Localization of light in lossless inhomogeneous dielectrics

Alexander Figotin

Department of Mathematics, University of North Carolina at Charlotte, Charlotte, North Carolina 28223

Abel Klein

Department of Mathematics, University of California, Irvine, Irvine, California 92697-3875

Received September 3, 1997; revised manuscript received November 21, 1997, accepted December 2, 1997

The localization of electromagnetic waves in lossless inhomogeneous dielectric media is studied. We consider a three-dimensional lossless periodic medium (photonic crystal) having a gap in the frequency spectrum (photonic bandgap). If such a medium is perturbed by either a single defect or a random array of defects, exponentially localized electromagnetic waves arise with frequencies in the gap. For a single defect, we derive equations for these midgap frequencies and estimate their number. For a random medium, we show the occurrence of Anderson localization for electromagnetic waves. © 1998 Optical Society of America
[S0740-3232(98)00305-6]

OCIS codes: 260.2110, 350.7420.

1. INTRODUCTION

Localization of classical electromagnetic (EM) waves has received much attention in recent years.[1–11] This phenomenon arises from coherent multiple scattering and interference and occurs when the scale of the coherent multiple scattering reduces to the wavelength itself. Numerous potential applications[4,6,7,12–15] and the fundamental significance of the localization of classical waves motivate the interest in this phenomenon.

The basic situation for localization that we consider here is as follows. We start with a perfectly periodic lossless dielectric medium, which we call a photonic crystal.[16] The propagation of EM waves in photonic crystals has been the subject of intensive study in recent years.[6,7,12,13,15–20] The most significant manifestation of coherent multiple scattering in the periodic medium is the rise of a gap in the spectrum, called a photonic band gap.[13,14] If a periodic dielectric medium with a bandgap is perturbed by either a single defect (impurity) or a random array of defects, localized EM waves can arise under some conditions. The frequencies of these localized waves lie in the gap. In the case of a single defect the localized eigenmodes are often called defect or impurity midgap eigenmodes. In the case of a random medium the phenomenon of localization has the same nature as the Anderson localization of electrons,[21–23] which is now well understood in the mathematical literature.[24–33]

The physical origin of photonic bandgaps and the localization of EM waves is the same: multiple scattering and destructive wave interference. The propagation (or nonpropagation) of EM waves in photonic crystals and EM wave localization are intimately related and are reflected in our mathematical studies.[8–11]

Localization of a wave caused by a single defect or by a random array of defects in a perfectly periodic medium is a general wave phenomenon. In addition to electron and EM waves, this phenomenon is also relevant to acoustic waves,[8,9,11,30,32–34] elastic waves,[35] acoustic phonons,[36] and more-complicated excitations involving coupled waves such as polaritons.[37]

The subject of this paper is the localization of classical EM waves in a lossless linear dielectric medium in three dimensions. The rigorous investigation of the propagation of EM waves in three-dimensional inhomogeneous media poses challenging mathematical problems. In this paper we intentionally select and deal only with those quantities that can be treated rigorously.

In spite of the burden of mathematical rigor, some physically important quantities can be, and have been, studied in detail. For instance, (i) we found sufficient conditions for a defect in a periodic dielectric with a spectral gap to generate midgap defect eigenmodes and gave *a priori* estimates on their radii of localization[10]; (ii) we derived equations for the midgap defect frequencies and obtained estimates of the number of those frequencies in the gap[11]; (iii) we proved the occurrence of Anderson localization of EM waves in random media.[9]

Our focus is on the mathematical concepts and methods that give a solid mathematical basis to the physical theory, as well as on the tools for reliable computational schemes for the quantities describing the localization of EM waves. The relevant statements are formulated in the form of theorems and have been rigorously proved.[8–11,32,33]

2. INHOMOGENEOUS DIELECTRIC MEDIA

We assume that the propagation of EM waves is described by the classical Maxwell equations

$$\frac{\partial}{\partial t}\mathbf{B} = -\nabla \times \mathbf{E}, \qquad \nabla \cdot \mathbf{B} = 0,$$

$$\frac{\partial}{\partial t}\mathbf{D} = \nabla \times \mathbf{H} \qquad \nabla \cdot \mathbf{D} = 0, \tag{1}$$

with the linear constitutive relationships

$$\mathbf{B}(\mathbf{x}, t) = \mu(\mathbf{x})\mathbf{H}(\mathbf{x}, t),$$
$$\mathbf{D}(\mathbf{x}, t) = \varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x}, t). \tag{2}$$

We use the Giorgi system of units. The vector fields **E**, **D**, **H**, and **B** are the position- and time-dependent electric field, electric induction, magnetic field, and magnetic induction, respectively.

Since we consider inhomogeneous media, the dielectric constant $\varepsilon = \varepsilon(\mathbf{x})$ and the magnetic permeability $\mu = \mu(\mathbf{x})$ are, in general, position dependent. Below we consider media for which the magnetic permeability is approximately constant, so we shall take it to be identically 1. As for the dielectric constant $\varepsilon(\mathbf{x})$, we neglect its frequency dependence, and, since the medium is lossless, $\varepsilon(\mathbf{x})$ is real valued. We always assume that

$$0 < \varepsilon_- \leqslant \varepsilon(\mathbf{x}) \leqslant \varepsilon_+ < \infty \tag{3}$$

for some constants ε_- and ε_+.

The energy density $\mathscr{E}(\mathbf{x}, t) = \mathscr{E}_{\mathbf{H},\mathbf{E}}(\mathbf{x}, t)$ and the (conserved) energy of a solution $(\mathbf{H}, \mathbf{E})$ of the Maxwell's equations (1) are given by

$$\mathscr{E}(\mathbf{x}, t) = 1/2[\mu(\mathbf{x})|\mathbf{H}(\mathbf{x}, t)|^2 + \varepsilon(\mathbf{x})|\mathbf{E}(\mathbf{x}, t)|^2],$$
$$\mathscr{E} = \mathscr{E}_{\mathbf{H},\mathbf{E}} = \int_{\mathbb{R}^3} \mathscr{E}(\mathbf{x}, t)\mathrm{d}\mathbf{x}, \tag{4}$$

where $\mathbb{R}^3$ is three-dimensional space.

Maxwell's equations can be recast as a Schrödinger-like equation, i.e., as a first-order conservative linear equation, of the form

$$-i \frac{\partial}{\partial t} \boldsymbol{\Psi}_t = \mathbb{M}\boldsymbol{\Psi}_t, \tag{5}$$

with

$$\boldsymbol{\Psi}_t = \begin{pmatrix} \mathbf{H}_t \\ \mathbf{E}_t \end{pmatrix}, \qquad \mathbb{M} = \begin{bmatrix} 0 & \dfrac{i}{\mu} \nabla^{\times} \\ \dfrac{-i}{\varepsilon} \nabla^{\times} & 0 \end{bmatrix}, \tag{6}$$

where $\nabla^{\times}$ is the symbol for the curl operator, i.e., $[\nabla^{\times}\mathbf{H}](\mathbf{x}) = \mathrm{curl}\ \mathbf{H}(\mathbf{x}) = \nabla \times \mathbf{H}(\mathbf{x})$.

It is convenient and appropriate to introduce scalar products for the fields $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ as follows:

$$(\mathbf{H}_1, \mathbf{H}_2)_\mu = \int_{\mathbb{R}^3} \overline{\mathbf{H}_1(\mathbf{x})}\mathbf{H}_2(\mathbf{x})\mu(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{7}$$

$$(\mathbf{E}_1, \mathbf{E}_2) = \int_{\mathbb{R}^3} \overline{\mathbf{E}_1(\mathbf{x})}\mathbf{E}_2(\mathbf{x})\varepsilon(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{8}$$

where $\bar{\mathbf{E}}$ is complex conjugate to **E**. The corresponding norms are defined as usual: $\|\mathbf{H}\|_\mu = \sqrt{(\mathbf{H}, \mathbf{H})_\mu}$ and $\|\mathbf{E}\|_\varepsilon = \sqrt{(\mathbf{E}, \mathbf{E})_\varepsilon}$.

$\mathbb{S}_\mu$ is defined as the Hilbert space of solenoidal magnetic fields $\mathbf{H}(\mathbf{x})$:

$$\|\mathbf{H}\|_\mu < \infty, \quad \nabla \cdot \mu(\mathbf{x})\mathbf{H}(\mathbf{x}) = 0. \tag{9}$$

Similarly, we introduce $\mathbb{S}_\varepsilon$ as the Hilbert space of solenoidal electric fields $\mathbf{E}(\mathbf{x})$:

$$\|\mathbf{E}\|_\varepsilon < \infty, \quad \nabla \cdot \varepsilon(\mathbf{x})\mathbf{E}(\mathbf{x}) = 0. \tag{10}$$

The Hilbert space $\mathbb{H}_{\mu,\varepsilon}$ of finite-energy EM fields is now defined as the set of pairs $\boldsymbol{\Psi} = (\mathbf{H}, \mathbf{E})$ such that **H** is in $\mathbb{S}_\mu$ and **E** is in $\mathbb{S}_\varepsilon$. The scalar product in $\mathbb{H}_{\mu,\varepsilon}$ is set to be

$$(\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2) = 1/2[(\mathbf{H}_1, \mathbf{H}_2)_\mu + (\mathbf{E}_1, \mathbf{E}_2)_\varepsilon], \tag{11}$$

so the energy of an EM field $\boldsymbol{\Psi}$ is given by the square of its norm:

$$\mathscr{E}_{\boldsymbol{\Psi}} = \|\boldsymbol{\Psi}\|^2 = (\boldsymbol{\Psi}, \boldsymbol{\Psi}) = 1/2(\|\mathbf{H}\|_\mu{}^2 + \|\mathbf{E}\|_\varepsilon{}^2). \tag{12}$$

The operator $\mathbb{M}$ governing the dynamics of EM fields is a self-adjoint operator in the Hilbert space $\mathbb{H}_{\mu,\varepsilon}$; hence the solution to Eq. (5) is given by

$$\boldsymbol{\Psi}_t = \exp(it\mathbb{M})\boldsymbol{\Psi}_0, \tag{13}$$

so we have energy conservation:

$$\mathscr{E}_{\boldsymbol{\Psi}_t} = \|\boldsymbol{\Psi}_t\|^2 = \|\boldsymbol{\Psi}_0\|^2 = \mathscr{E}_{\boldsymbol{\Psi}_0}. \tag{14}$$

If $\boldsymbol{\Psi}_t$ is a solution of Eq. (5), it must also satisfy the second-order equation $(\partial^2/\partial t^2)\boldsymbol{\Psi}_t = -\mathbb{M}^2\boldsymbol{\Psi}_t$, so the magnetic and electric fields satisfy the second-order equations

$$\frac{\partial^2}{\partial t^2}\mathbf{H}_t = -\frac{1}{\mu}\nabla^{\times}\frac{1}{\varepsilon}\nabla^{\times}\mathbf{H}_t, \quad \mathbf{H}_t \text{ in } \mathbb{S}_\mu, \tag{15}$$

$$\frac{\partial^2}{\partial t^2}\mathbf{E}_t = -\frac{1}{\varepsilon}\nabla^{\times}\frac{1}{\mu}\nabla^{\times}\mathbf{E}_t, \quad \mathbf{E}_t \text{ in } \mathbb{S}_\varepsilon. \tag{16}$$

It is natural to introduce the Maxwell operators

$$\mathfrak{M}_{\mathbf{H}} = \frac{1}{\mu}\nabla^{\times}\frac{1}{\varepsilon}\nabla^{\times}, \qquad \mathfrak{M}_{\mathbf{E}} = \frac{1}{\varepsilon}\nabla^{\times}\frac{1}{\mu}\nabla^{\times}, \tag{17}$$

which are nonnegative self-adjoint operators on $\mathbb{S}_\mu$ and $\mathbb{S}_\varepsilon$, respectively. The two Maxwell operators are unitarily equivalent; more precisely,

$$\mathfrak{M}_{\mathbf{E}} = U\mathfrak{M}_{\mathbf{H}}U^*, \tag{18}$$

where U is the unitary operator mapping $\mathbb{S}_\mu$ onto $\mathbb{S}_\varepsilon$, given by

$$U\mathbf{H} = \frac{-i}{\varepsilon}\nabla^{\times}\frac{1}{\sqrt{\mathfrak{M}_H}}\mathbf{H}, \tag{19}$$

for **H** in the range of the operator $\sqrt{\mathfrak{M}_{\mathbf{H}}}$. Thus, if $\sigma(A)$ stands for the spectrum of the operator A, we have

$$\sigma(\mathbb{M}) = \sigma(\sqrt{\mathfrak{M}_{\mathbf{H}}}) \cup [-\sigma(\sqrt{\mathfrak{M}_{\mathbf{H}}})].$$

In particular, if a frequency ω belongs to the spectrum of the operator $\mathbb{M}$, then $-\omega$ must also be in this spectrum. Moreover, to obtain solutions of Eq. (5), we may set

$$\boldsymbol{\Psi}_{\pm,t} = [\exp(\pm it\sqrt{\mathfrak{M}_{\mathbf{H}}})\mathbf{H}_0,$$
$$\pm U\exp(\pm it\sqrt{\mathfrak{M}_{\mathbf{H}}})\mathbf{H}_0], \quad \mathbf{H}_0 \text{ in } \mathbb{S}_\mu. \tag{20}$$

Conversely, any solution of Eq. (5) can be written as a linear combination of at most four solutions of this form.

It follows that, to find all the eigenvalues and eigenmodes for $\mathbb{M}$, it is necessary and sufficient to find all the eigenvalues and eigenmodes for $\mathfrak{M}_{\mathbf{H}}$. For if ω^2 $(\omega > 0)$ and $\mathbf{H}_\omega$ are, respectively, an eigenvalue and the corresponding eigenmode of $\mathfrak{M}_{\mathbf{H}}$, i.e.,

$$\mathfrak{M}_{\mathbf{H}}\mathbf{H}_\omega = \omega^2\mathbf{H}_\omega, \quad \mathbf{H}_\omega \text{ in } \mathbb{S}_\mu, \tag{21}$$

then we have

$$U\mathbf{H}_\omega = \frac{-i}{\omega\varepsilon}\nabla^\times \mathbf{H}_\omega, \tag{22}$$

so Eqs. (6), (20), and (21) imply that

$$\mathbb{M}\mathbf{\Psi}_{\pm\omega} = \pm\omega\mathbf{\Psi}_{\pm\omega}, \tag{23}$$

where

$$\mathbf{\Psi}_{\pm\omega} = (\mathbf{H}_{\pm\omega}, \mathbf{E}_{\pm\omega}),$$

$$\mathbf{E}_{\pm\omega} = \pm U\mathbf{H}_{\pm\omega} = \pm\frac{-i}{\omega\varepsilon}\nabla^\times \mathbf{H}_{\pm\omega}. \tag{24}$$

Conversely, if $\mathbb{M}(\mathbf{H}_{\pm\omega}, \mathbf{E}_{\pm\omega}) = \pm\omega(\mathbf{H}_{\pm\omega}, \mathbf{E}_{\pm\omega})$, with $\omega > 0$ and nonzero vector $(\mathbf{H}_{\pm\omega}, \mathbf{E}_{\pm\omega})$ in $\mathbb{H}_{\mu,\varepsilon}$, then it follows that $\mathfrak{M}_\mathbf{H}\mathbf{H}_{\pm\omega} = \omega^2\mathbf{H}_{\pm\omega}$ and $\mathbf{E}_{\pm\omega} = \pm U\mathbf{H}_{\pm\omega} = \pm(-i/\omega\varepsilon)\nabla^\times\mathbf{H}_{\pm\omega}$.

An important conclusion of the previous considerations is that to study the spectral properties of the medium described by the Maxwell equations (1) and (2), it is sufficient to study the spectral properties of the self-adjoint operator $\mathfrak{M}_\mathbf{H}$ defined in Eqs. (17).

From here on, we shall work primarily with the operator $\mathfrak{M}_\mathbf{H}$. Below we assume that

$$\mu(\mathbf{x}) = 1, \tag{25}$$

and we simply write

$$\mathfrak{M} = \mathfrak{M}_\mathbf{H} = \nabla^\times\frac{1}{\varepsilon}\nabla^\times, \tag{26}$$

with the operator $\mathfrak{M}$ acting on the Hilbert space $\mathbb{S} = \mathbb{S}_1$, a closed subspace of the Hilbert space of square-integrable, vector-valued functions (L^2). We denote by $P_\mathbb{S}$ the orthogonal projection onto $\mathbb{S}$, so $I_\mathbb{S} = P_\mathbb{S}{}^*$ takes a function in $\mathbb{S}$ into the same function as an element of L^2.

We also need the unrestricted Maxwell operator:

$$M = M_\mathbf{H} = \nabla^\times\frac{1}{\varepsilon}\nabla^\times \quad \text{acting on} \quad L^2. \tag{27}$$

We have $\mathfrak{M} = P_\mathbb{S}MI_\mathbb{S}$; $\mathfrak{M}$ is the restriction of M to the orthogonal complement of its kernel. Note that $\mathfrak{M}$ and M have the same spectrum, so we can work with M to answer questions about the spectrum of $\mathfrak{M}$.

3. LOCALIZED WAVES

A localized EM wave is a finite-energy solution of Maxwell's equations with the property that almost all the wave's energy remains in a fixed bounded region of space at all times, e.g.,

$$\lim_{R\to\infty}\inf_t \frac{1}{\mathscr{E}}\int_{|\mathbf{x}|\leq R}\mathscr{E}(\mathbf{x}, t)\mathrm{d}\mathbf{x} = 1. \tag{28}$$

In spectral terms the localization phenomenon can be characterized as follows. The dielectric medium in which the waves propagate is fully described by the linear Maxwell operator $\mathbb{M}$ defined by Eqs. (6). The basic spectral attributes of $\mathbb{M}$ are its eigenvalues ω and eigenmodes $\mathbf{\Psi}_\omega(\mathbf{x}) = [\mathbf{H}_\omega(\mathbf{x}), \mathbf{E}_\omega(\mathbf{x})]$, i.e., $\mathbb{M}\mathbf{\Psi}_\omega = \omega\mathbf{\Psi}_\omega$. If for some ω the eigenmode $\mathbf{\Psi}_\omega$ has finite energy, i.e.,

$$\int_{\mathbb{R}^3} {}^1\!/\!{}_2[\mu(\mathbf{x})|\mathbf{H}_\omega(\mathbf{x})|^2 + \varepsilon(\mathbf{x})|\mathbf{E}_\omega(\mathbf{x})|^2]\mathrm{d}\mathbf{x} < \infty, \tag{29}$$

we call it a localized eigenmode. Observe that, for extended eigenmodes, which are similar to plane EM waves in a homogeneous medium, condition (29) is clearly violated, since the total energy of the wave will be infinite.

It turns out that in many cases inequality (29) follows from a stronger property: exponential decay of $\mathbf{\Psi}_\omega(\mathbf{x})$ as $|\mathbf{x}| \to \infty$ (Refs. 8–11), i.e.,

$$|\mathbf{\Psi}_\omega(\mathbf{x})| \leq \text{const.}\exp\left(-\frac{|\mathbf{x}|}{L_\omega}\right). \tag{30}$$

The smallest L_ω for which inequality (30) still holds is called the radius of localization of the eigenmode $\mathbf{\Psi}_\omega$.

Having a localized eigenmode $\mathbf{\Psi}_\omega$, we can easily construct a localized EM wave in the sense of Eq. (28). Namely, observe that $\mathbf{\Psi}_{\omega,t} = \exp(it\omega)\mathbf{\Psi}_\omega$ is a localized EM wave, i.e., it satisfies Eqs. (5) and (28). Note, in addition, that in this case $-\omega$ is also an eigenvalue of $\mathbb{M}$ with eigenmode $\bar{\mathbf{\Psi}}_\omega$, so $\bar{\mathbf{\Psi}}_{\omega,t} = \exp(it\omega)\bar{\mathbf{\Psi}}_\omega$ is also a localized wave, since if $\mathbf{J}$ denotes the antiunitary involution corresponding to complex conjugation on $\mathbb{H}_{\mu,\epsilon}$, i.e., $\mathbf{J}\mathbf{\Psi} = \bar{\mathbf{\Psi}}$, we have $\mathbf{J}\mathbb{M}\mathbf{J} = -\mathbb{M}$. It also follows that the spectrum of $\mathbb{M}$ is symmetric, i.e., $\sigma(\mathbb{M}) = -\sigma(\mathbb{M})$, with $\mathbf{J}\mathbb{M}_+\mathbf{J} = \mathbb{M}_-$, where $\mathbb{M}_\pm$ is the positive and negative parts of $\mathbb{M}$. Furthermore, using the same arguments, we find that any linear combinations of localized eigenmodes of $\mathbb{M}$ give rise to localized EM waves.

4. PERIODIC DIELECTRIC MEDIA

There is a well-known relationship between the localization of a wave propagating in an inhomogeneous medium and the rise of spectral gaps (stop bands) in a periodic medium. Both phenomena are caused by multiple scattering and destructive wave interference. To obtain wave localization, we start with a periodic dielectric medium with a spectral gap and then perturb it by either a single defect or a random array of defects.

In this section we describe the basic properties of the periodic dielectric medium (photonic crystal). For many dielectric materials of interest the magnetic permeability is close to unity, so we shall assume from now on that Eq. (25) holds.

The periodic dielectric medium is described by a periodic dielectric function $\varepsilon_0(\mathbf{x})$. If $\mathbb{L}$ is the lattice of periods, we have

$$\varepsilon_0(\mathbf{x} + \mathbf{n}) = \varepsilon_0(\mathbf{x}) \quad \text{for any } \mathbf{n} \text{ from } \mathbb{L}.$$

We assume for simplicity that the primitive cell (PC) of $\mathbb{L}$ is a cube.

The corresponding Maxwell operator takes the form

$$\mathfrak{M}_0\mathbf{\Psi}(\mathbf{x}) = \nabla\times\frac{1}{\varepsilon_0(\mathbf{x})}\nabla\times\mathbf{\Psi}(\mathbf{x}), \tag{31}$$

where $\nabla\cdot\mathbf{\Psi}(\mathbf{x}) = 0$. In view of the periodicity, the Floquet–Bloch theory can be applied, so the spectrum of $\mathfrak{M}_0$ has band structure, and the eigenmodes have Bloch form:

$$\mathfrak{M}_0\mathbf{\Psi}_{\omega_n(\mathbf{k})}(\mathbf{x}) = \omega_n{}^2(\mathbf{k})\mathbf{\Psi}_{\omega_n(\mathbf{k})}(\mathbf{x}), \quad n = 1, 2, ...,$$

$$\mathbf{\Psi}_{\omega_n(\mathbf{k})}(\mathbf{x}) = \exp(i\mathbf{k}\cdot\mathbf{x})\mathbf{\Phi}_{\omega_n(\mathbf{k})}(\mathbf{x}),$$

$$\mathbf{\Phi}_{\omega_n(\mathbf{k})}(\mathbf{x} + \mathbf{n}) = \mathbf{\Phi}_{\omega_n(\mathbf{k})}(\mathbf{x}), \quad \mathbf{n} \text{ in } \mathbb{L},$$

where the quasi-momentum $\mathbf{k}$ belongs to the Brillouin zone (BZ), the PC of the lattice reciprocal belongs to $\mathbb{L}$, and n is the index of a zone. The frequency function $\omega_n(\mathbf{k})$ is called the dispersion relation of the nth zone, and $\mathbf{\Psi}_{\omega_{n(\mathbf{k})}}$ is the eigenmode of the nth zone. We naturally order the frequencies $\omega_n(\mathbf{k})$ such that

$$\omega_1(\mathbf{k}) \leqslant \omega_2(\mathbf{k}) \leqslant \ldots \leqslant \omega_n(\mathbf{k}) \leqslant \ldots .$$

If I_n denotes the interval of values of the function $\omega_n{}^2(\mathbf{k})$, where $\mathbf{k}$ runs the BZ, then the spectrum $\sigma(\mathfrak{M}_0)$ of the periodic operator $\mathfrak{M}_0$ coincides with the union of these intervals, i.e.,

$$\sigma(\mathfrak{M}_0) = \bigcup_{n=1,2,\ldots} I_n . \tag{32}$$

It can easily be verified that the spectrum $\sigma(\mathfrak{M}_0)$ is a closed subset of the positive semiaxis $[0, \infty)$ and that 0 is in its spectrum. It may happen that the intervals I_n do not cover all the semiaxis $[0, \infty)$, so there will be an interval $(\omega_a{}^2, \omega_b{}^2)$, called a spectral gap, that does not belong to the spectrum of $\mathfrak{M}_0$. The physical significance of the existence of a gap lies in the fact that a wave with frequency in the gap cannot propagate in the medium.

The zone structure of the spectrum is a generic property owing to the periodicity. But the existence of spectral gaps is not a generic property. It depends in a subtle way on the geometry and the distribution of the dielectric materials in the periodic medium. It is quite evident that high-contrast periodic media favor the rise of spectral gaps, but this observation alone is far from sufficient to establish the existence of a gap. In particular, the mathematical clarification of the very concept of high contrast is, we believe, a nontrivial and important problem.

The existence of gaps for some periodic dielectric and acoustic media has been rigorously proved.[19,20] The cited papers also give a constructive approach to what may be called a high-contrast medium.

Since our interest is primarily in the phenomenon of localization, we simply assume from here on that the background periodic medium has at least one spectral gap.

Assumption 1 (a gap in the spectrum). There exist frequencies $0 < \omega_a < \omega_b$ such that $\omega_a{}^2$ and $\omega_b{}^2$ are in the spectrum of $\sigma(\mathfrak{M}_0)$, and the interval $(\omega_a{}^2, \omega_b{}^2)$ is a spectral gap, i.e., $(\omega_a{}^2, \omega_b{}^2)$ has no intersection with $\sigma(\mathfrak{M}_0)$.

Why does a wave with frequency in a spectral gap not propagate in the medium? In fact, that kind of medium response is a very general property of a linear conservative medium occupying infinite space. If we excite the medium locally at a frequency ω that is not an eigenfrequency, the amplitude of the forced oscillations will die out exponentially away from the location of the excitation.

The quantity that describes the amplitude of these forced oscillations is the corresponding Green's function $G_0(\omega; \mathbf{x}, \mathbf{y})$, defined by

$$G_0(\omega; \mathbf{x}, \mathbf{y}) = (\mathfrak{M}_0 - \omega^2 I)^{-1}(\mathbf{x}, \mathbf{y}), \tag{33}$$

where $\mathbf{y}$ is the location of the source of the harmonic excitation at frequency ω. The Green's function $G_0(\omega; \mathbf{x}, \mathbf{y})$ gives the amplitude of the medium response at an observation point $\mathbf{x}$. If the frequency of the forced oscillation ω falls in a spectral gap, we expect $G_0(\omega; \mathbf{x}, \mathbf{y})$ to decay exponentially; i.e., for some constants C_ω and L_ω we have

$$|G_0(\omega; \mathbf{x}, \mathbf{y})| \leqslant C_\omega \exp\left(-\frac{|\mathbf{x} - \mathbf{y}|}{L_\omega}\right), \quad |\mathbf{x} - \mathbf{y}| \to \infty. \tag{34}$$

In this case it is easy to see that we have a nonpropagation regime. Indeed, let us take a sphere centered at $\mathbf{y}$ of a large radius R. Then, in view of relation (34), the energy flow through the sphere can be estimated as follows:

$$\pi R^2 |G_0(\omega; \mathbf{x}, \mathbf{y})|^2|_{|\mathbf{x}-\mathbf{y}|=R} \leqslant \pi R^2 C_\omega \times \exp\left(-\frac{2R}{L_\omega}\right) \xrightarrow[R\to\infty]{} 0. \tag{35}$$

This lack of energy flow can be interpreted as a nonpropagation regime. In contrast, if the frequency ω is in the spectrum, then $|G_0(\omega; \mathbf{x}, \mathbf{y})|$ is proportional to $|\mathbf{x} - \mathbf{y}|^{-1}$, which results in a nonzero energy flow.

The rate of the exponential decay of the Green's function $G_0(\omega; \mathbf{x}, \mathbf{y})$, for ω in a spectral gap, can be estimated rigorously as follows. Let

$$\chi(\mathbf{y}) = \begin{cases} 1 & \text{if } \mathbf{y} \text{ is in PC,} \\ 0 & \text{otherwise} \end{cases} \tag{36}$$

be the characteristic function of the PC of our lattice of periods, and let

$$\chi_{\mathbf{x}}(\mathbf{y}) = \chi_{\mathbf{x}}(\mathbf{y} - \mathbf{x}). \tag{37}$$

Consider the operators

$$G(\omega) = (\mathfrak{M} - \omega^2 I)^{-1}, \quad G_0(\omega) = (\mathfrak{M}_0 - \omega^2 I)^{-1} \tag{38}$$

acting in the Hilbert space $\mathbb{S}$. We recall that the norm of an operator A is defined by

$$\|A\| = \sup_{\mathbf{\Psi} \neq 0} \frac{\|A\mathbf{\Psi}\|}{\|\mathbf{\Psi}\|}. \tag{39}$$

The following statement holds[9]:

Theorem 1. Let $\varepsilon(\mathbf{x})$ satisfy relation (3), let $\mathfrak{M}$ be an operator of the form given in Eq. (26) having a spectral gap, and let ω^2 fall in this gap. Then there exists a finite absolute constant C_0 such that for all $\mathbf{x}$ and $\mathbf{y}$ the following inequality holds:

$$\|\chi_{\mathbf{x}} G(\omega) \chi_{\mathbf{y}}\| \leqslant \frac{C_0}{\eta} \exp\left(-\frac{|\mathbf{x} - \mathbf{y}|}{L_\omega}\right), \tag{40}$$

where

$$L_\omega = \frac{4(2\varepsilon_-{}^{-1} + \omega^2 + \eta)}{\eta}, \tag{41}$$

with η being the distance from ω^2 to the edges of the spectral gap. Moreover,

$$\|\chi_{\mathbf{x}} \nabla^\times G(\omega) \chi_{\mathbf{y}}\| \leqslant \frac{C(1 + \omega^2)}{\eta} \exp\left(-\frac{|\mathbf{x} - \mathbf{y}|}{L_\omega}\right), \tag{42}$$

where C is a constant depending on ε_- and ε_+. In particular, inequalities (40) and (42) hold for $G_0(\omega)$.

Another important property of the periodic medium is a certain regularity of the gap edges, which we define as follows. Let us consider, for instance, the edge $\omega_a{}^2$. Let $n_1 \leqslant n \leqslant n_0$, be the indices of all the bands such that $\omega_a{}^2$ is their right-hand edge, i.e.,

$$\max_{\mathbf{k}\ \text{in BZ}} \omega_n{}^2(\mathbf{k}) = \omega_a{}^2.$$

A natural regularity condition at $\omega_a{}^2$ is a nondegeneracy condition: For each $n_1 \leqslant n \leqslant n_0$ we have $\omega_n{}^2(\mathbf{k}) = \omega_a{}^2$ for only finitely many $\mathbf{k}$, say, $k_{n,1}, \ldots, k_{n,s_n}$; and for any $i = 1, ..., s_n$ we have

$$a - \omega_n{}^2(\mathbf{k}) \geqslant c_i|\mathbf{k} - \mathbf{k}_{n,i}|^2 \tag{43}$$

for small $|\mathbf{k} - \mathbf{k}_{n,i}|$, with $c_i > 0$.

This nondegeneracy condition is a common assumption in the physical literature.[14] It was also used in the study of Lifshitz tails in spectral gaps of periodic Schrödinger operators.[38] In fact, the verification of the regularity of a gap edge is not a simple matter even for Schrödinger operators. For our purposes a weaker notion of regularity of a spectral edge suffices. To motivate the condition, note that for $\omega_a{}^2 < \omega^2 < \omega_b{}^2$ we always have[11] $\mathrm{Tr}\{[\chi I_{\mathrm{S}}|G_0(\omega^2)|P_{\mathrm{S}}\chi]^2\} < \infty$, where Tr A denotes the trace of the operator A. Our definition of regularity at an edge of the gap requires that this quantity remain finite as we approach this edge. Let us introduce the quantities

$$\xi(\omega_a{}^2) = \limsup_{n\downarrow 0} \mathrm{Tr}\{[\chi I_{\mathrm{S}}|G_0(\omega_a{}^2 + \eta)|P_{\mathrm{S}}\chi]^2\}, \tag{44}$$

$$\xi(\omega_b{}^2) = \limsup_{n\downarrow 0} \mathrm{Tr}\{[\chi I_{\mathrm{S}}|G_0(\omega_b{}^2 - \eta)|P_{\mathrm{S}}\chi]^2\}, \tag{45}$$

where χ is as in Eq. (36). These quantities, in terms of which we shall give a rigorous definition of regularity of an edge, appear in estimates of the number of eigenvalues that a defect can create in the gap. Note that with the following definition nondegenerate edges will always be regular.[11]

Definition 1 (regularity at an edge). The left-hand edge $\omega_a{}^2$ of the gap $(\omega_a{}^2, \omega_b{}^2)$ in the spectrum of the periodic operator $\mathfrak{M}_0$ is regular if $\xi(\omega_a{}^2) < \infty$. Similarly, the right-hand edge $\omega_b{}^2$ is regular if $\xi(\omega_b{}^2) < \infty$.

5. MIDGAP DEFECT EIGENMODES

It is a well-known fact in solid-state physics that, in three dimensions, a potential well of depth U and of radius a generates an exponentially localized state if

$$a^2 U > \frac{\pi^2 \hbar^2}{8m}, \tag{46}$$

where m is the mass of the quantum particle.[39] The question is whether an analogous condition can be found for classical EM waves.

In spite of the fundamental similarity between the creation of localized eigenmodes for classical and electron waves, there are some important differences. First, for the electron it suffices to perturb locally a homogeneous medium (i.e., a constant potential) to generate a localized eigenmode. For classical waves a local perturbation of a homogeneous medium [i.e., $\varepsilon_0(x)$ is constant] cannot generate a localized eigenmode. This can easily be seen from the consideration of a one-dimensional model. Indeed, in that case we consider the eigenvalue problem $-\{1/[\varepsilon(x)]u'(x)\}' = \xi u'(x)$, $-\infty < x < \infty$, where $\varepsilon(x) =$ const. if $|x| > R$ for some R and ξ is a positive number. It is clear that this equation cannot have square-integrable solutions. Since, in general, the one-dimensional case is the most favorable for localization, we should not expect localization in analogous circumstances in the multidimensional case.

The reason for this difference between classical waves and electrons can be explained as follows. The motion of the electron in a homogeneous medium is described by the Schrödinger operator $H_0 = -\Delta + V_0$ with a constant potential $V_0(x) \equiv v_0$. Clearly the spectrum $\sigma(H_0)$ of the operator H_0 is the interval $[v_0, \infty)$, so we may consider the infinite interval $(-\infty, v_0)$ as a gap in the spectrum of the operator H_0. Note that the edge v_0 of the gap depends on the homogeneous medium. Hence, if we perturb this homogeneous medium by a defect, say, a potential well, the spectrum can expand in the interior of the gap $(-\infty, v_0)$, and if this happens the corresponding eigenmodes will be exponentially localized.

For EM waves in a homogeneous medium that are described by the Maxwell operator $\mathfrak{M}$ with constant $\varepsilon(\mathbf{x})$, we always have $\sigma(\mathfrak{M}) = [0, \infty)$, so, as for Schrödinger operators, we may consider the infinite interval $(-\infty, 0)$ as a gap in the spectrum. But for classical waves the bottom 0 of the spectrum does not depend at all on the coefficient $\varepsilon(\mathbf{x})$ of the medium. This is why a local perturbation of any medium by a defect does not expand the spectrum into the gap $(-\infty, 0)$, as we saw in the one-dimensional model.

Thus, to employ a mechanism for localization of EM waves similar to the one for electronic localization, we have to start with a medium described by a coefficient $\varepsilon_0(\mathbf{x})$ such that the corresponding Maxwell operator has a gap inside its spectrum, and the edges of the gaps must depend on the medium, i.e., on the coefficient $\varepsilon_0(\mathbf{x})$. Such media can be perturbed locally by a defect, giving rise to exponentially localized eigenmodes with corresponding eigenvalues in the interior of the gaps.

A defect is a perturbation of a given medium in a finite domain (see Fig. 1). Defects in the medium generate localized waves by creating localized eigenmodes of the operator $\mathfrak{M}$.

Let us say that the medium described by $\varepsilon(\mathbf{x})$ is obtained from the background medium by the insertion of a defect, if $\varepsilon(\mathbf{x})$ and $\varepsilon_0(\mathbf{x})$ differ only in a bounded domain Λ. In this case we shall say that $\varepsilon(\mathbf{x})$ and $\varepsilon_0(\mathbf{x})$ differ by a defect. A simple way to tailor these defects is as follows. Let Λ be a bounded domain containing the origin 0. Typically, we take Λ to be the cube of side 1 centered at the origin, or the unit ball centered at the origin. Let us set $\Lambda_\ell = \ell\Lambda$ for $\ell > 0$, so Λ_ℓ is the cube of side ℓ centered at the origin, etc. We insert a defect by changing the value of $\varepsilon_0(\mathbf{x})$ inside Λ_ℓ to a given constant $\epsilon > 0$, i.e.,

$$\varepsilon(\mathbf{x}) = \varepsilon_{\varepsilon,\ell}(\mathbf{x}) = \begin{cases} \epsilon & \text{if } \mathbf{x} \text{ is in } \Lambda_\ell, \\ \varepsilon_0(\mathbf{x}) & \text{otherwise.} \end{cases} \tag{47}$$

We recall that the essential spectrum $\sigma_{\mathrm{ess}}(\mathfrak{M})$ of an operator $\mathfrak{M}$ consists of all the points of its spectrum $\sigma(\mathfrak{M})$

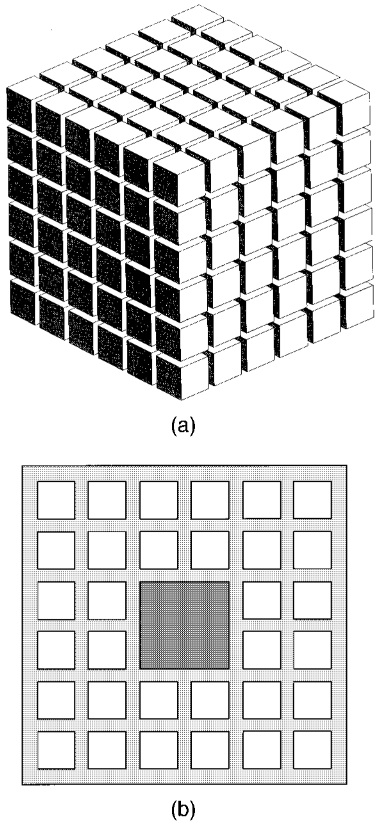

Fig. 1. (a) Slab of a photonic crystal with an interior defect of higher dielectric constant. (b) The defect is shown on the cross section of the slab as a darker square.

that are not isolated eigenvalues with finite multiplicity. Essential spectra are not changed by defects.[10]

Theorem 2 (stability of essential spectrum). Assume that $\varepsilon(\mathbf{x})$ and $\varepsilon_0(\mathbf{x})$ differ by a defect. Then

$$\sigma_{\text{ess}}(\mathfrak{M}) = \sigma_{\text{ess}}(\mathfrak{M}_0).$$

If $(\omega_a^2,\ \omega_b^2)$ is a gap in the spectrum of $\mathfrak{M}_0$, the spectrum of $\mathfrak{M}$ in $(\omega_a^2,\ \omega_b^2)$ consists at most of isolated eigenvalues with finite multiplicity, with the corresponding eigenmodes decaying exponentially fast, with a rate depending on the distance from the eigenvalue to the edges of the gap.

Theorem 2 has been proved rigorously,[10] and it says that a finite defect can create only isolated eigenvalues in the gap $(\omega_a^2,\ \omega_b^2)$ with exponentially decaying eigenmodes. Here we consider just the basic arguments. Recall that the spectrum $\sigma(\mathfrak{M})$ of a self-adjoint operator $\mathfrak{M}$ can be defined as the set of real numbers ω^2 such that for any positive δ we can find a square integrable $\boldsymbol{\Psi}$ such that

$$\|(\mathfrak{M} - \omega^2 I)\boldsymbol{\Psi}\| \leqslant \delta\|\boldsymbol{\Psi}\|, \quad \|\boldsymbol{\Psi}\|^2 = \int |\boldsymbol{\Psi}(\mathbf{x})|^2 \mathrm{d}\mathbf{x}, \tag{48}$$

where I is the identity operator. In particular, for the periodic operator $\mathfrak{M}_0$ we can always choose $\boldsymbol{\Psi}$ vanishing outside a ball B_δ of sufficiently large radius such that B_δ does not intersect the domain Λ of the defect and relations (48) hold for $\mathfrak{M}_0$. But for this $\boldsymbol{\Psi}$ we clearly have $\mathfrak{M}\boldsymbol{\Psi} = \mathfrak{M}_0\boldsymbol{\Psi}$, so relations (48) also hold for $\mathfrak{M}$; hence we have $\sigma_{\text{ess}}(\mathfrak{M}) \supset \sigma_{\text{ess}}(\mathfrak{M}_0)$. Now, let $\boldsymbol{\Phi}_\omega$ be an eigenmode (extended or localized) of $\mathfrak{M}$ corresponding to the eigenvalue ω^2 in the gap $(\omega_a^2,\ \omega_b^2)$. Let us show that $\boldsymbol{\Phi}_\omega$ must be localized and, in addition, exponentially decaying away from the location of the defect Λ. Indeed, from

$$\mathfrak{M}\boldsymbol{\Phi}_\omega = \omega^2\boldsymbol{\Phi}_\omega \tag{49}$$

we obtain

$$\boldsymbol{\Phi}_\omega(\mathbf{x}) = -\int G_0(\omega;\ \mathbf{x},\ \mathbf{y})[(\mathfrak{M} - \mathfrak{M}_0)\boldsymbol{\Phi}_\omega](\mathbf{y})\mathrm{d}\mathbf{y}, \tag{50}$$

where $G_0(\omega;\ \mathbf{x},\ \mathbf{y}) = (\mathfrak{M}_0 - \omega I)^{-1}(\mathbf{x},\ \mathbf{y})$ is the Green's function of the periodic operator $\mathfrak{M}_0$. Observe that, since $\mathfrak{M}$ and $\mathfrak{M}_0$ differ only on the domain Λ, Eq. (50) implies that

$$\boldsymbol{\Phi}_\omega(\mathbf{x}) = -\int_\Lambda G_0(\omega;\ \mathbf{x},\ \mathbf{y})[(\mathfrak{M} - \mathfrak{M}_0)\boldsymbol{\Phi}_\omega](\mathbf{y})\mathrm{d}\mathbf{y}. \tag{51}$$

Recall now that, since the eigenvalue ω^2 is in the gap of the operator $\mathfrak{M}_0$, the Green's function $G_0(\omega;\ \mathbf{x},\ \mathbf{y})$ must decay exponentially, i.e.,

$$|G_0(\omega;\ \mathbf{x},\ \mathbf{y})| \leqslant C_1 \exp(-C_2|\mathbf{x} - \mathbf{y}|). \tag{52}$$

It follows from general considerations that the eigenmode $\boldsymbol{\Phi}_\omega(\mathbf{x})$ must be effectively bounded in any finite domain regardless of whether it is extended or localized. This last comment, Eq. (51), and the inequality (52), imply that $\boldsymbol{\Phi}_\omega(\mathbf{x})$ must be an exponentially decaying function and hence that all the eigenmodes of $\mathfrak{M}$ with corresponding eigenvalues in the gap must be exponentially localized.

Theorem 2 states that, if we have any spectrum generated by a defect in the spectral gap $(\omega_a^2,\ \omega_b^2)$ of the original periodic operator $\mathfrak{M}_0$, it must be associated with exponentially decaying eigenmodes. But the question whether the defect creates any spectrum at all in the spectral gap of $\mathfrak{M}_0$ remains. Using simple space scaling, we observe that a sufficient condition for the existence of localized eigenmodes generated by the defect should have the form

$$\ell^2\epsilon > C(\Lambda,\ \omega_a,\ \omega_b), \tag{53}$$

where the constant $C(\Lambda,\ \omega_a,\ \omega_b)$, in general, depends on the shape Λ of the defect and on the location of the gap $(\omega_a^2,\ \omega_b^2)$.

When Λ is a cube of side 1 and the defect is as in Eq. (47), a more delicate and rigorous analysis[10] leads to the following sufficient condition for the existence of localized eigenmodes. In particular, it gives inequality (53) with

$$C(\Lambda,\ \omega_a,\ \omega_b) = \frac{158(\omega_a^2 + \omega_b^2)}{(\omega_b^2 - \omega_a^2)^2}. \tag{54}$$

Theorem 3 (creation of defect eigenmodes). Let $(\omega_a^2,\ \omega_b^2)$ be a gap in the spectrum of $\mathfrak{M}_0$; select τ in $(\omega_a^2,\ \omega_b^2)$; and pick $0 < \gamma < 1$ such that the interval $[\tau(1 - \gamma),\ \tau(1 + \gamma)]$ is contained in the gap, i.e., $[\tau(1 - \gamma),\ \tau(1 + \gamma)] \subset (\omega_a^2,\ \omega_b^2)$. If $\varepsilon(x) \equiv \epsilon$ in a cube of side ℓ, with

$$\ell^2\epsilon > \frac{79}{\tau\gamma^2}, \tag{55}$$

the corresponding operator $\mathfrak{M}$ has at least one defect eigenmode with corresponding eigenvalue inside the interval $[\tau(1-\gamma),\ \tau(1+\gamma)]$.

Note that condition (55) is analogous to condition (46).

We shall call a defect positive if $\varepsilon(\mathbf{x}) \leqslant \varepsilon_0(\mathbf{x})$, in which case we have $\mathfrak{M} - \mathfrak{M}_0 \geqslant 0$. Similarly, a defect is negative if $\varepsilon(\mathbf{x}) \geqslant \varepsilon_0(\mathbf{x})$, so $\mathfrak{M} - \mathfrak{M}_0 \leqslant 0$. If the defect is either positive or negative, we can say more about how the eigenvalues are distributed in the gap. To do so, we use a modified Birman–Schwinger method to reduce the problem to the study of the eigenvalues of a compact operator (in fact, a Hilbert–Schmidt operator).[11] The usual Birman–Schwinger method[40,41] cannot be used directly, as $\mathfrak{M} - \mathfrak{M}_0$ is not relatively compact with respect to $\mathfrak{M}_0$. The solution is to use the resolvents: We set $H = (\mathfrak{M} + I)^{-1}$ and $H_0 = (\mathfrak{M}_0 + I)^{-1}$, prove that $V = H - H_0$ is a Hilbert–Schmidt operator in dimension 3 or less, and then use the Birman–Schwinger method for $H = H_0 + V$. This method gives equations for the defect eigenmodes and corresponding midgap eigenvalues in terms of the spectral attributes of an auxiliary Hilbert–Schmidt operator.[11]

Let $0 < \omega_a < \omega_b$ such that $({\omega_a}^2,\ {\omega_b}^2)$ is a gap in the spectrum of the operator $\mathfrak{M}_0$, and let us insert a negative defect such that $\varepsilon(\mathbf{x}) \geqslant \varepsilon_0(\mathbf{x})$ and $V \geqslant 0$. Let us consider the eigenvalue problem for the operator $\mathfrak{M}$ in the gap:

$$\mathfrak{M}\boldsymbol{\Psi} = \omega^2\boldsymbol{\Psi}, \qquad \omega \text{ is in } (\omega_a,\ \omega_b). \tag{56}$$

This is clearly equivalent to the eigenvalue problem

$$H\boldsymbol{\Psi} = (\mathfrak{M} + I)^{-1}\boldsymbol{\Psi} = (\omega^2 + 1)^{-1}\boldsymbol{\Psi}, \qquad \omega \text{ is in } (\omega_a,\ \omega_b). \tag{57}$$

In contrast, the eigenvalue problem

$$H\boldsymbol{\Psi} = H_0\boldsymbol{\Psi} + V\boldsymbol{\Psi} = \xi\boldsymbol{\Psi}, \qquad \xi \text{ is not in } \sigma(H_0), \tag{58}$$

can be rewritten as

$$\boldsymbol{\Psi} = -R_0(\xi)V\boldsymbol{\Psi}, \qquad R_0(\xi) = (H_0 - \xi I)^{-1}. \tag{59}$$

Setting

$$\mathscr{R}(\xi) = -\sqrt{V}R_0(\xi)\sqrt{V}, \tag{60}$$

we obtain the eigenvalue problem

$$\mathscr{R}(\xi)\boldsymbol{\Phi} = \boldsymbol{\Phi}, \qquad \boldsymbol{\Phi} = \sqrt{V}\boldsymbol{\Psi}, \tag{61}$$

which is equivalent to the eigenvalue problem [Eq. (58)]. The Birman–Schwinger operator $\mathscr{R}(\xi)$ is a self-adjoint Hilbert–Schmidt operator (V is a Hilbert–Schmidt operator). The original eigenvalue problem [Eq. (56)] for $\mathfrak{M}$ can now be rewritten as follows:

$$\mathscr{S}(\omega)\boldsymbol{\Phi} = \boldsymbol{\Phi}, \qquad \boldsymbol{\Phi} = \sqrt{V}\boldsymbol{\Psi}, \qquad \omega \text{ is in } (\omega_a,\ \omega_b), \tag{62}$$

and $\mathscr{S}(\omega)$ is the self-adjoint Hilbert–Schmidt operator given by

$$\mathscr{S}(\omega) = \mathscr{R}[(\omega^2 + 1)^{-1}] = (\omega^2 + 1)\sqrt{V}\,\frac{\mathfrak{M}_0 + I}{\mathfrak{M}_0 - \omega^2 I}\,\sqrt{V}. \tag{63}$$

In the case of a positive defect such that $\varepsilon(\mathbf{x}) \leqslant \varepsilon_0(\mathbf{x})$ and $V \leqslant 0$, the analog of Eqs. (62) and (63) takes the form

$$\mathscr{S}(\omega)\boldsymbol{\Phi} = -\boldsymbol{\Phi}, \qquad \boldsymbol{\Phi} = \sqrt{-V}\boldsymbol{\Psi}, \qquad \omega \text{ is in } (\omega_a,\ \omega_b), \tag{64}$$

with

$$\mathscr{S}(\omega) = (\omega^2 + 1)\sqrt{-V}\,\frac{\mathfrak{M}_0 + I}{\mathfrak{M}_0 - \omega^2 I}\,\sqrt{-V}. \tag{65}$$

Observe now that, since $(\mathfrak{M}_0 - \omega^2 I)^{-1}$ is a monotonically increasing, norm-continuous operator function of ω^2 in $({\omega_a}^2,\ {\omega_b}^2)$, the operator $\mathscr{S}(\omega)$ is also a monotonically increasing, norm-continuous function of ω in $(\omega_a,\ \omega_b)$ for both negative and positive defects. Since $\mathscr{S}(\omega)$ is a self-adjoint Hilbert–Schmidt operator, its spectrum consists of eigenvalues of finite multiplicity, with 0 being the only possible point of accumulation. Let ${\rho_1}^+(\omega) \geqslant {\rho_2}^+(\omega) \geqslant \ldots \geqslant 0$ and ${\rho_1}^-(\omega) \leqslant {\rho_2}^-(\omega) \leqslant \ldots \leqslant 0$ be the sequences of the positive and the negative eigenvalues of the operator $\mathscr{S}(\omega)$, respectively, repeated according to their multiplicity. If we have a finite number of either positive or negative eigenvalues, we complete the sequence by assigning the value 0. The functions ${\rho_n}^{\pm}(\omega)$ are monotonically increasing and continuous in ω in the interval $(\omega_a,\ \omega_b)$ (see Fig. 2).

We have[11]

Theorem 4 (equations for defect eigenvalues). Let $\varepsilon_0(\mathbf{x})$ be a periodic function satisfying relation (3), with the interval $({\omega_a}^2,\ {\omega_b}^2)$ being a gap in the spectrum of $\mathfrak{M}_0$, and let $\varepsilon(\mathbf{x})$ be obtained from $\varepsilon_0(\mathbf{x})$ by the insertion of a defect. Then

(i) If the defect is negative, the only possible point of accumulation of the defect eigenvalues of $\mathfrak{M}$ in $({\omega_a}^2,\ {\omega_b}^2)$ is ${\omega_b}^2$. In this case the frequencies $\omega_1 \leqslant \omega_2 \leqslant \ldots$ in $(\omega_a,\ \omega_b)$, such that ${\omega_1}^2 \leqslant {\omega_2}^2 \leqslant \ldots$ are the eigenvalues of the operator $\mathfrak{M}$ in the gap $({\omega_a}^2,\ {\omega_b}^2)$, coincide with the set of the solutions of the equations

$${\rho_n}^+(\omega) = 1, \qquad n = 1, 2, \ldots, \tag{66}$$

where ${\rho_n}^+(\omega)$ are the positive eigenvalues of the operator $\mathscr{S}(\omega)$ defined by Eq. (63). Moreover, if φ_i is an eigenmode of the operator $\mathscr{S}(\omega_i)$ with eigenvalue ${\rho_{n_i}}^+(\omega_i) = 1$, then

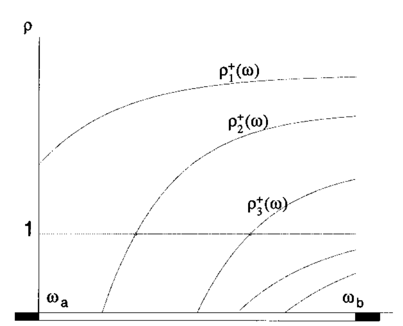

Fig. 2. Equations for the eigenvalues for the defect eigenmodes take the form ${\rho_n}^+(\omega) = 1$, where the functions ${\rho_n}^+(\omega)$ are the eigenvalues of an auxiliary compact operator depending on the spectral parameter ω.

$$\mathbf{\Psi}_i = \frac{\mathfrak{M}_0 + I}{\mathfrak{M}_0 - \omega_i^2 I} \sqrt{V}\mathbf{\Phi}_i \tag{67}$$

is an exponentially localized eigenmode of the operator $\mathfrak{M}$ with eigenvalue ω_i^2.

(ii) If the defect is positive, the only possible point of accumulation of the defect eigenvalues of $\mathfrak{M}$ in (ω_a^2, ω_b^2) is ω_a^2. In this case the frequencies $\omega_1 \geqslant \omega_2 \geqslant \ldots$ in (ω_a, ω_b), such that $\omega_1^2 \geqslant \omega_2^2 \geqslant \ldots$ are the eigenvalues of the operator $\mathfrak{M}$ in the gap (ω_a^2, ω_b^2), coincide with the set of the solutions of the equations

$$\rho_n^-(\omega) = -1, \qquad n = 1, 2, ..., \tag{68}$$

where $\rho_n^-(\omega)$ are the negative eigenvalues of the operator $\mathscr{S}(\omega)$ defined by Eq. (65). Moreover, if φ_i is an eigenmode of the operator $\mathscr{S}(\omega_i)$ with eigenvalue $\rho_{n_i}^-(\omega_i) = -1$, then

$$\mathbf{\Psi}_i = -\frac{\mathfrak{M}_0 + I}{\mathfrak{M}_0 - \omega_i^2 I} \sqrt{-V}\mathbf{\Phi}_i \tag{69}$$

is an exponentially localized eigenmode of the operator $\mathfrak{M}$ with eigenvalue ω_i^2.

It follows from Theorem 3 that, for defects as in Eq. (47), which satisfy condition (55) with $\tau = (\omega_a^2 + \omega_b^2)/2$ and $\gamma = (\omega_b^2 - \omega_a^2)/(\omega_a^2 + \omega_b^2)$, defect eigenmodes and midgap eigenvalues always exist, so in this case we can guarantee the existence of a solution for some of Eqs. (66) and (68).

Theorem 4 reduces the search for defect eigenmodes and midgap eigenvalues of the perturbed operator $\mathfrak{M}$ to the investigation of the spectral attributes of the relevant Hilbert–Schmidt operator $\mathscr{S}(\omega)$. When it comes to numerical estimations, the reduction to the Hilbert–Schmidt operator $\mathscr{S}(\omega)$ is quite valuable, since this compact operator is more suitable for truncations than the original unbounded differential operator $\mathfrak{M}$ with nonsmooth coefficient $\varepsilon(\mathbf{x})$.

To estimate the number of defect eigenvalues, we need a function that counts eigenvalues. For a given self-adjoint operator H and interval (α, β), we define the counting function by the formula

$$N_H(\alpha, \beta) = \operatorname{Tr} \chi_{(\alpha,\beta)}(H). \tag{70}$$

Note that $N_H(\alpha, \beta)$ is always a nonnegative integer unless it is infinite. If H has discrete spectrum in (α, β), $N_H(\alpha, \beta)$ gives the number of eigenvalues of H in (α, β), counted according to their multiplicity.

It is convenient to write

$$\varepsilon(\mathbf{x}) = \frac{\varepsilon_0(\mathbf{x})}{1 + \theta(\mathbf{x})}, \tag{71}$$

where $\theta(x)$ is a bounded measurable function with compact support satisfying

$$-1 < \theta_- \leqslant \theta(x) \leqslant \theta_+ < \infty, \tag{72}$$

for some constants θ_- and θ_+. Note that a defect is positive if $\theta(\mathbf{x}) \geqslant 0$, in which case we have $\theta_- = 0$. Similarly, a defect is negative if $\theta(\mathbf{x}) \leqslant 0$, so $\theta_+ = 0$.

Our estimate of the number of eigenvalues in a gap is given by the following theorem[11]:

Theorem 5. Let $\varepsilon_0(\mathbf{x})$ be a periodic function satisfying relation (3), with the interval (ω_a^2, ω_b^2) being a gap in the spectrum of $\mathfrak{M}_0$. Let us insert a defect by taking $\varepsilon(\mathbf{x})$ as in Eq. (71), where $\theta(\mathbf{x})$ satisfies inequality (72) and vanishes outside a cube Λ_ℓ of side $\ell > 0$. Letting $0 < \delta \leqslant 1$, we have

(i) Assuming that the left-hand edge ω_a^2 is regular and that the defect is positive,

$$N_{\mathfrak{M}}(\omega_a^2, \omega_b^2) \leqslant C_{\omega_a^2,\varepsilon_{0,\pm},\theta_+,\delta}\theta_+^2(\ell + 3)^{9+\delta}\xi(\omega_a^2) < \infty, \tag{73}$$

with

$$C_{\omega_a^2,\varepsilon_{0,\pm},\theta_+,\delta} = \frac{(\omega_a^2 + 1)^2}{\delta} C[1 + \varepsilon_-^{-(6+\delta)}]C', \tag{74}$$

$$C' = \left[1 + 6\frac{\sqrt{\varepsilon_{0,+}}}{\varepsilon_{0,-}}\left(\sqrt{\varepsilon_-} + \frac{1}{\sqrt{\varepsilon_-}}\right)\right]^2, \tag{75}$$

where C is some constant, independent of δ and of the other parameters; $\varepsilon_- = \varepsilon_{0,-}/(1 + \theta_+)$; and $\xi(\omega_a^2)$ is as given in Eq. (44).

(ii) Assuming that the right-hand edge ω_b^2 is regular and that the defect is negative,

$$N_{\mathfrak{M}}(\omega_a^2, \omega_b^2) \leqslant C_{\omega_b^2,\varepsilon_{0,\pm},\theta_-,\delta}\theta_-^2(\ell + 3)^{9+\delta}\xi(\omega_b^2) < \infty, \tag{76}$$

with

$$C_{\omega_b^2,\varepsilon_{0,\pm},\theta_-,\delta} = \frac{(b + 1)^2}{\delta} C[1 + \varepsilon_{0,-}^{-(6+\delta)}]C', \tag{77}$$

$$C' = \left[1 + 6\frac{\sqrt{\varepsilon_+}}{\varepsilon_{0,-}}\left(\sqrt{\varepsilon_{0,-}} + \frac{1}{\sqrt{\varepsilon_{0,-}}}\right)\right]^2, \tag{78}$$

where C is some constant, independent of δ and of the other parameters; $\varepsilon_+ = \varepsilon_{0,+}/(1 + \theta_-)$; and $\xi(\omega_b^2)$ is as given in Eq. (45).

Since the function $N_{\mathfrak{M}}(\omega_a^2, \omega_b^2)$ is integer valued, $N_{\mathfrak{M}}(\omega_a^2, \omega_b^2) < 1$ implies that $N_{\mathfrak{M}}(\omega_a^2, \omega_b^2) = 0$. Thus we have the following immediate corollary, which tells us that there are no midgap eigenvalues if the defect is small.

Corollary 1. Let $\varepsilon_0(\mathbf{x})$ be a periodic function satisfying relation (3), with the interval (ω_a^2, ω_b^2) being a gap in the spectrum of $\mathfrak{M}_0$. Let us insert a defect by taking $\varepsilon(\mathbf{x})$ as in Eq. (71), where $\theta(\mathbf{x})$ satisfies inequality (72) and vanishes outside a cube Λ_ℓ of side $\ell > 0$. Then

(i) If the left-hand edge ω_a^2 is regular and the defect is positive, we must have $N_{\mathfrak{M}}(\omega_a^2, \omega_b^2) = 0$ for small $\theta_+ > 0$; how small depends only on ω_a^2, $\varepsilon_{0,\pm}$, and ℓ. In other words, there are no eigenvalues in the gap for weak defects.

(ii) If the right-hand edge ω_b^2 is regular and the defect is negative, we must have $N_{\mathfrak{M}}(\omega_a^2, \omega_b^2) = 0$ for small $\theta_- < 0$; how small depends only on ω_b^2, $\varepsilon_{0,\pm}$, and ℓ, so there are no eigenvalues in the gap for weak defects.

6. LOCALIZATION IN DISORDERED MEDIA

As we have seen, a strong enough single defect in a periodic dielectric medium with a spectral gap creates exponentially localized EM waves. If we have a random array of such defects, then, under some natural conditions, the localized waves created by individual defects do not couple (i.e., the EM wave tunneling becomes inefficient), so we get an infinite number of localized waves whose frequencies are dense in an interval contained in the spectral gap of the underlying periodic medium. This phenomenon is analogous to the Anderson localization of electron waves in random media, which has been studied intensively for the past four decades, in both the physics and the mathematics literature.[21–25,27–30]

The relevant mathematical problems led to the study of the spectral properties of differential, partial differential, and matrix linear operators with random position-dependent coefficients. For electron waves the coefficient describing the medium is a random potential $V(\mathbf{x})$, where $\mathbf{x}$ is the position in space. In our setting the medium is described by a random dielectric constant $\varepsilon(\mathbf{x})$. The randomness means (1) that for any given $\mathbf{x}$ we know that $\varepsilon(\mathbf{x})$ is a random quantity, and (2) that the random field $\varepsilon(\mathbf{x})$ is statistically homogeneous and ergodic.[23,27,29,30] Conditions (1) and (2) can be formulated as follows. The random medium will be a randomization of an underlying periodic medium with a lattice of periods $\mathbb{L}$. For any set of positions $\mathbf{x}_1, \ldots, \mathbf{x}_N$ the joint probability distribution $P_{\mathbf{x}_1,\ldots,\mathbf{x}_N}$ of the random quantities $\varepsilon(\mathbf{x}_1), \ldots, \varepsilon(\mathbf{x}_N)$ is $\mathbb{L}$ invariant; namely, for any vector $\mathbf{m}$ in the lattice $\mathbb{L}$ we have

$$P_{\mathbf{x}_1,\ldots,\mathbf{x}_N} = P_{\mathbf{x}_1+\mathbf{m},\ldots,\mathbf{x}_N+\mathbf{m}}. \tag{79}$$

This statistical homogeneity condition is very natural from a physical point of view. The ergodicity of the probability measure $\mathbb{P}$ describing all the random quantities $\varepsilon(\mathbf{x})$, with $\mathbf{x}$ running over three-dimensional space, means that all $\mathbb{L}$-invariant events must be trivial, i.e., their probability must be either 0 or 1. Lack of statistical correlations between $\varepsilon(\mathbf{x})$ and $\varepsilon(\mathbf{y})$ for large $|\mathbf{x} - \mathbf{y}|$ is usually sufficient for ergodicity. We always assume that a random dielectric constant $\varepsilon(\mathbf{x})$ is statistically homogeneous and ergodic.

Let us turn now to the spectral properties of a random self-adjoint differential or matrix operator A. It is customary to classify a spectral point ω based on whether the corresponding eigenmode u_ω is square integrable. In the first case, i.e., $\int |u_\omega(\mathbf{x})|^2 \mathrm{d}\mathbf{x} < \infty$, we say that ω is an eigenvalue (a true eigenvalue) and that u_ω is a localized eigenmode. We denote by $\sigma_{\mathrm{pp}}(A)$ the set of eigenvalues of A. It is a well-known basic fact that $\sigma_{\mathrm{pp}}(A)$ is at most countable, i.e.,

$$\sigma_{\mathrm{pp}}(A) = \{\omega_1, \omega_2, \ldots\}.$$

If $\int |u_\omega(\mathbf{x})|^2 \mathrm{d}\mathbf{x} = \infty$, as, for instance, for a plane wave, we call u_ω an extended eigenmode and classify ω as a point of the continuous spectrum $\sigma_c(A)$, a closed set. The entire spectrum $\sigma(A)$ is given by

$$\sigma(A) = \overline{\sigma_{\mathrm{pp}}(A)} \cup \sigma_c(A),$$

where $\overline{\sigma_{\mathrm{pp}}(A)}$ is the closure of the set $\sigma_{\mathrm{pp}}(A)$.

For a random (differential or matrix) operator A, we list below some of its unusual spectral properties, which may seem peculiar and exotic but are truly typical and always hold for a physically meaningful random medium. These spectral properties are due to ergodicity and self-averaging and hold in large generality with probability 1.[23,27,29,30]

- The sets $\sigma(A)$, $\overline{\sigma_{\mathrm{pp}}(A)}$, and $\sigma_c(A)$ are nonrandom.
- The integrated density of states (the number of states per unit volume) is nonrandom.
- Any interval contains either zero or infinitely many points of the spectrum.
- The probability that a fixed number θ is an eigenvalue of A is 0. In other words, in spite of the fact that the closure of the pure point spectrum $\overline{\sigma_{\mathrm{pp}}(A)}$ is nonrandom, the countable set $\sigma_{\mathrm{pp}}(A)$ of eigenvalues ω_j is random. The eigenvalues ω_j are sensitive to a particular sample (easily movable), so they never hit θ with probability 1.

In view of these properties, random operators typically exhibit spectral behavior unusual in classical spectral theory. For instance, assume that an interval I belongs to the spectrum $\sigma(A)$ of a random operator A (we can speak unambiguously of the spectrum, or of the closure of the pure point spectrum, since they are the same for almost all realizations of the random operator), and assume that A has only pure point spectrum in I. Then, with probability 1, the eigenvalues ω_j of the operator A form a dense countable subset in the interval I, i.e., any subinterval I' of I contains infinitely many eigenvalues ω_j. But there is zero probability that a fixed ω in I will be an eigenvalue.

In many interesting cases one can prove that the eigenmodes of the pure point spectrum decay exponentially. One can look at these unusual, but typical for random operators, spectral properties as merely a reflection of the complexity of the random environment.

Our strategy for proving Anderson localization[21] of EM waves is as follows[9]:

1. We study the effect of random perturbations on a spectral gap of the underlying periodic medium; we obtain estimates on the size of the spectral gap of the perturbed medium, showing that the gap does not close for random perturbations that are not too large.
2. The Maxwell operator $\mathfrak{M}_{\mathbf{H}}$ of the random medium is shown to have pure point spectrum in some closed subinterval I of the spectral gap of the underlying periodic medium, with all the corresponding eigenfunctions being exponentially decaying (in the sense of having exponentially decaying local L^2 norms[9]). For this operator we prove that the curl of an exponentially decaying eigenfunction is also exponentially decaying, so it follows from Eqs. (18) and (22) that the corresponding operator $\mathfrak{M}_{\mathbf{E}}$ also has pure point spectrum in the closed interval I, with all the corresponding eigenfunctions being exponentially decaying.
3. We conclude from Eq. (23) that the operator $\mathbb{M}$ has pure point spectrum for real ω such that ω^2 is in the interval I, with all the corresponding eigenfunctions being

exponentially decaying, so the energy densities of the corresponding solutions of Eqs. (1) are also exponentially decaying, uniformly in the time t, satisfying Eq. (28).

The localization of EM waves in random media is thus a consequence of Anderson localization for the random operator $\mathfrak{M} = \mathfrak{M}_{\mathbf{H}} = \nabla^{\times}(1/\varepsilon)\nabla^{\times}$ on $\mathbb{S}$, i.e., the existence of closed intervals where this random operator has pure point spectrum with exponentially decaying eigenfunctions, with probability 1.

We model a random array of defects in a periodic medium by a random dielectric constant $\varepsilon(\mathbf{x})$. For simplicity, we take the lattice $\mathbb{L}$ of periods to be $q\mathbb{Z}^3$, with q being a positive integer and $\mathbb{Z}^3$ the usual cubic lattice.

Assumption 2 (the random media). The dielectric constant $\varepsilon_g(\mathbf{x}) = \varepsilon_{g,\zeta}(\mathbf{x})$ is a random function of the form

$$\varepsilon_{g,\zeta}(\mathbf{x}) = \varepsilon_0(\mathbf{x})\gamma_{g,\zeta}(\mathbf{x}), \tag{80}$$

with

$$\gamma_{g,\zeta}(\mathbf{x}) = 1 + g \sum_{i \text{ in } \mathbb{Z}^3} \zeta_i u_i(\mathbf{x}), \tag{81}$$

where

(i) $\varepsilon_0(\mathbf{x})$ is a measurable real-valued function that is q periodic for some q in $\mathbb{N}$; i.e., $\varepsilon_0(\mathbf{x}) = \varepsilon_0(\mathbf{x} + q\mathbf{i})$ for all $\mathbf{x}$ in $\mathbb{R}^3$ and all $\mathbf{i}$ in $\mathbb{Z}^3$, with

$$0 < \varepsilon_{0,-} \leq \varepsilon_0(\mathbf{x}) \leq \varepsilon_{0,+} < \infty \tag{82}$$

for some constants $\varepsilon_{0,-}$ and $\varepsilon_{0,+}$.

(ii) $u_i(\mathbf{x}) = u(\mathbf{x} - \mathbf{i})$ for each $\mathbf{i}$ in $\mathbb{Z}^3$, with u being a nonnegative measurable real-valued function with compact support, say, $u(\mathbf{x}) = 0$ if $\|x\|_\infty \leq r_u$ for some $r_u < \infty$, such that

$$0 < U_- \leq U(\mathbf{x}) \equiv \sum_{i \text{ in } \mathbb{Z}^3} u_i(\mathbf{x}) \leq U_+ < \infty \tag{83}$$

for some constants U_- and U_+.

(iii) $\zeta = \{\zeta_i;\, i \text{ is in } \mathbb{Z}^3\}$ is a family of independent, identically distributed random variables taking values in the interval $[-1, 1]$, whose common probability distribution has a bounded density $\rho(t) = \rho_{\zeta_i}(t) > 0$ almost everywhere in $[-1, 1]$.

(iv) g, satisfying $0 \leq g < 1/U_+$, is the disorder parameter.

Notice that assumption 2 implies that each $\varepsilon_{g,\zeta}$ satisfies relation (3), with

$$\varepsilon_\pm = \varepsilon_{g,\pm} = \varepsilon_{0,\pm}(1 \pm gU_+) \tag{84}$$

The periodic operators associated with the periodic dielectric constant $\varepsilon_0(\mathbf{x})$ will carry the subscript 0, i.e., $M_0 = M(\varepsilon_0)$, $\mathfrak{M}_0 = \mathfrak{M}(\varepsilon_0)$. We study the random operators

$$M_g = M_{g,\zeta} = M(\varepsilon_{g,\zeta}), \tag{85}$$

$$\mathfrak{M}_g = \mathfrak{M}_{g,\zeta} = \mathfrak{M}(\varepsilon_{g,\zeta}). \tag{86}$$

It is a consequence of ergodicity that the spectrum of these operators is nonrandom; i.e., there exists a nonrandom set Σ_g such that $\sigma(\mathfrak{M}_{g,\zeta}) = \sigma(M_{g,\zeta}) = \Sigma_g$ with probability one. In addition, the decompositions of $\sigma(\mathfrak{M}_{g,\zeta})$ and $\sigma(M_{g,\zeta})$ into pure point spectrum, absolutely continuous spectrum, and singular continuous spectrum are also independent of the choice of ζ with probability 1.[27,29,30]

The following theorem[9] gives information on the location of Σ_g, the (nonrandom) spectrum of the random Maxwell operator $\mathfrak{M}_g$. It shows that, for sufficiently small values of the coupling constant g, the random array of defects shrinks the gap but does not close it.

We recall that a function $f(g)$ is called Lipschitz continuous on an interval if $|f(g) - f(g')| \leq C|g - g'|$ for all g, g' in the interval.

Theorem 6 (location of the spectrum). Let the random operator $\mathfrak{M}_g$ defined by Eq. (85) satisfy assumptions 2 and 1. Then there exists g_0, with

$$\frac{1}{U_+}\left(1 - \frac{\omega_a}{\omega_b}\right) \tag{87}$$

$$\leq g_0 \leq \frac{1}{U_+} \min\left\{1, \left[\left(\frac{\omega_b}{\omega_a}\right)^{U_+/U_-} - 1\right]\right\}, \tag{88}$$

and there exist strictly increasing Lipschitz continuous real-valued functions ${\omega_a}^2(g)$ and $-{\omega_b}^2(g)$ on the interval $[0, 1/U_+)$, with ${\omega_a}^2(0) = {\omega_a}^2$, ${\omega_b}^2(0) = {\omega_b}^2$, and ${\omega_a}^2(g) \leq {\omega_b}^2(g)$, such that

(i) Under the random perturbation, the spectrum expands into the gap $[{\omega_a}^2(g), {\omega_b}^2(g)]$:

$$\Sigma_g \cap [{\omega_a}^2, {\omega_b}^2] = [{\omega_a}^2, {\omega_a}^2(g)] \cup [{\omega_b}^2(g), {\omega_b}^2]. \tag{89}$$

(ii) For $g < g_0$, we have ${\omega_a}^2(g) < {\omega_b}^2(g)$, so $[{\omega_a}^2(g), {\omega_b}^2(g)]$ is a gap in the spectrum of the random operator $\mathfrak{M}_g$, located inside the gap $({\omega_a}^2, {\omega_b}^2)$ of the unperturbed periodic operator $\mathfrak{M}_0$. Moreover, we have

$${\omega_a}^2 \leq {\omega_a}^2(1 + gU_+)^{U_-/U_+} \leq {\omega_a}^2(g) \leq \frac{{\omega_a}^2}{1 - gU_+}, \tag{90}$$

$${\omega_b}^2(1 - gU_+) \leq {\omega_b}^2(g) \leq \frac{{\omega_b}^2}{(1 + gU_+)^{U_-/U_+}} \leq {\omega_b}^2. \tag{91}$$

(iii) If $g_0 < 1/U_+$, we have ${\omega_a}^2(g) = {\omega_b}^2(g)$ for all g in $[g_0, (1/U_+))$, and the random operator $\mathfrak{M}_g$ has no gap inside the gap $({\omega_a}^2, {\omega_b}^2)$ of the unperturbed periodic operator $\mathfrak{M}_0$, i.e., $[{\omega_a}^2, {\omega_b}^2] \subset \Sigma_g$.

We proceed with the proof of this theorem[9] by first approximating the (nonrandom) spectrum of the random operator by spectra of (nonrandom) periodic operators, which are then approximated by spectra of operators on finite cubes with periodic boundary condition. The latter operators have compact resolvents (i.e., Green's functions), and bounds on their eigenvalues are obtained by the min–max principle.

To state our results on localization, we need the following definition:

Definition 2 (exponential localization). We say that the random operator $\mathfrak{M}_g$ exhibits localization in an interval $I \subset \Sigma_g$, if $\mathfrak{M}_g$ has only pure point spectrum in I with probability 1. We have exponential localization in I if we have localization and if, with probability 1, all the eigenfunctions corresponding to eigenvalues in I are exponentially decaying (in the sense of having exponentially decaying local L^2 norms).

Remark 1. The curls of exponentially decaying eigenfunctions of $\mathfrak{M}_g$ always have exponentially decaying local L^2 norms.[9] Thus the corresponding energy densities [see Eqs. (4)] also have exponentially decaying local L^2 norms, uniformly in the time t.

We show that random perturbations create exponentially localized eigenfunctions near the edges of the gap. Our method of proof requires low probability of extremal values for the random variables. The results given below are formulated for the left-hand edge of the gap, with similar results holding at the right-hand edge.[9]

Theorem 7 (localization at the edge). Let the random operator $\mathfrak{M}_g$ defined by Eq. (85) satisfy assumptions 2 and 1, with

$$\int_{1-\gamma}^{1} \rho(t)\mathrm{d}t \leqslant K\gamma^{\eta} \quad \text{for } 0 \leqslant \gamma \leqslant 1, \tag{92}$$

where $K < \infty$ and $\eta > 3/2$. For any $g < g_0$ there exists $\delta(g) > 0$, such that the random operator $\mathfrak{M}_g$ exhibits exponential localization in the interval $[\omega_a{}^2(g) - \delta(g),\ \omega_a{}^2(g)]$.

Theorem 7 is proved by a multiscale analysis,[9,28,33] which reduces the proof to the verification of exponential decay of the (random) Green's function in a given finite scale, with high probability. This decay with high (appropriate for the scale) probability is then shown to hold for larger and larger scales. Finally, the exponential decay of the Green's function in all scales is used to show exponential localization. We discuss some of the key steps in the proof.

Given an open cube Λ in three-dimensional space and M as in Eq. (27), we denote by M_Λ the restriction of M to Λ with Dirichlet boundary condition, i.e., $M_\Lambda = \nabla^{\times}(1/\varepsilon)\nabla^{\times}$ acting on square-integrable functions on Λ with zero tangential component at the boundary. The corresponding Green's function is given by

$$G_\Lambda(\omega) = (M_\Lambda - \omega^2 I)^{-1}. \tag{93}$$

The multiscale analysis requires control of the norm of the Green's functions of the operators $M_{g,\zeta,\Lambda}$, with high probability. This is given by a Wegner-type estimate,[9] which says that the probability that the corresponding Green's functions $G_{g,\zeta,\Lambda}(\omega)$ are bigger than a given number $1/\eta$, $0 < \eta \leqslant \omega^2$, is no more than proportional to $\eta|\Lambda|^2$; in fact,

$$\mathbb{P}\left\{\left\|G_{g,\zeta,\Lambda}(\omega)\right\| \geqslant \frac{1}{\eta}\right\} \leqslant Q\omega\eta|\Lambda|^2, \tag{94}$$

where Q is some constant. This estimate is typically used when Λ is a cube of size L and $\eta = L^{-s}$ for suitable $s > 6$, so $1/\eta = L^s$ is large and $\eta|\Lambda|^2 = L^{-s+6}$ is small. It already indicates that, with high probability, the eigenvalues of $\mathfrak{M}_{g,\zeta,\Lambda}$ do not want to be too close to any given ω^2, a precursor of Anderson localization.

Now let $\Lambda_L(\mathbf{x})$ denote the cube of side L centered at $\mathbf{x}$. We write $G_{g,\zeta,\mathbf{x},L}(\omega) = G_{g,\zeta,\Lambda_L(\mathbf{x})}(\omega)$. Given that $m > 0$ and $\omega > 0$, we can say that the cube $\Lambda_L(\mathbf{x})$ is regular (for a fixed g, ζ), if

$$\|\Gamma_{\mathbf{x},L} G_{g,\zeta,\mathbf{x},L}(\omega)\chi_{\mathbf{x}}\| \leqslant \exp[-m(L/2)], \tag{95}$$

where $\chi_{\mathbf{x}}$ is as in Eq. (37) and $\Gamma_{\mathbf{x},L}$ is the characteristic function of $\Lambda_{L-q}(\mathbf{x}) - \Lambda_{L-3q}(\mathbf{x})$, which plays the role of a thick boundary. Note that relation (95) says that the finite-cube Green's function is localized in the sense that it decays exponentially from the center of the cube to its boundary, with the given rate m. It turns out that this regularity of the finite-cube Green's function is an indication of the exponential localization of the random operator and that, to prove localization, it suffices to show that it occurs with high probability at a sufficiently large scale.

For a given scale L, let us call $P(L)$ the probabilistic statement:

$$\mathbb{P}\{\Lambda_L(\mathbf{x}) \text{ is regular}\} \geqslant 1 - \frac{1}{L^p}, \tag{96}$$

where $p > 3$ is some fixed number of our choice. The multiscale analysis states that if we can verify $P(L_0)$ for some sufficiently large scale L_0, then $P(L)$ is also true for all scales $L = L_k$, where $L_{k+1} = L_k{}^{\alpha}$ for some appropriate $\alpha > 1$, $k = 0, 1, 2, \ldots$. This is shown by an induction argument,[28] by use of relation (94). In fact, one proves a stronger statement at all scales L_k, from which we get the conclusions of Theorem 7.

Thus, to prove Theorem 7, it suffices to verify $P(L)$ for some sufficiently large scale $L = L_0$. This is done by means of assumption (92) and a finite-cube version of Theorem 1. We call $P(L_0)$ the starting hypothesis for the multiscale analysis.

Theorem 7 can be extended to the situation in which the gap is filled by the spectrum of the random operator.[9] In this case we establish the existence of an interval (inside the original gap) where the random Maxwell operator exhibits exponential localization. Under somewhat different assumptions we can arrange for localization in any fraction of the gap as we want.

Theorem 8 (localization at the meeting of the edges). Let the random operator $\mathfrak{M}_g$ defined by Eq. (85) satisfy assumptions 2 and 1, with

$$\int_{1-\gamma}^{1} \rho(t)\mathrm{d}t, \quad \int_{-1}^{-1+\gamma} \rho(t)\mathrm{d}t \leqslant K\gamma^{\eta} \qquad \text{for } 0 \leqslant \gamma \leqslant 1, \tag{97}$$

where $K < \infty$ and $\eta > 3$. Assume that $g_0 < 1/U_+$ [e.g., if $(\omega_b/\omega_a)^{U_+/U_-} < 2$], so the random operator $\mathfrak{M}_g$ has no gap inside $(\omega_a{}^2, \omega_b{}^2)$ for g in $[g_0, (1/U_+)$. Then there exist $0 < \varepsilon < (1/U_+) - g_0$ and $\delta > 0$, such that the random operator $\mathfrak{M}_g$ exhibits exponential localization in the interval $[\omega_a{}^2(g_0) - \delta, \omega_a{}^2(g_0) + \delta]$ for all $g_0 \leqslant g < g_0 + \varepsilon$.

The proof of Theorem 8 is analogous to the proof of Theorem 7 if one takes into account both edges of the gap.

Remark 2. Theorems 7 and 8 should be true without the extra hypotheses (92) and (97), at least if the edges of the spectral gap are nondegenerate. They are used to obtain the starting hypothesis for the multiscale analysis in the proof of localization. If the edge of the gap is nondegenerate, one may expect estimates similar to Lifshitz tails[30] for the density of states inside the gap, which would replace hypotheses (92) and (97) in the proofs. This is how the starting hypothesis is obtained for random Schrödinger operators at the bottom of the spectrum.[26] Estimates of Lifshitz tails in spectral gaps of periodic Schrödinger operators have been obtained at nondegenerate edges.[38]

7. CONCLUSIONS

We discussed the localization of electromagnetic waves in lossless inhomogeneous dielectric media. Our starting point was a three-dimensional lossless periodic dielectric medium (photonic crystal) exhibiting a gap in the frequency spectrum (photonic bandgap). If such a medium is perturbed by either a single defect or a random array of defects, exponentially localized electromagnetic waves may arise with frequencies in the gap.

For a single defect, we gave a simple condition to ensure the rise of exponentially localized electromagnetic waves with frequency in a specified subinterval of the photonic bandgap. We derived equations for these midgap frequencies and estimated their number.

For a random array of defects, we showed that, under some natural conditions, the gap shrinks but does not close, and we get an infinite number of localized electromagnetic waves with frequencies dense in an interval contained in the spectral gap of the underlying periodic medium. This phenomenon is analogous to the Anderson localization of electron waves in random media.

An important technical achievement of our proofs is that no assumptions are made about the smoothness of the function $\varepsilon(\mathbf{x})$, which are so common in almost all classical results on partial differential elliptic operators. This was possible owing to a variational approach to the problems and to the treatment of the relevant operators as quadratic forms. Such general conditions on $\varepsilon(\mathbf{x})$, i.e., the bounds in relation (3) and the lack of smoothness, are required on physical grounds. In practice, only a few materials are used in the fabrication of periodic and disordered media, in which case $\varepsilon(\mathbf{x})$ takes just a finite number of values, so $\varepsilon(\mathbf{x})$ is piecewise constant and hence discontinuous.

ACKNOWLEDGMENTS

The effort of A. Figotin is sponsored by the U.S. Air Force Office of Scientific Research, Air Force Materiel Command, under grant F49620-97-1-0019. The work of A. Klein was supported in part by National Science Foundation grant DMS-9500720. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Air Force Office of Scientific Research or the U.S. Government.

REFERENCES

1. P. W. Anderson, "A question of classical localization. A theory of white paint," Philos. Mag. B **53**, 505–509 (1985).
2. P. Sheng, ed., *Scattering and Localization of Classical Waves* (World Scientific, Singapore, 1990).
3. S. John, "Localization of light," Phys. Today **44**(5), 32–40 (1991).
4. S. John, "The localization of light," in *Photonic Band Gaps and Localization*, C. M. Soukoulis, ed., Vol. 308 of NATO ASI Ser. B. (Plenum, New York, 1993), pp. 1–22.
5. R. Meade, D. Brommer, A. Rappe, and J. Joannopoulos, "Photonic bound states in periodic materials," Phys. Rev. Lett. **67**, 3380–3384 (1991).
6. J. Rarity and C. Weisbuch, eds., *Microcavities and Photonic Bandgaps: Physics and Applications* (Kluwer Academic, Dordrecht, The Netherlands, 1995).
7. C. Soukoulis, ed., *Photonic Band Gap Materials* (Kluwer Academic, Dordrecht, The Netherlands, 1996).
8. A. Figotin and A. Klein, "Localization of electromagnetic and acoustic waves in random media. Lattice model," J. Stat. Phys. **76**, 985–1003 (1994).
9. A. Figotin and A. Klein, "Localization of classical waves II: Electromagnetic waves," Commun. Math. Phys. **184**, 411–441 (1997).
10. A. Figotin and A. Klein, "Localized classical waves created by defects," J. Stat. Phys. **86**, 165–177 (1997).
11. A. Figotin and A. Klein, "Midgap defect modes in dielectric and acoustic media," SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. (to be published).
12. Special issue on development and applications of materials exhibiting photonic bandgaps, J. Opt. Soc. Am. B **10**, 279–413 (1993).
13. P. R. Villeneuve and M. Piché, "Photonic band gaps in periodic dielectric structures," Prog. Quantum Electron. **18**, 153–200 (1994).
14. P. M. Hui and N. F. Johnson, "Photonic band-gap materials," in *Solid State Physics*, H. Ehrenreich and F. Spaepen, eds. (Academic, New York, 1995), Vol. 49, pp. 151–203.
15. P. R. Villeneuve and J. Joannopoulos, "Working at the speed of light," Sci. Spectra **9**, 18–24 (1997).
16. J. Joannopoulos, R. Meade, and J. Winn, *Photonic Crystals* (Princeton U. Press, Princeton, N.J., 1995).
17. C. Soukoulis, ed., *Photonic Band Gaps and Localization* (Plenum, New York, 1993).
18. E. Yablonovitch, T. Gmitter, R. Meade, D. Brommer, A. Rappe, and J. Joannopoulos, "Donor and acceptor modes in photonic periodic structure," Phys. Rev. B **44**, 13772–13774 (1991).
19. A. Figotin and P. Kuchment, "Band-gap structure of spectra of periodic dielectric and acoustic media. I. Scalar model," SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **56**, 68–88 (1996).
20. A. Figotin and P. Kuchment, "Band-gap structure of spectra of periodic dielectric and acoustic media. II. 2D photonic crystals," SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **56**, 1561–1620 (1996).
21. P. W. Anderson, "Absence of diffusion in certain random lattice," Phys. Rev. **109**, 1492–1505 (1958).
22. B. Shklovskii and A. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Heidelberg, 1984).
23. I. M. Lifshits, S. A. Greduskul, and L. A. Pastur, *Introduction to the Theory of Disordered Systems* (Wiley, New York, 1988).
24. J. Fröhlich and T. Spencer, "Absence of diffusion in the Anderson tight binding model for large disorder or low energy," Commun. Math. Phys. **88**, 151–184 (1983).
25. J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer, "Constructive proof of localization in the Anderson tight binding model," Commun. Math. Phys. **101**, 21–46 (1985).
26. H. Holden and F. Martinelli, "On absence of diffusion near the bottom of the spectrum for a random Schrödinger operator on $L^2(\mathbb{R}^\nu)$," Commun. Math. Phys. **93**, 197–217 (1984).

27. H. Cycon, R. Froese, W. Kirsch, and B. Simon, *Schrödinger Operators* (Springer-Verlag, Heidelberg, 1987).
28. H. Dreifus and A. Klein, "A new proof of localization in the Anderson tight binding model," Commun. Math. Phys. **124**, 285–299 (1989).
29. R. Carmona and J. Lacroix, *Spectral Theory of Random Schrödinger Operators* (Birkhäuser, Boston, Mass., 1990).
30. L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators* (Springer-Verlag, Heidelberg, 1991).
31. J. Combes and P. Hislop, "Localization for some continuous, random Hamiltonians in d-dimensions," J. Funct. Anal. **124**, 149–180 (1994).
32. A. Figotin and A. Klein, "Localization phenomenon in gaps of the spectrum of random lattice operators," J. Stat. Phys. **75**, 997–1021 (1994).
33. A. Figotin and A. Klein, "Localization of classical waves I: Acoustic waves," Commun. Math. Phys. **180**, 439–482 (1996).
34. J. Maynard, "Acoustic Anderson localization," in *Random Media and Composites*, B. V. Kohn and G. W. Milton, eds., (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1988), pp. 206–207.
35. W. Kohler, G. Papanicolaou, and B. White, "Localization and mode conversion for elastic waves in randomly layered media," Wave Motion **23**, 1–22 and 181–201 (1996).
36. A. Maradudin, E. Montroll, and G. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic, New York, 1963).
37. L. Deych and A. Lisyansky, "Impurity localization of electromagnetic waves in polariton region," Phys. Rev. Lett. (to be published).
38. Research done by F. Klopp on Internal Lifshits tails for random perturbations of periodic Schrödinger operators.
39. L. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).
40. M. Reed and B. Simon, *Analysis of Operators*, Vol. 4 of Methods of Modern Mathematical Physics (Academic, New York, 1978).
41. M. Klaus, "Some applications of the Birman–Schwinger principle," Helv. Phys. Acta **55**, 49–68 (1982).