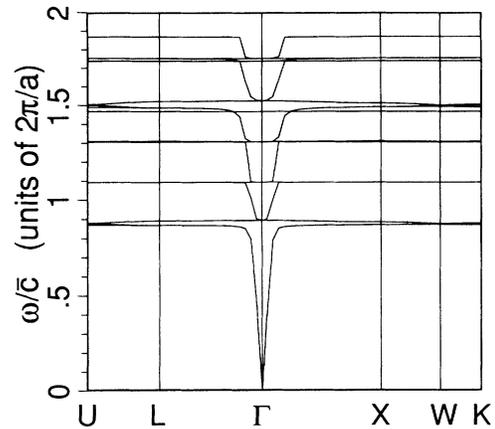

FIG. 11. Band structure for scalar waves for Gaussian spheres in the fcc lattice. $\epsilon_b = 1$, $\epsilon_a = 10^6$, $\sigma = 0.5$, $\|\epsilon_r\| = 3.18$, $N = 1243$.
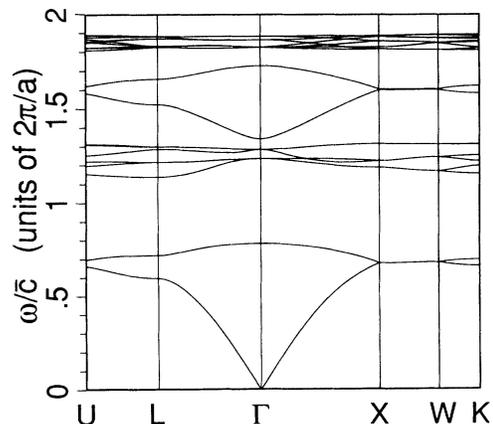

FIG. 13. Band structure for scalar waves for Gaussian spheres in the diamond structure. $\epsilon_b = 1$, $\epsilon_a = 10^6$, $\sigma = 0.5$, $\|\epsilon_r\| = 2.14$, $N = 749$.
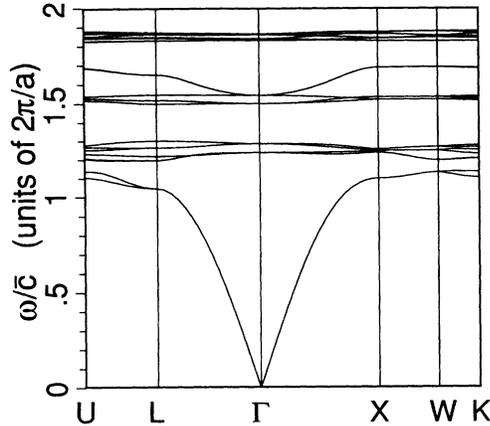
FIG. 14. Band structure for EM waves for Gaussian spheres in the diamond structure. $\epsilon_b = 1$, $\epsilon_a = 10^6$, $\sigma = 0.5$, $\|\epsilon_r\| = 2.14$, $N = 749$.

ably the most interesting feature of these bands is their apparent simplicity. Although convergence of the higher frequencies is not as good, the pattern is clear. One sees the unmistakable quantization of frequencies at this limit. Also notable is that for a given type of lattice, the stationary frequencies for scalar and vector waves coincide most of the time. This is to be expected for those solutions of the EM wave equation for which $\nabla \cdot \mathbf{E} = 0$, since then the scalar and the vector eigenvalue problems become identical.

V. CONCLUSION

We found that the discontinuous nature of $\epsilon(\mathbf{r})$ and the EM fields for a periodic array of dielectric spheres severely limits the accuracy of the plane-wave method. Fluctuation of the truncated series from the actual case is large and the convergence is very slow.

We have also formulated the problem in the form of a generalized Hermitian eigenvalue problem that has many features in common with the ordinary Hermitian problem that gives rise to electronic band structures. It is, however, limited to the case where the dielectric function is frequency independent. We observe that the scattering strength parameter $\|\epsilon_r\|$ plays a crucial role in the photonic band structure. When it is weak, the photonic band structure resembles the free photon case, and when it is strong the bands become degenerate and flat and resemble the harmonic-oscillator spectrum.

The Gaussian sphere model gave us good convergence and provided a good model upon which we can draw reliable conclusions. We find for this model that, for EM

waves in a fcc lattice, there is no gap between the second and third bands consistent with the results of previous numerical work on hard spheres. For Gaussian spheres in diamond structure we find large gaps between the second and third bands.

With hard spheres, we find a sizable gap for the air spheres in fcc between the eighth and ninth bands but none between the second and third bands. For hard spheres in the diamond structure, packed with $\beta < 0.34$ there are no gaps. At the close-packing fraction of $\beta = 0.34$, our best estimate of the gap is $\Delta_{23} < 3.5\%$ for $\epsilon_a \approx 13$. This, we believe, is due to a change in the topology of the medium.

The plane-wave method is clearly an attractive method because of its simplicity and its applicability, at least in principle, to any type of $\epsilon(\mathbf{r})$. We have tried to make the best of the method by using a "Cauchy sequence" approach and the complementary E and H methods, which is necessary when all the fields $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{H}$ behave about equally badly. Hence, there is a need to solve the problem for hard spheres with other methods, such as KKR or APW (augmented plane wave), although we suspect that the discontinuous nature of $\epsilon(\mathbf{r})$ and the fields would still pose problems. This need is particularly acute for high dielectric contrasts and near close-packing ratios and for higher frequencies, an accurate calculation of which would bring invaluable insight into the problem.

Future work is planned to be devoted to spheroid-shaped particles where the polarization degeneracy for EM waves is broken at the long-wavelength limit and gaps are observed in fcc structures.[16]

Note added. A structure consisting of cylindrical holes drilled along the [110], [101], and [011] directions in fcc, another similar structure in the diamond lattice, and a similar structure in the simple cubic lattice have recently been shown to have large gaps.[17] We have also recently become aware of a criterion alternative to Eq. (21) for the volume fraction that maximizes the band gap, $\beta_0 = 1/2n$, given by John,[10] who considers a one-dimensional structure and imposes the condition that the Bragg and the Mie resonances occur at the same frequency.

---

*Present address: Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071.

†Present address: Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153, Japan.

[1] E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987).

[2] S. John, Phys. Rev. Lett. **58**, 2486 (1987).

[3] E. Yablonovitch and T. J. Gmitter, Phys. Rev. Lett. **63**, 1950 (1989).

[4] S. John and R. Rangarajan, Phys. Rev. B **38**, 10 101 (1988).

[5] E. N. Economou and A. Zdetsis, Phys. Rev. B **40**, 1334 (1989).

[6] S. Satpathy, Z. Zhang, and M. R. Salehpour, Phys. Rev. Lett. **64**, 1239 (1990).

[7] K. M. Leung and Y. F. Liu, Phys. Rev. Lett. **65**, 2646 (1990).

[8] Z. Zhang and S. Satpathy, Phys. Rev. Lett. **65**, 2650 (1990).

[9] K. M. Ho, C. T. Chan, and C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990).

[10] G. Kurizki and A. Z. Genack, Phys. Rev. Lett. **61**, 2269 (1988); R. J. Glauber and M. Lewenstein, Phys. Rev. A **43**, 467 (1991); S. John, Phys. Today **44** (5), 32 (1991); S. John and J. Wang, Phys. Rev. Lett. **64**, 2418 (1990). In the last reference, the calculations are based on the assumption that the electromagnetic field modes, i.e., the vector potential **A**, in a dielectric medium are transverse plane waves.

[11] See, for example, J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), pp. 75–78.

[12] See, for example, Ivar Stakgold, *Green's Functions and Boundary Value Problems* (Wiley-Interscience, London, 1979), p. 134.

[13] W. Lamb, D. M. Wood, and N. W Ashcroft, Phys. Rev. B **21**, 2248 (1980).

[14] R. C. McPhedran and D. R. McKenzie, Proc. R. Soc. London Ser. A **359**, 45 (1978).

[15] H. S. Sözüer, J. W. Haus, and R. Inguva (unpublished).

[16] J. W. Haus, H. S. Sözüer, and R. Inguva (unpublished).

[17] E. Yablonovitch and K. M. Leung, Physica B **175**, 81 (1991); C. T. Chan, K. M. Ho, and C. M. Soukoulis, Europhys. Lett. **16**, 563 (1991); H. S. Sözüer and J. W. Haus (unpublished).
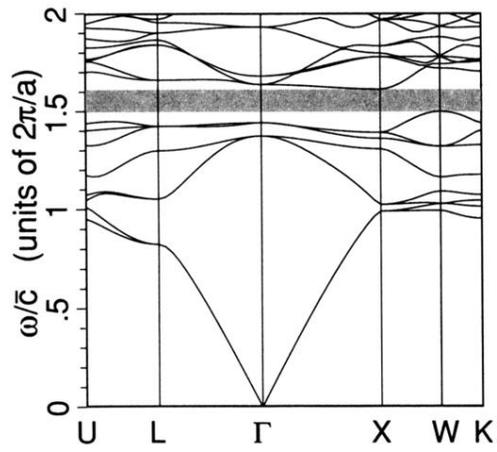
FIG. 7.   Band structure for EM waves for air spheres in a fcc lattice. $\epsilon_b = 16$, $\epsilon_a = 1$, $\beta = 0.74$, $\|\epsilon_r\| = 1.34$, $N = 1211$.