MIT Open Access Articles

Speed-of-light limitations in passive linear media

Speed-of-light limitations in passive linear media

Aaron Welters,[*] Yehuda Avniel, and Steven G. Johnson

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Received 1 May 2014; published 26 August 2014)

We prove that well-known speed-of-light restrictions on electromagnetic energy velocity can be extended to a new level of generality, encompassing even nonlocal chiral media in periodic geometries, while at the same time weakening the underlying assumptions to only passivity and linearity of the medium (either with a transparency window or with dissipation). As was also shown by other authors under more limiting assumptions, passivity alone is sufficient to guarantee causality and positivity of the energy density (with no thermodynamic assumptions). Our proof is general enough to include a very broad range of material properties, including anisotropy, bianisotropy (chirality), nonlocality, dispersion, periodicity, and even delta functions or similar generalized functions. We also show that the "dynamical energy density" used by some previous authors in dissipative media reduces to the standard Brillouin formula for dispersive energy density in a transparency window. The results in this paper are proved by exploiting deep results from linear-response theory, harmonic analysis, and functional analysis that had previously not been brought together in the context of electrodynamics.

I. INTRODUCTION

Because relativity implies that information should travel more slowly than c (the speed of light in vacuum), there is a long history of interest in proving how this subluminal restriction arises for electromagnetic energy velocity $\mathbf{v}_e$, and various authors have proved that the speed-of-light limitation $||\mathbf{v}_e|| \leqslant c$ holds for homogeneous causal dispersive media [1–6] and [7, Sec. 84], for homogeneous passive isotropic or anisotropic dispersive media [8–10], or for periodic nondispersive media [11, p. 41], usually assuming negligible loss (i.e., in a "transparency window"). For lossy media various authors have proved similar bounds, but with some disagreement over the definition of energy velocity [2,4–6,8,12]. In this paper, we extend those results by both greatly simplifying the underlying assumptions and by generalizing the reach of the results. Similar in spirit to [8–10] (which were limited to homogeneous media), we prove $||\mathbf{v}_e|| \leqslant c$ assuming only that we have a *passive* medium (one which does no net work on the fields) with or without a transparency window, and from this we obtain not only $||\mathbf{v}_e|| \leqslant c$ but also causality of the polarization response and positivity of the energy density (which some other authors justified on thermodynamic grounds [7, Sec. 80]). Furthermore, our results apply to arbitrary anisotropic and *bianisotropic* (chiral) dispersive media, both homogeneous (Sec. II, similar to Yaghjian [9,10] and Glasgow *et al.* [8]) and *periodic* (Sec. III) media, include certain classes of spatially *nonlocal* media and only require that the medium be passive "on average" in the unit cell of periodicity. Mathematically, we generalize the concept of a linear susceptibility to arbitrary distributions over the vector-valued Hilbert space of electromagnetic fields in the unit cell, and extend the notion of a transparency window to a very general measure-theoretic definition that we connect to traditional formulations of electromagnetic dissipation. The key elements of our proof (whose technical details are in Appendix B) are derived from important previous theorems

on passive convolution operators and Herglotz functions from linear-response theory and functional analysis [13–20]. The energy velocity $\mathbf{v}_e$ in a transparency window is equivalent to the well-known group velocity $\mathbf{v}_g$ [2,7,11,21–23], and in Appendix A we extend this well-known equivalence $\mathbf{v}_e = \mathbf{v}_g$ [2,7,11,21,22,24–32]. Furthermore, we strengthen the $||\mathbf{v}_e|| \leqslant c$ bound by showing that $||\mathbf{v}_e|| = c$ can only hold for a very special class of materials and field solutions (Sec. IV). In Sec. V, we extend these bounds to lossy media following Glasgow *et al.* [8], and we show that the "dynamical energy density" used for lossy media in previous work [8,9,12,33–35] reduces to the familiar dispersive energy density [2,7,21,24] in the limit of a transparency window (greatly simplifying and generalizing earlier proofs [9,33–35]). Finally, in Sec. VI we discuss several open problems.

Several important results were derived in previous work on energy-velocity limitations in electromagnetism. Brillouin [1,2] assumed a homogeneous, isotropic, linear medium with a Lorentzian model of material dispersion. The original derivation of energy density for the single-resonance Lorentz medium by Brillouin, however, included an algebraic error. This was first pointed out by Loudon [3,4] who, after correcting this error, obtained a subluminal expression for the energy velocity in a Lorentz medium, and Loudon's formulation was subsequently extended by other authors [5,6]. However, some later works argued that one should include an additional term in the energy density (which appears in the denominator of the energy velocity), equal to a total dissipated energy [8,9,34,35], and this also results in a subluminal energy velocity [8]; our bounds apply regardless of whether this term is included. Whereas the early work was limited to a specific (usually Lorentz "oscillator") model of dispersion, the newer "dynamical" energy density has the advantage of being model independent. In the case of negligible loss (where there is no disagreement about the energy density), the energy velocities derived by Brillouin and Loudon are shown by Loudon to be equivalent and bounded by c. Landau and Lifshitz [7] assumed homogeneous, isotropic, linear dispersive materials satisfying causality, positivity of energy density due to thermodynamic considerations, and some regularity

———————
[*]awelters@math.mit.edu

conditions on the permittivity and permeability as a function of frequency. Their derivation of $\|\mathbf{v}_e\| \leqslant c$ in a transparency window is then based on the Kramers-Kronig relations and the positivity of the energy density. Yaghjian [9] was able to show positivity of the energy density (with some later corrections [10]) and $\|\mathbf{v}_e\| \leqslant c$ assuming only passivity of the medium in a transparency window, anticipating our results from Sec. II. Although he derived a positive energy density for arbitrary bianisotropic homogeneous media, Yaghjian [9] only applied this to a speed-of-light bound in the isotropic case; also, that derivation did not explicitly specify the allowed functional forms of the susceptibilities, nor did it connect those results to causality or to the theorems of linear-response theory (some of which are rederived by [9] in limited forms). Nevertheless, the results of Yaghjian can be viewed as a precursor of our results. For lossy media, Glasgow *et al.* [8] employed similar passivity assumptions in order to obtain positivity of the dissipated energy and hence subluminal energy velocity. All of these previous results, however, were for energy velocity defined pointwise in space, or equivalently for light propagation in a homogeneous medium (or possibly a homogenized approximation to an inhomogeneous "metamaterial"), but a more general setting is that of *periodic* media in which Bloch waves propagate with a well-defined group and energy velocity [11,26–32,36]. Joannopoulos *et al.* [11] proved $\|\mathbf{v}_e\| \leqslant c$ in this periodic setting, but only under the assumptions of isotropic, dispersionless, linear media with real permittivity and permeability bounded below by one. In this paper, we combine the weaker assumption of passivity with a very general setting of periodic media in order to unify and generalize these past results.

II. HOMOGENEOUS MEDIA

In this section, we prove the speed-of-light limitation in the simplified case of homogeneous passive linear media, in preparation for the more difficult case of periodic media in Sec. III. In Sec. II A, we review the definitions of planewave, group velocity, and energy velocity. In Sec. II B, we set up the basic properties of susceptibility and define passivity. In Sec. II C, we give a definition of a transparency window. In Sec. II D, we derive the speed-of-light limitation in a transparency window.

A. Planewaves and energy velocity

In a homogeneous medium, because of the continuous translational symmetry, Maxwell's equations admit nontrivial solutions which are time-harmonic electromagnetic planewaves

$$\mathbf{E}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{H}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{1}$$

where $\mathbf{E}$, $\mathbf{H}$ are constant vectors, with real frequency ω and real wavevector $\mathbf{k}$ (in bandwidth of negligible loss), satisfying a dispersion relation $\omega = \omega(\mathbf{k})$. The group velocity is defined as $\mathbf{v}_g = \nabla_\mathbf{k}\omega(\mathbf{k})$, where $\nabla_\mathbf{k}$ denotes the gradient, which is the velocity of narrow-bandwidth wave packets [2,21]. The energy velocity $\mathbf{v}_e$ of such a field is defined [2,7,11,21–23] as the ratio of the time-averaged energy flux to the time-averaged energy

density U, i.e.,

$$\mathbf{v}_e = \frac{\text{Re}\,\mathbf{S}}{U}, \quad \text{where} \quad \mathbf{S} = \frac{c}{8\pi}\mathbf{E}\times\mathbf{H}^* \tag{2}$$

is the complex Poynting vector in Gaussian units. It is known [2,7,11,21–23] that $\mathbf{v}_g = \mathbf{v}_e$ for transparent media and we review the fact that this is true even for dispersive media in Appendix A. In dispersionless isotropic or anisotropic media, the energy density is $U = \frac{1}{16\pi}(\mathbf{E}^\dagger\epsilon\mathbf{E} + \mathbf{H}^\dagger\mu\mathbf{H})$, where ϵ, μ are the frequency-independent permittivity and permeability tensors, respectively, but when the media are dispersive with negligible loss then the energy density is given by the Brillouin formula [2,7,21,24], namely, $U = \frac{1}{16\pi}(\mathbf{E}^\dagger\frac{d\omega\epsilon(\omega)}{d\omega}\mathbf{E} + \mathbf{H}^\dagger\frac{d\omega\mu(\omega)}{d\omega}\mathbf{H})$.

More generally, we consider anisotropic, bianisotropic, and dispersive linear media, which are described by a 6×6 susceptibility χ:

$$\begin{bmatrix} \mathbf{P} \\ \mathbf{M} \end{bmatrix} = \chi * \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix} = \int_{-\infty}^{\infty} \chi(t')\begin{bmatrix} \mathbf{E}(t-t') \\ \mathbf{H}(t-t') \end{bmatrix}dt' \tag{3}$$

for the electric and magnetic polarizations $\mathbf{P} = \frac{1}{4\pi}(\mathbf{D} - \mathbf{E})$ and $\mathbf{M} = \frac{1}{4\pi}(\mathbf{B} - \mathbf{H})$. Mathematically, the convolution is defined in the distribution sense as described in Appendix B 3 a, allowing the susceptibility to be a generalized function (such as a delta function), i.e., a distribution on the smooth vector-valued functions $[\mathbf{E}(t) \quad \mathbf{H}(t)]^{\text{T}}$ with compact support. For such general materials, when the frequency ω belongs to a transparency window (as defined in Sec. II C), the Brillouin formula for the energy density U can be generalized [9,10,25,33–35] in terms of the Fourier transform of electromagnetic susceptibility $\widehat{\chi}(\omega)$ (a 6×6 matrix-valued function of frequency ω) by

$$U = \frac{1}{16\pi}\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}^\dagger \frac{d}{d\omega}\omega\,[I + 4\pi\widehat{\chi}(\omega)]\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}, \tag{4}$$

where I is the identity matrix.

B. Passivity

The key property of the susceptibility that will be used to derive the energy-velocity bound $\|\mathbf{v}_e\| \leqslant c$ and other properties is *passivity*, which is the statement that *polarization currents do not do work*. Mathematically, this is the following condition: the inequality

$$0 \leqslant \text{Re}\int_{-\infty}^{t}\mathbf{E}(t')^\dagger\frac{d\mathbf{P}(t')}{dt'} + \mathbf{H}(t')^\dagger\frac{d\mathbf{M}(t')}{dt'}\,dt' \tag{5}$$

must hold for all time t and all smooth vector-valued functions $[\mathbf{E}(t) \quad \mathbf{H}(t)]^{\text{T}}$ with compact support. Physically, the integral in Eq. (5) represents the work that the fields do on the bound currents [22]. Thus, a passive medium is one satisfying the passivity condition (5) or, in other words, is one in which energy can only be absorbed by the material but never generated by it.

A key point here, which is known in a closely related distributional setting of Zemanian [16] and proved in this setting in Appendix B 3 a, is that the convolution equation (3) and the passivity condition (5) mean that $\frac{d\chi}{dt} *$ is a special kind of convolution operator known as a "passive" convolution operator [16, Def. 8.2-2], a fact that has many important consequences. We will now describe three of these consequences

which are proved in Appendix B 3 b: causality, analyticity, and positivity. As we shall see later in this paper, it is the latter two properties along with a transparency window that are key to deriving positivity of the energy density, which is used to derive the speed-of-light limitation $\|\mathbf{v}_e\| \leqslant c$.

First, passivity implies causality: the polarization only depends on fields in the past, i.e., $\chi*$ is a causal convolution operator [16, Def. 4.6-1] or, equivalently, $\chi(t) = 0$ for $t < 0$ (in the distributional sense [16, Sec. 3.3, p. 55]). Technically, passivity only implies that $\frac{d\chi}{dt}*$ is a causal convolution operator, but this means that since χ is an antiderivative of $\frac{d\chi}{dt}$ then there exists some antiderivative X that differs from χ by a constant such that $X*$ is a causal convolution operator. Our proof below does not require that additive constant to be zero, but on physical grounds it of course must be (as a nonzero constant would mean the polarizations would depend on the fields at all times in the future and past).

Second, the Fourier transform $\widehat{\chi}(\omega) = (\mathscr{F}\chi)(\omega)$ exists for $\operatorname{Im}\omega > 0$ and is an analytic 6×6 matrix-valued function of the complex frequency ω in the upper half-plane. Third, the matrix

$$h(\omega) = \omega\widehat{\chi}(\omega), \quad \operatorname{Im}\omega > 0 \qquad (6)$$

has a positive-semidefinite imaginary part, i.e.,

$$\operatorname{Im}h(\omega) \geqslant 0, \quad \operatorname{Im}\omega > 0, \qquad (7)$$

where the imaginary part is defined as $\operatorname{Im}h(\omega) = \frac{1}{2i}[h(\omega) - h(\omega)^\dagger]$. The above two properties mean that $h(\omega)$ is what is known as a matrix-valued Herglotz function [19], i.e., a matrix-valued function analytic in the upper half-plane with positive-semidefinite imaginary part in that domain. In fact, a necessary and sufficient condition for $h(\omega) = \omega\widehat{\chi}(\omega)$ to be a Herglotz function is that $\frac{d\chi}{dt}*$ is a passive convolution operator, i.e, that the passivity condition (5) holds. Herglotz functions have been heavily studied (see, for instance, [17,19,20,37], and references within) and their properties, along with a transparency window, are the key to the positivity of the energy density and the speed-of-light limitation.

C. Transparency window

A transparency window for a passive linear media is defined as a frequency interval $(\omega_1, \omega_2) \subseteq \mathbb{R}$ where losses are negligible [7, Secs. 80 and 84]. A precise definition in terms of the Herglotz function $h(\omega)$ will be given in this section. Recall that we require a transparency window for a well-defined energy velocity and for the group velocity to equal the energy velocity. In typical circumstances, the condition of a transparency window is simply that $\operatorname{Im}h(\omega) = 0$ in the interval, but more generally the definition is given in terms of a measure induced by $\operatorname{Im}h(\omega)$ which we discuss briefly in this section. Two important properties of the function $h(\omega)$ in the transparency window will also be discussed. In particular, we show that the Herglotz function $h(\omega)$ is a monotonically increasing and differentiable Hermitian-matrix-valued function of frequency in the window. These properties are critical since they imply the positivity of the energy density, which is used to derive the speed-of-light limitation on the energy velocity in the next section.

Now, as is to be expected, there is a correspondence between the imaginary part $\operatorname{Im}h(\omega)$ of the Herglotz function and electromagnetic losses (i.e., dissipation of energy) due to the transfer of electromagnetic energy into matter by absorption. The simplest case in which to understand this correspondence is that of isotropic media with integrable susceptibility $\chi(t)$ and time-harmonic fields $\mathbf{E}(t) = \mathbf{E}e^{-i\omega t}$, $\mathbf{H}(t) = \mathbf{H}e^{-i\omega t}$ (with spatial dependence suppressed). In this case, the permittivity $\epsilon(\omega)$ and permeability $\mu(\omega)$ are scalar functions which are related to the susceptibility by $\widehat{\chi}(\omega) = \begin{bmatrix} \frac{\epsilon(\omega)-1}{4\pi}I & 0 \\ 0 & \frac{\mu(\omega)-1}{4\pi}I \end{bmatrix}$. It follows from (3) that the polarization $\mathbf{P}(t) = \mathbf{P}e^{-i\omega t}$ and magnetization $\mathbf{M}(t) = \mathbf{M}e^{-i\omega t}$ are time harmonic with $\mathbf{P} = \frac{\epsilon(\omega)-1}{4\pi}\mathbf{E}$ and $\mathbf{M} = \frac{\mu(\omega)-1}{4\pi}\mathbf{H}$. Now, the value of the integrand in Eq. (5) for the real parts of these time-harmonic fields represents the power dissipated per unit time and so its time average is just

$$\frac{1}{p}\int_0^p \operatorname{Re}\mathbf{E}(t')^\dagger \frac{d\operatorname{Re}\mathbf{P}(t')}{dt'} + \operatorname{Re}\mathbf{H}(t')^\dagger \frac{d\operatorname{Re}\mathbf{M}(t')}{dt'}\, dt'$$

$$= \frac{1}{2}\operatorname{Re}[\mathbf{E}^\dagger(-i\omega\mathbf{P}) + \mathbf{H}^\dagger(-i\omega\mathbf{M})]$$

$$= \frac{1}{2}\begin{bmatrix}\mathbf{E}\\\mathbf{H}\end{bmatrix}^\dagger \operatorname{Im}h(\omega)\begin{bmatrix}\mathbf{E}\\\mathbf{H}\end{bmatrix}$$

$$= \frac{\omega}{8\pi}[\operatorname{Im}\epsilon(\omega)\|\mathbf{E}\|^2 + \operatorname{Im}\mu(\omega)\|\mathbf{H}\|^2], \qquad (8)$$

where $p = 2\pi/\omega$. But (8) is just the well-known expression for the electric and magnetic losses in a dispersive isotropic medium for monochromatic electromagnetic fields [7, Sec. 80, Eq. (80.4)].

The next simplest case to understand is that of isotropic media with nonmonochromatic fields in which the electromagnetic fields $\mathbf{E}(t)$, $\mathbf{H}(t)$ are square integrable and both $\chi(t)$, $\frac{d\chi}{dt}(t)$ are integrable. Then, by Young's convolution theorem [38, Theorem 23.44] and Plancherel's theorem [39, Theorem IX.6], the convolution in Eq. (3) can be Fourier transformed and the total energy dissipated by such fields is the quantity (5) with $t = \infty$, which is just the value

$$\operatorname{Re}\int_{-\infty}^{\infty}\mathbf{E}(t')^\dagger \frac{d\mathbf{P}(t')}{dt'} + \mathbf{H}(t')^\dagger \frac{d\mathbf{M}(t')}{dt'}\, dt'$$

$$= \operatorname{Re}\int_{-\infty}^{\infty}\widehat{\mathbf{E}}(\omega)^\dagger[-i\omega\widehat{\mathbf{P}}(\omega)] + \widehat{\mathbf{H}}(\omega)^\dagger[-i\omega\widehat{\mathbf{M}}(\omega)]\frac{d\omega}{2\pi}$$

$$= \int_{-\infty}^{\infty}\begin{bmatrix}\widehat{\mathbf{E}}(\omega)\\\widehat{\mathbf{H}}(\omega)\end{bmatrix}^\dagger \operatorname{Im}h(\omega)\begin{bmatrix}\widehat{\mathbf{E}}(\omega)\\\widehat{\mathbf{H}}(\omega)\end{bmatrix}\frac{d\omega}{2\pi}$$

$$= \frac{1}{4\pi}\int_{-\infty}^{\infty}\omega[\operatorname{Im}\epsilon(\omega)\|\widehat{\mathbf{E}}(\omega)\|^2 + \operatorname{Im}\mu(\omega)\|\widehat{\mathbf{H}}(\omega)\|^2]\frac{d\omega}{2\pi}. \qquad (9)$$

But, (9) is just the well-known expression for the electric and magnetic losses in a dispersive isotropic medium for non-monochromatic electromagnetic fields which go sufficiently rapidly to zero as $t \to \pm\infty$ [7, Sec. 80, Eq. (80.6)].

In the most general case of passive linear media, the susceptibility $\chi(t)$ is a generalized function and so it may not be integrable. Thus, the correspondence between the imaginary part $\operatorname{Im}h(\omega)$ and losses is more complicated, involving

concepts of measures which we discuss in Appendix B 3 c in more detail. In particular, it is shown in Appendix B 3 c that there exists a non-negative matrix-valued measure Ω on the bounded Borel subsets of $\mathbb{R}$ such that (5), with $t = \infty$ and for any smooth vector-valued function $F(t) = [\mathbf{E}(t) \quad \mathbf{H}(t)]^{\mathrm{T}}$ with compact support, is given in terms of an integral involving only this measure and the Fourier transform of F integrated over all real frequencies, namely,

$$\mathrm{Re} \int_{-\infty}^{\infty} \mathbf{E}(t')^{\dagger} \frac{d\mathbf{P}(t')}{dt'} + \mathbf{H}(t')^{\dagger} \frac{d\mathbf{M}(t')}{dt'} \, dt' \qquad (10)$$

$$= \frac{1}{2} \int_{\mathbb{R}} d\widehat{F}(\omega)^{\dagger} \Omega_{\omega} \widehat{F}(\omega). \qquad (11)$$

Moreover, we prove in Appendix B 3 c that the value of this measure Ω on any finite open interval is given below by (12), which again comes from the imaginary part $\mathrm{Im}\, h(\omega)$ by taking a limit as we approach the real axis from the upper half-plane.

From the discussion above we can conclude that a way to quantify how lossy a material is in a certain finite frequency range $(\omega_1, \omega_2) \subseteq \mathbb{R}$ is to use the limits

$$\Omega((\omega_1, \omega_2)) = \lim_{\delta \downarrow 0} \lim_{\eta \downarrow 0} \int_{\omega_1+\delta}^{\omega_2-\delta} \frac{1}{\pi} \mathrm{Im}\, h(\omega + i\eta) \, d\omega. \quad (12)$$

The fact that these limits always exist, which we prove in Appendix B 3 c, is just based on fundamental properties of matrix-valued Herglotz functions described in Ref. [19]. The reason that it is necessary to define the limits in Eq. (12) as such is to handle the most general possible cases for h that can arise from the theory of Herglotz functions [19]. One can see how this measures losses in the frequency range (ω_1, ω_2) for isotropic media under the conditions on $\chi(t)$ which lead to the identities (8) and (9) since in this case Eq. (12) simplifies to

$$\Omega((\omega_1, \omega_2)) = \lim_{\delta \downarrow 0} \lim_{\eta \downarrow 0} \int_{\omega_1+\delta}^{\omega_2-\delta} \frac{1}{\pi} \mathrm{Im}\, h(\omega + i\eta) \, d\omega$$

$$= \lim_{\delta \downarrow 0} \int_{\omega_1+\delta}^{\omega_2-\delta} \frac{1}{\pi} \lim_{\eta \downarrow 0} \mathrm{Im}\, h(\omega + i\eta) \, d\omega$$

$$= \int_{\omega_1}^{\omega_2} \frac{1}{\pi} \mathrm{Im}\, h(\omega) \, d\omega. \qquad (13)$$

Now, we give the precise definition of a transparency window using (12). The interval (ω_1, ω_2) is said to be a transparency window for the medium if losses are negligible in this frequency range, i.e., if

$$\Omega((\omega_1, \omega_2)) = 0. \qquad (14)$$

As shown in Appendix B 3 c, it follows from this definition that if (ω_1, ω_2) is a transparency window, then $h(\omega)$ can be analytically continued into the lower half-plane through this interval (ω_1, ω_2) to a function satisfying the Hermitian and monotonicity conditions

$$\mathrm{Im}\, h(\omega) = 0, \quad \frac{dh}{d\omega}(\omega) \geqslant 0, \quad \omega \in (\omega_1, \omega_2). \qquad (15)$$

As mentioned above, these properties, summarized in Fig. 1, are key since, as we shall see in the next section, they imply the positivity of the energy density and the speed-of-light limitation on the energy velocity.
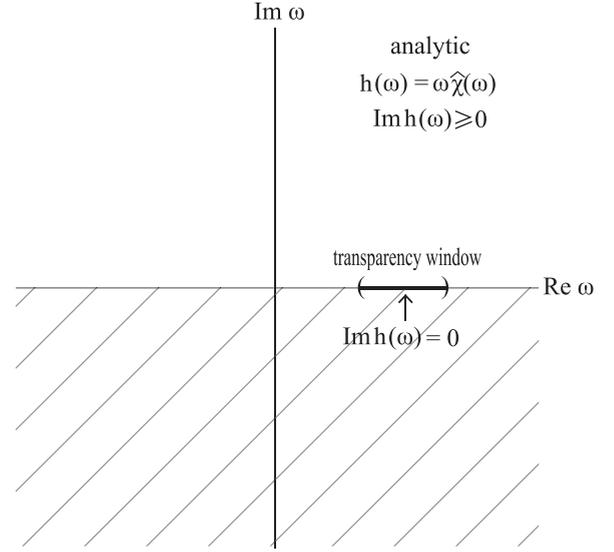

FIG. 1. The analytic properties of the function $h(\omega) = \omega\widehat{\chi}(\omega)$ along with a transparency window. This is all that is needed to prove the speed-of-light limitation $\|\mathbf{v}_e\| \leqslant c$. Moreover, the analytic properties of $h(\omega)$ are necessary and sufficient conditions for $\frac{d\chi}{dt}*$ to be a passive convolution operator.

D. Speed-of-light limitation

We now show that the speed-of-light limitation

$$\|\mathbf{v}_e\| \leqslant c \qquad (16)$$

holds for any time-harmonic electromagnetic planewave with frequency ω in a transparency window (ω_1, ω_2). To do this, we break the proof into two steps. First, we bound the numerator $\mathrm{Re}\,\mathbf{S}$ in Eq. (2) using elementary inequalities. Second, we bound the denominator U using more sophisticated inequalities that follow from the properties of the Herglotz function $h(\omega)$.

To bound the numerator in Eq. (2), we first use the Lagrange's identity for complex numbers [40, p. 30, Prob. 1.48] to conclude that

$$\|\mathbf{E} \times \mathbf{H}^*\|^2 = \|\mathbf{E}\|^2 \|\mathbf{H}\|^2 - |\mathbf{E}^{\mathrm{T}} \mathbf{H}|^2. \qquad (17)$$

From this identity and the elementary inequality $\|\mathbf{E}\|\|\mathbf{H}\| \leqslant \frac{1}{2}(\|\mathbf{E}\|^2 + \|\mathbf{H}\|^2)$, we obtain the following upper bound on the energy flux:

$$\|\mathrm{Re}\,\mathbf{S}\| \leqslant \frac{c}{16\pi}(\|\mathbf{E}\|^2 + \|\mathbf{H}\|^2). \qquad (18)$$

Next, we deal with the denominator in Eq. (2). It follows from the monotonicity property in Eq. (15) that the energy density U in Eq. (4) is well defined for any electromagnetic planewave with frequency ω in the transparency window (ω_1, ω_2) and satisfies

$$U = \frac{1}{16\pi} \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}^{\dagger} \frac{d}{d\omega} \omega \left[I + 4\pi\widehat{\chi}(\omega) \right] \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}$$

$$= \frac{1}{16\pi}(\|\mathbf{E}\|^2 + \|\mathbf{H}\|^2) + \frac{1}{4} \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}^{\dagger} \frac{dh}{d\omega}(\omega) \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}$$

$$\geqslant \frac{1}{16\pi}(\|\mathbf{E}\|^2 + \|\mathbf{H}\|^2). \qquad (19)$$

In particular, from this inequality it follows that the energy density is positive and so the energy velocity $\mathbf{v}_e$ defined by (2) does not blow up. Finally, from the two bounds (18) and (19), we deduce the upper bound on the energy velocity

$$\|\mathbf{v}_e\| = \frac{\|\mathrm{Re}\,\mathbf{S}\|}{U} \leqslant \frac{\frac{c}{16\pi}(\|\mathbf{E}\|^2 + \|\mathbf{H}\|^2)}{\frac{1}{16\pi}(\|\mathbf{E}\|^2 + \|\mathbf{H}\|^2)} = c. \qquad (20)$$

III. PERIODIC MEDIA

In this section, we prove the speed-of-light limitation in periodic passive linear media. This section is a strict generalization of Sec. II and as such the presentation here will follow the same outline. First, we review Bloch waves, group velocity, and energy velocity. Next, we set up the basic properties of susceptibility, define passivity, and then give a definition of a transparency window. Finally, we derive the speed-of-light limitation in a transparency window.

The main difference between this section and the special case, Sec. II, is that planewaves must be replaced with Bloch waves and, as such, the Fourier transform of the susceptibility in general no longer takes on values that are matrices, but instead its values are bounded linear operators on an infinite-dimensional Hilbert space (where, for generality, we include the possibility of spatially nonlocal operators). Nevertheless, the basic outline is the same even if the theorems are a bit more complicated.

A. Bloch waves and energy velocity

In a periodic medium, because of the discrete translational symmetry, Maxwell's equations admit nontrivial solutions which are time-harmonic electromagnetic Bloch waves [11,41]

$$\mathbf{E}(\mathbf{r})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{H}(\mathbf{r})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \qquad (21)$$

where $\mathbf{E}(\mathbf{r})$, $\mathbf{H}(\mathbf{r})$ are periodic functions on the lattice, with real frequency ω and real wavevector $\mathbf{k}$ (in a transparency window), satisfying a dispersion relation $\omega = \omega(\mathbf{k})$. The group velocity is $\mathbf{v}_g = \nabla_{\mathbf{k}}\omega(\mathbf{k})$ as before. In Sec. II A, the energy velocity was defined as the ratio of the energy flux to the energy density at a single point in space since all points were the same for a planewave, but now we must average over a unit cell, i.e., the energy velocity $\mathbf{v}_e$ of such a field is [11,26]

$$\mathbf{v}_e = \frac{\int_V \mathrm{Re}\,\mathbf{S}(\mathbf{r})d\mathbf{r}}{\int_V U(\mathbf{r})d\mathbf{r}}, \qquad (22)$$

where $\mathbf{S}, U$ are defined as in Sec. II A (but now are functions of position $\mathbf{r}$), V denotes a unit cell of the periodic medium, and the vector-valued functions $\mathbf{E}(\mathbf{r})$, $\mathbf{H}(\mathbf{r})$ are assumed to be square integrable in V. Just as in Sec. II A, the equality $\mathbf{v}_g = \mathbf{v}_e$ holds for transparent media, under certain reservations, a fact that we review in Appendix A [see (A20)]. As discussed in Appendix A, V can be unbounded in certain directions, for example, in a periodic waveguide, in which case the energy velocity is only defined in the periodic directions.

Now, we can handle any type of periodic linear media, including isotropic, anisotropic, bianisotropic, and dispersive media that are described by the same convolution form as (3), but with a spatial dependence on $\mathbf{r}$. As before, the convolution is defined in the distribution sense (as described

in Appendix B 3 a) except that now the susceptibility $\chi(t)$ is a distribution on a different set of test functions, namely, the smooth functions of time with compact support that take values in the Hilbert space of square-integrable vector-valued functions in the unit cell V. For such general materials, when the frequency ω belongs to a transparency window (as defined in Sec. III C), the energy density U is more complicated and is given in Appendix A by Eq. (A22). This is a generalization of the Brillouin formula (4) to arbitrary periodic passive linear media for frequencies in a transparency window.

The reason for the complication in the definition of the energy density U is that the Fourier transform of electromagnetic susceptibility $\widehat{\chi}(\omega)$ will be a function of frequency whose values are bounded linear operators in this infinite-dimensional Hilbert space. That just means the polarizations at a point $\mathbf{r}$ can be a function of $[\mathbf{E}(\mathbf{r}') \quad \mathbf{H}(\mathbf{r}')]^{\mathrm{T}}$ at other points in the unit cell. This generalization allows us to include responses nonlocal within the unit cell. As nonlocal responses are often used to describe metals at short length scales [42–49], this generalization could be used to describe, for instance, a periodic array of disjoint metallic objects. However, the most common case is that of local media, in which the polarizations at a point $\mathbf{r}$ only depend upon the field at the point $\mathbf{r}$. This means the susceptibility in the frequency domain is typically just matrix multiplication by $\widehat{\chi}(\mathbf{r},\omega)$, a 6×6 matrix-valued function of frequency ω and space $\mathbf{r}$ which is spatially periodic. In this case, the energy density is essentially identical to the Brillouin formula (4) except that the derivative $\frac{d}{d\omega}$ is replaced by the partial derivative $\frac{\partial}{\partial\omega}$ (since now there is a dependence on $\mathbf{r}$).

B. Passivity

In a periodic medium, just as in the homogeneous case, the key property of the susceptibility that will be used to derive the energy-velocity bound $\|\mathbf{v}_e\| \leqslant c$ and other properties is *passivity*. But, it turns out that we do not need to require passivity at every point individually as in (5), and instead we can require a weaker condition: that the medium is passive on average over the unit cell, i.e., *polarization currents do not do work on average within the unit cell V*. Mathematically, this means our condition is like that of (5), but with an integral over the unit cell as well as over time. More precisely, this is the following condition: the inequality

$$0 \leqslant \mathrm{Re} \int_{-\infty}^{t} \int_V \mathbf{E}(t')^{\dagger} \frac{d\mathbf{P}(t')}{dt'} + \mathbf{H}(t')^{\dagger} \frac{d\mathbf{M}(t')}{dt'} \, d\mathbf{r}\, dt' \quad (23)$$

(suppressing the spatial dependency $\mathbf{r}$) holds for all time t and all smooth functions of time $[\mathbf{E}(t) \quad \mathbf{H}(t)]^{\mathrm{T}}$ with compact support which take values in the Hilbert space of square-integrable vector-valued functions in the unit cell V.

Again, the key point here is that the convolution equation (3) and the passivity condition (23) mean that $\frac{d\chi}{dt}*$ is a "passive" convolution operator. The consequences are essentially the same as those described in Sec. II (and proved in Appendix B 3), with the only difference being in interpretation and proof, which is just due to the fact that the Hilbert space we are now working in is infinite dimensional. For instance, just like before (23) implies causality, i.e., $\chi(t) = 0$ for $t < 0$,

with the same technical caveat as previously mentioned. Also, its Fourier transform $\widehat{\chi}(\omega) = (\mathscr{F}\chi)(\omega)$ exists for $\mathrm{Im}\,\omega > 0$, but now it is an analytic bounded-operator-valued function. Similarly $h(\omega)$, defined as in Eq. (6), is a bounded-operator-valued Herglotz function, meaning it is analytic (in the operator norm topology) in the upper half-plane with (7) satisfied, i.e., the imaginary part of the operator is positive semidefinite (in the operator sense rather than the matrix sense) with respect to the inner product (B7). And, like before, a necessary and sufficient condition for $h(\omega) = \omega\widehat{\chi}(\omega)$ to be a Herglotz function is that $\frac{d\chi}{dt}*$ is a passive convolution operator, i.e., the passivity condition (23) holds.

C. Transparency window

A transparency window for a periodic passive linear media is still defined as a frequency interval $(\omega_1, \omega_2) \subseteq \mathbb{R}$ where losses are negligible [7, Secs. 80 and 84], but now electromagnetic losses are due to the transfer of electromagnetic energy into matter by absorption in the unit cell V, and a precise definition in terms of the imaginary part $\mathrm{Im}\,h(\omega)$ of the Herglotz function $h(\omega)$ is the same as in Sec. II C, i.e., (14) must be satisfied, the only difference is that the limits in Eq. (12), which still exist as proven in Appendix B 3 c, are now in the strong operator topology [39, Sec. VI.1] of the Hilbert space with inner product (B7). Similar to before, the two important properties of the function $h(\omega)$ in a transparency window (ω_1, ω_2) still continue to hold in this case, namely, $h(\omega)$ can be analytically continued (in operator norm topology) into the lower half-plane through this interval (ω_1, ω_2) to a bounded-operator-valued function satisfying the self-adjoint and monotonicity conditions in Eq. (15), where now analyticity, self-adjointness, and positive semidefiniteness is with respect to Hilbert space with inner product (B7). These properties of the function $h(\omega)$, along with the transparency window, all of which do not differ in form from Sec. II, are depicted in Fig. 1.

D. Speed-of-light limitation

We now will show that speed-of-light limitation

$$\|\mathbf{v}_e\| \leqslant c \tag{24}$$

holds for any time-harmonic electromagnetic Bloch wave provided the periodic medium is passive in its unit cell V and the frequency ω is in a transparency window (ω_1, ω_2). Deriving the speed-of-light limitation is done in essentially the same way as was done in Sec. II D.

First, we derive the upper bound on the numerator in Eq. (22) in exactly same way as (18) with one additional step, namely, we use the elementary inequality

$$\left\|\int_V \mathrm{Re}\,\mathbf{S}(\mathbf{r})d\mathbf{r}\right\| \leqslant \int_V \|\mathrm{Re}\,\mathbf{S}(\mathbf{r})\|\,d\mathbf{r}. \tag{25}$$

Second, deriving the lower bound on the denominator in Eq. (22) is done essentially in the same way as (19), except that now we must use the fact that $\frac{dh}{d\omega}(\omega) \geqslant 0$ is in the operator sense so that

$$\int_V U(\mathbf{r})d\mathbf{r} \geqslant \int_V \frac{1}{16\pi}[\|\mathbf{E}(\mathbf{r})\|^2 + \|\mathbf{H}(\mathbf{r})\|^2]d\mathbf{r}. \tag{26}$$

Again, this inequality means that the energy density is positive so that the energy velocity $\mathbf{v}_e$ given by (22) does not blow up. Finally, from the two bounds (25) and (26) we deduce the upper bound on the energy velocity

$$\|\mathbf{v}_e\| = \frac{\|\int_V \mathrm{Re}\,\mathbf{S}(\mathbf{r})d\mathbf{r}\|}{\int_V U(\mathbf{r})d\mathbf{r}} \leqslant c. \tag{27}$$

IV. ACHIEVING c

Here, we analyze the speed-of-light limitation $\|\mathbf{v}_e\| \leqslant c$ for necessary conditions on the fields and the media for the upper bound c to be achieved in a transparency window. As we shall show, $\|\mathbf{v}_e\| = c$ is never achieved except in vacuum or possibly in a few very special cases which we discuss.

We can see from the inequalities (17)–(20) and (25)–(27) that the upper bound $\|\mathbf{v}_e\| = c$ is achieved by a time-harmonic electromagnetic planewave (1) or Bloch wave (21) exactly when the numerator and denominator in Eqs. (20) or (27), respectively, achieve simultaneously the upper and lower bounds of the inequalities (18) and (19), respectively, in the case of a planewave or (25), (18), and (26), respectively, in the case of a Bloch wave. First, we consider under what circumstances the numerator achieves its upper bound, and we see a certain relationship between ϵ and μ (via the electric and magnetic parts of $\widehat{\chi}$) is required that is difficult to achieve outside of vacuum. Second, we examine when the denominator achieves its lower bound, and it turns out that this is only possible for a very specific form of the dispersion relation $\widehat{\chi}(\omega)$ which is proportional to $1/\omega$.

Necessary conditions for the numerator to achieve the upper bound are that the complex amplitude vector $F = [\mathbf{E} \quad \mathbf{H}]^{\mathrm{T}}$ of the corresponding planewave or Bloch wave satisfy the following three conditions:

$$\mathbf{E}^{\mathrm{T}}\mathbf{H} = 0, \quad \|\mathbf{E}\| = \|\mathbf{H}\|, \quad \mathrm{Im}\,(\mathbf{E}\times\mathbf{H}^*) = 0 \tag{28}$$

(almost everywhere in the unit cell), which follows from the identity (17) and the inequality (18). In the case of planewaves, the conditions (28) are both necessary and sufficient for the numerator in Eq. (20) to achieve the upper bound in Eq. (18). The conditions (28) are very restrictive: for instance, for planewave solutions (1) at a nonzero real wavevector $\mathbf{k}$ in a homogeneous isotropic medium, these conditions imply that the relative permittivity and permeability are equal. More complicated possibilities, including nontransverse solutions to (28), arise in anisotropic media. For example, one such solution is

$$\omega_0 = \frac{c}{3}\sqrt{2}, \quad \mathbf{k} = -\sqrt{2}\mathbf{e}_3, \quad \mathbf{E} = \sqrt{2}\mathbf{e}_2, \quad \mathbf{H} = \mathbf{e}_1 + \mathbf{e}_3, \tag{29}$$

$$\epsilon(\omega_0) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mu(\omega_0) = \begin{bmatrix} 5 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 2 \end{bmatrix}, \tag{30}$$

where $\mathbf{e}_j$, $j = 1,2,3$, denote the standard orthonormal basis vectors in $\mathbb{R}^3$.

We derive another constraint on the materials by considering (19) or (26) from the denominator in Eqs. (20) or (27), respectively. The lower bound in the inequalities (19) or (26), respectively, is achieved exactly when the complex

amplitude vector $F = [\mathbf{E} \quad \mathbf{H}]^{\mathrm{T}}$ of the wave in Eqs. (1) or (21), respectively, with frequency ω is in the null space of the operator $h'(\omega)$, i.e.,

$$h'(\omega)F = 0, \tag{31}$$

which follows from the monotonicity property (15) and the definition of the energy density (4) for homogeneous media or (A22) for periodic media in a transparency window. But, from a deep result in the theory of Herglotz functions (which we describe and prove in Appendix B 3 c), it follows that the equality (31) being true for a single frequency in a transparency window implies that (31) is true for *all* ω, real or complex, so that there exists a constant operator h_0 such that

$$h(\omega)F = h_0 F, \quad \mathrm{Im}\, h_0 = 0 \tag{32}$$

is satisfied for all ω. In terms of the Fourier transform of the susceptibility $\widehat{\chi}(\omega)$, we therefore must have

$$\widehat{\chi}(\omega)F = \frac{1}{\omega} h_0 F, \tag{33}$$

for *all* $\omega \neq 0$.

Equation (33) is a very strong, and rather unrealistic, constraint upon the material dispersion (especially since the susceptibility is usually argued to decay faster than $1/\omega$ as $\omega \to \infty$, on physical grounds [22, Sec. 7.5], [22, Sec. 7.10 (c)], [7, Sec. 78], or to prove the Kramers-Kronig relations by contour integration [22, Sec. 7.10 (d)], [7, Sec. 82], [50, Sec. 123]). It includes vacuum, of course ($h_0 = 0$), and can also be viewed as including perfect conductors as a limiting case. For example, a perfect electric conductor (PEC), which imposes a Dirichlet condition of vanishing tangential electric field, can be obtained by taking the limit $s \to \infty$ of an isotropic electric susceptibility $\frac{\epsilon-1}{4\pi} = \frac{s}{\omega}$ (see [51–53]). Inhomogeneous geometries consisting only of vacuum and PEC are known to support modes with $\|\mathbf{v}_e\| = c$, such as the TEM mode of a coaxial PEC waveguide [23, Sec. 9.20].

Equation (33) superficially resembles the susceptibility $i\sigma/\omega$ of a material with a constant scalar conductivity σ (see [22, Sec. 7.5 (c)]), but this case is excluded by the condition of a transparency window, which requires that h_0 be self-adjoint, i.e., $\mathrm{Im}\, h_0 = 0$. The coefficient h_0 can be further constrained if one imposes the usual condition that real fields produce real polarizations in the time domain [22, Sec. 7.10 (c)], [7, Sec. 82], [50, Sec. 123] which, in the case of homogeneous media, implies that h_0 must be of the form $h_0 = i\sigma$ where σ is a real antisymmetric matrix. For isotropic media, this implies that $h_0 = 0$, which is the vacuum case. For a nonchiral medium, it would mean that a nonzero h_0 would necessarily break electromagnetic reciprocity [54,55], [7, Sec. 89], [56,57] [due to $\widehat{\chi}(\omega) \neq \widehat{\chi}(\omega)^{\mathrm{T}}$]. For a chiral medium, it is possible to have such an $h_0 = i\sigma$ which is actually consistent with reciprocity [57]. Thus, a nonzero h_0 would mean if the medium is nonchiral, then it is nonreciprocal (having the form of a gyrotropic medium [7, Sec. 79, Prob. 1], [7, Sec. 101]) or it can be reciprocal, but then it is chiral.

V. VELOCITY IN LOSSY MEDIA

In the previous sections, we focused on the case of materials with negligible losses (specifically, a transparency window)

which corresponds to the situation in which group velocity is a meaningful energy velocity, but of course the problem of energy transport velocity in lossy media is also important. In lossy media, Poynting's theorem can be generalized to define a "dynamical" energy density [8,9,12,33–35], from which one can define a "dynamical" energy velocity via the ratio of the Poynting flux to the energy density [8,12]. Recently, Glasgow *et al.* [8] showed that this dynamical energy velocity is $\leqslant c$ in passive media, showed that it bounds the front velocity [2], [23, Sec. 5.18] (the speed at which a region of nonzero fields expands) by generalizing a result of Sommerfeld and Brillouin [2,58–60] via a derivation similar to that in [61, § 2.4.3], and also bounded the velocity of the mean energy-weighted position. In this section, we generalize and extend some of these results in several ways.

We begin in Sec. V A by reviewing the dynamical energy density and velocity constructions, and we delineate the types of susceptibilities for which the inequalities hold-unlike previous authors, we include distributions and furthermore generalize to the case of periodic and nonlocal media (as in Sec. III).

It is essential that the dynamical energy density reduce to the familiar dispersive energy density (4) in a transparency window, and for passive linear media this equivalence was shown by Yaghjian and Best [33], Yaghjian [9], and Figotin and Schenker [34,35]. The proof in Refs. [9,33], which applies to certain bianisotropic homogeneous media, does not include periodic and nonlocal media, nor does it specify the regularity of susceptibilities for which their argument applies. The proof of equivalence in Refs. [34,35] is rather cumbersome, requiring several pages and an extension of Maxwell's equations in lossy media to a conservative system by introducing auxiliary fields. It was also limited to square-integrable fields over all space (excluding Bloch waves, for example), and required more restrictive assumptions about the susceptibility (excluding distributions, for example). Therefore, in Sec. V B we provide an alternative proof, which is much shorter and (we believe) more intuitive and does not require the conservative extension, and applies to the full generality of fields and susceptibilities considered in this paper.

A. Dynamical energy density and velocity

We begin by reviewing the generalization of Poynting's theorem to arbitrary lossy media, from which the dynamical energy density is defined as in Ref. [8], but we do so in the context of our generalized susceptibility definitions of Secs. II and III.

Consider the macroscopic Maxwell's equations (in Gaussian units)

$$\boldsymbol{\nabla} \cdot \mathbf{D} = 4\pi\rho, \quad \boldsymbol{\nabla} \times \mathbf{H} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}, \tag{34}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0, \quad \boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0. \tag{35}$$

We are using complex fields which are not necessarily time harmonic so we want to derive Poynting's theorem [22, Sec. 6.8], i.e., the energy conservation law, for such fields in which the rate of doing work per unit volume by the fields

is taken to be

$$\text{Re}(\mathbf{J}^\dagger \mathbf{E}). \tag{36}$$

In order to exhibit this conservation law explicitly, we use Maxwell's equations to express (36) in other terms using the method from [22, Sec. 6.8]. First, we begin by taking the complex conjugate of both sides of the equality in Eq. (34) and take the dot product with $\mathbf{E}$ to yield

$$\mathbf{J}^\dagger \mathbf{E} = \frac{c}{4\pi}\mathbf{E}^{\text{T}}(\nabla \times \mathbf{H}^*) - \frac{1}{4\pi}\mathbf{E}^{\text{T}}\left(\frac{\partial \mathbf{D}}{\partial t}\right)^*. \tag{37}$$

Next, we use the vector identity

$$\nabla \cdot (a \times b) = b^{\text{T}}(\nabla \times a) - a^{\text{T}}(\nabla \times b) \tag{38}$$

with $a = \mathbf{E}$, $b = \mathbf{H}^*$, and use (35) to get that the right side of (37) is

$$\mathbf{J}^\dagger \mathbf{E} = -\nabla \cdot \mathbf{S} - \frac{1}{4\pi}\mathbf{H}^\dagger\frac{\partial \mathbf{B}}{\partial t} - \frac{1}{4\pi}\mathbf{E}^{\text{T}}\left(\frac{\partial \mathbf{D}}{\partial t}\right)^*, \tag{39}$$

where $\mathbf{S} = \frac{c}{4\pi}\mathbf{E} \times \mathbf{H}^*$. Finally, we take the real part of both sides of (39) to get

$$\text{Re}(\mathbf{J}^\dagger \mathbf{E}) = -\nabla \cdot \text{Re}\,\mathbf{S} - \frac{1}{4\pi}\text{Re}\left(\mathbf{H}^\dagger\frac{\partial \mathbf{B}}{\partial t} + \mathbf{E}^\dagger\frac{\partial \mathbf{D}}{\partial t}\right), \tag{40}$$

where the term $U(t)$ (suppressing spatial dependency) in

$$\frac{\partial U}{\partial t} = \frac{1}{4\pi}\text{Re}\left(\mathbf{H}^\dagger\frac{\partial \mathbf{B}}{\partial t} + \mathbf{E}^\dagger\frac{\partial \mathbf{D}}{\partial t}\right) \tag{41}$$

is interpreted to be a *dynamical energy density*. Notice that this result (40) is the standard form for the Poynting's theorem in differential form if we use real fields and assume the usual condition that real fields produce real polarizations in the time domain.

Now, we can express this as an energy conservation law in an integral form by

$$-\frac{\partial u}{\partial t} = \int_V \text{Re}(\mathbf{J}^\dagger \mathbf{E})d\mathbf{r} + \int_V \nabla \cdot \text{Re}\,\mathbf{S}\,d\mathbf{r}, \tag{42}$$

where V is the periodic unit cell,

$$\begin{aligned} u(t) &= \int_{-\infty}^t \int_V \frac{1}{4\pi}\text{Re}\left(\mathbf{H}^\dagger\frac{\partial \mathbf{B}}{\partial t'} + \mathbf{E}^\dagger\frac{\partial \mathbf{D}}{\partial t'}\right)d\mathbf{r}\,dt' \\ &= \frac{1}{8\pi}[\|\mathbf{E}(t)\|^2 + \|\mathbf{H}(t)\|^2] \\ &\quad + \text{Re}\int_{-\infty}^t \int_V \mathbf{E}^\dagger\frac{\partial \mathbf{P}}{\partial t'} + \mathbf{H}^\dagger\frac{\partial \mathbf{M}}{\partial t'}\,d\mathbf{r}\,dt', \end{aligned} \tag{43}$$

assuming sufficiently smooth fields which go to zero sufficiently rapidly as $t \to \pm\infty$ in the norm $\|\ldots\|$ with respect to the inner product (A4) (for homogeneous media we just drop the integral over V in the definitions). From the passivity property (23), it follows that the total *dynamical energy u* (in V) defined above is always a non-negative quantity and satisfies the inequality

$$u(t) \geqslant \frac{1}{8\pi}[\|\mathbf{E}(t)\|^2 + \|\mathbf{H}(t)\|^2] \tag{44}$$

for all $t \in \mathbb{R}$.

We define the *dynamical energy velocity* to be

$$\mathbf{v}_e(t) = \frac{\int_V \text{Re}\,\mathbf{S}(t)\,d\mathbf{r}}{u(t)}, \tag{45}$$

whenever $u(t) \neq 0$ for $t \in \mathbb{R}$, where

$$\mathbf{S}(t) = \frac{c}{4\pi}\mathbf{E}(t) \times \mathbf{H}(t)^*. \tag{46}$$

As we will now show, the speed-of-light limitation

$$\|\mathbf{v}_e(t)\| \leqslant c \tag{47}$$

holds for every $t \in \mathbb{R}$ in which $u(t) \neq 0$.

To prove this inequality, we first derive the upper bound on the numerator in Eq. (45) similar to the inequality (18) by using the inequalities (17) and (25). Second, the lower bound (44) as already been proven for the denominator in Eq. (45). Finally, from these bounds we derive the upper bounded on the dynamical energy velocity

$$\|\mathbf{v}_e(t)\| = \frac{\|\int_V \text{Re}\,\mathbf{S}(t)\,d\mathbf{r}\|}{u(t)} \leqslant c \tag{48}$$

for every $t \in \mathbb{R}$ in which $u(t) \neq 0$.

B. Connection to dispersive energy density and velocity in a transparency window

In this section, we will show that the dynamical energy (43) and velocity (45) reduce to the dispersive energy and velocity for a transparency window. We prove this result by adapting an approach from [15,16]: we construct a complex-frequency (exponentially growing) waveform that smoothly "turns off" after some time T (in order to have a well-defined polarization response), and then we take a limit as the frequency approaches a transparency window. Combined with the analyticity and self-adjointness of $h(\omega)$, we obtain the familiar Brillouin formula (involving $dh/d\omega$) for the dispersive energy density (4) and the energy velocity (2) for homogeneous media and their generalizations (A22) and (22) for periodic media.

Choose any $T \in \mathbb{R}$ and any $\lambda(t) \in C^\infty(\mathbb{R})$ satisfying

$$\lambda(t) = \begin{cases} 1, & \text{if } t < T \\ 0, & \text{if } T + 1 < t. \end{cases} \tag{49}$$

Define the function $\theta(t)$, for $t \in \mathbb{R}$, by

$$\theta(t) = e^{-i\omega t}\lambda(t), \quad \text{Im}\,\omega > 0. \tag{50}$$

Consider now the Hilbert space of all vector-valued functions which are square integrable in V with inner product $(\ldots,\ldots)$ defined in Eq. (A4) and let $F(\mathbf{r}) = [\mathbf{E}(\mathbf{r}) \quad \mathbf{H}(\mathbf{r})]^{\text{T}}$ be a nonzero element in this space. Then, for the fields

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}(\mathbf{r})\theta(t), \quad \mathbf{H}(\mathbf{r},t) = \mathbf{H}(\mathbf{r})\theta(t) \tag{51}$$

we can perform the time integral in Eq. (43), relying on the proofs of [15, Theorem 10.1] and [16, Theorem 8.12-1] that these integrals exist. This yields

$$\text{Re}\int_{-\infty}^t \int_V \mathbf{E}^\dagger\frac{\partial \mathbf{P}}{\partial t'} + \mathbf{H}^\dagger\frac{\partial \mathbf{M}}{\partial t'}\,d\mathbf{r}\,dt' = \frac{e^{2\,\text{Im}\,\omega t}}{2\,\text{Im}\,\omega}(F,\text{Im}\,h(\omega)F) \tag{52}$$

for all $t < T$. From this and the definition (43) we find for the fields (51) that

$$u(t) = \frac{e^{2\,\mathrm{Im}\,\omega t}}{8\pi}\left[(F,F) + \frac{4\pi}{\mathrm{Im}\,\omega}(F, \mathrm{Im}\,h(\omega)F)\right] \quad (53)$$

for all $t < T$ and is positive (at the end of this section, we will get an additional $\frac{1}{2}$ factor by time averaging).

From the formula (53) and the definition (45) we find for the fields (51) that the dynamical energy velocity is

$$\mathbf{v}_e(t) = \frac{\int_V \mathrm{Re}\,\mathbf{S}(t)\,d\mathbf{r}}{u(t)}$$

$$= \frac{\int_V \mathrm{Re}\left[\frac{c}{8\pi}\mathbf{E}(\mathbf{r}) \times \mathbf{H}(\mathbf{r})^*\right]d\mathbf{r}}{\frac{1}{16\pi}\left[(F,F) + \frac{4\pi}{\mathrm{Im}\,\omega}(F, \mathrm{Im}\,h(\omega)F)\right]} \quad (54)$$

for all $t < T$.

Now, though, if $\omega_0 \in (\omega_1, \omega_2)$ and $(\omega_1, \omega_2) \subseteq \mathbb{R}$ is a transparency window, then it follows that for $\omega = \omega_0 + i\eta$, $\eta > 0$ we have

$$\lim_{\eta \downarrow 0} \frac{1}{\mathrm{Im}\,\omega}\mathrm{Im}\,h(\omega) = h'(\omega_0), \quad (55)$$

with the limit converging in the operator norm topology. Hence, we find that

$$\lim_{\eta \downarrow 0}\mathbf{v}_e(t) = \frac{\int_V \mathrm{Re}\left[\frac{c}{8\pi}\mathbf{E}(\mathbf{r}) \times \mathbf{H}(\mathbf{r})^*\right]d\mathbf{r}}{\frac{1}{16\pi}\left[(F,F) + 4\pi(F, h'(\omega_0)F)\right]} \quad (56)$$

$$= \frac{\int_V \mathrm{Re}\left[\frac{c}{8\pi}\mathbf{E}(\mathbf{r}) \times \mathbf{H}(\mathbf{r})^*\right]d\mathbf{r}}{\left(F, \frac{1}{16\pi}\frac{d}{d\omega}\omega[I + 4\pi\widehat{\chi}(\omega)]\big|_{\omega=\omega_0}F\right)} \quad (57)$$

independent of t, provided $t < T$. The denominator here is in the form of the Brillouin formula for dispersive energy density. Now, all that remains is to verify the consequence for Bloch waves and time averaging.

Now, suppose we had a time-harmonic Bloch wave in Eq. (21) of the form

$$\mathbf{E}_{\mathbf{k}_0}(\mathbf{r})e^{i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)}, \quad \mathbf{H}_{\mathbf{k}_0}(\mathbf{r})e^{i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)}, \quad (58)$$

where $\mathbf{E}_{\mathbf{k}_0}(\mathbf{r})$, $\mathbf{H}_{\mathbf{k}_0}(\mathbf{r})$ are periodic functions on the lattice which are square integrable in the periodic unit cell V, with real frequency ω_0 in a transparency window, and real wavevector $\mathbf{k}_0$. Then, for the F above we can take $F(\mathbf{r}) = [\mathbf{E}_{\mathbf{k}_0}(\mathbf{r})e^{i\mathbf{k}_0 \cdot \mathbf{r}} \quad \mathbf{H}_{\mathbf{k}_0}(\mathbf{r})e^{i\mathbf{k}_0 \cdot \mathbf{r}}]^{\mathrm{T}}$. And so by the argument just given for this F and ω_0 we find that

$$\lim_{\eta \downarrow 0}\mathbf{v}_e(t) = \frac{\int_V \mathrm{Re}\left[\frac{c}{8\pi}\mathbf{E}_{\mathbf{k}_0}(\mathbf{r}) \times \mathbf{H}_{\mathbf{k}_0}(\mathbf{r})^*\right]d\mathbf{r}}{\int_V U(\mathbf{r})\,d\mathbf{r}} = \mathbf{v}_e, \quad (59)$$

independent of t, provided $t < T$, where U is defined by (A22) and $\mathbf{v}_e$ is the energy velocity (22) of this Bloch wave.

Using a similar approach above we will now prove that the dynamical energy (43) for the real parts of the fields (51) reduces after time averaging to the dispersive energy (A22) and (A14). In spirit, this is similar to the fact that for time-harmonic fields a and b the time average of their scalar product is $\langle \mathrm{Re}\,a^{\mathrm{T}}\mathrm{Re}\,b\rangle = \frac{1}{2}\mathrm{Re}\,(a^{\dagger}b)$, but some care is required here because of the exponential growth in time for complex ω. The proof requires we impose the usual condition that real fields produce real polarizations in the time domain [22, Sec. 7.10(c)], [7, Sec. 82], [50, Sec. 123]. Let $\Phi(t) =$

$\theta(t)F$, where as before $F(\mathbf{r}) = [\mathbf{E}(\mathbf{r}) \quad \mathbf{H}(\mathbf{r})]^{\mathrm{T}}$. Then, its polarization response is $[\mathbf{P}(t) \quad \mathbf{M}(t)]^{\mathrm{T}} = \chi * \Phi(t)$ and the latter condition implies that

$$[\mathrm{Re}\,\mathbf{P}(t) \quad \mathrm{Re}\,\mathbf{M}(t)]^{\mathrm{T}} = \mathrm{Re}\,[\chi * \Phi(t)] = \chi * \mathrm{Re}\,\Phi(t). \quad (60)$$

This implies the dynamical energy (43) for $\mathrm{Re}\,\Phi$, which we will denote by u_{real}, is given by

$$u_{\mathrm{real}}(t) = \frac{1}{8\pi}[\|\mathrm{Re}\,\mathbf{E}(t)\|^2 + \|\mathrm{Re}\,\mathbf{H}(t)\|^2]$$

$$+ \int_{-\infty}^{t}\int_V \mathrm{Re}\,\mathbf{E}^{\mathrm{T}}\mathrm{Re}\,\frac{\partial\mathbf{P}}{\partial t'} + \mathrm{Re}\,\mathbf{H}^{\mathrm{T}}\mathrm{Re}\,\frac{\partial\mathbf{M}}{\partial t'}\,d\mathbf{r}\,dt'. \quad (61)$$

Next, we use the vector identity

$$\mathrm{Re}\,a^{\mathrm{T}}\mathrm{Re}\,b = \frac{1}{2}\mathrm{Re}\,(a^{\dagger}b) + \frac{1}{2}\mathrm{Re}\,(a^{\mathrm{T}}b) \quad (62)$$

to write u_{real} in terms of the dynamical energy u of F:

$$u_{\mathrm{real}}(t) = \frac{1}{2}u(t) + \frac{1}{16\pi}\mathrm{Re}\int_V \mathbf{E}(t)^{\mathrm{T}}\mathbf{E}(t) + \mathbf{H}(t)^{\mathrm{T}}\mathbf{H}(t)d\mathbf{r}$$

$$+ \frac{1}{2}\mathrm{Re}\int_{-\infty}^{t}\int_V \mathbf{E}^{\mathrm{T}}\frac{\partial\mathbf{P}}{\partial t'} + \mathbf{H}^{\mathrm{T}}\frac{\partial\mathbf{M}}{\partial t'}\,d\mathbf{r}\,dt'. \quad (63)$$

Then, similar to the proofs of [15, Theorem 10.1] and [16, Theorem 8.12-1], it can be easily shown that

$$\mathrm{Re}\int_{-\infty}^{t}\int_V \mathbf{E}^{\mathrm{T}}\frac{\partial\mathbf{P}}{\partial t'} + \mathbf{H}^{\mathrm{T}}\frac{\partial\mathbf{M}}{\partial t'}\,d\mathbf{r}\,dt' = \mathrm{Im}\frac{e^{-2i\omega t}}{-2i\omega}(F^*, h(\omega)F) \quad (64)$$

for all $t < T$. From this identity and (63) it follows that

$$u_{\mathrm{real}}(t) = \frac{1}{2}u(t) + \frac{1}{16\pi}\mathrm{Re}\,e^{-2i\omega t}(F^*, F)$$

$$+ \frac{1}{2}\mathrm{Im}\frac{e^{-2i\omega t}}{-2i\omega}(F^*, h(\omega)F) \quad (65)$$

for all $t < T$, where $u(t)$ is given in terms of F by (53).

Now, though, if $\omega_0 \in (\omega_1, \omega_2)$, $\omega_0 \neq 0$, and $(\omega_1, \omega_2) \subseteq \mathbb{R}$ is a transparency window, then it follows from (53), (55), and (65) that for $\omega = \omega_0 + i\eta$, $\eta > 0$ we have

$$\lim_{\eta \downarrow 0}u_{\mathrm{real}}(t) = \frac{1}{16\pi}[(F,F) + 4\pi(F, h'(\omega_0)F)]$$

$$+ \frac{1}{16\pi}\mathrm{Re}\,e^{-2i\omega_0 t}(F^*, F)$$

$$+ \frac{1}{2}\mathrm{Im}\frac{e^{-2i\omega_0 t}}{-2i\omega_0}(F^*, h(\omega_0)F), \quad (66)$$

provided $t < T$. Letting $\langle f\rangle$ denote the time average of a periodic function f with period $2\pi/\omega_0$ over a time interval of length $2\pi/\omega_0$ contained in $(-\infty, T)$, we find that

$$\left\langle \lim_{\eta \downarrow 0}u_{\mathrm{real}}(t)\right\rangle = \frac{1}{16\pi}[(F,F) + 4\pi(F, h'(\omega_0)F)]$$

$$= \left(F, \frac{1}{16\pi}\frac{d}{d\omega}\omega[I + 4\pi\widehat{\chi}(\omega)]\big|_{\omega=\omega_0}F\right). \quad (67)$$

For homogeneous media, (67) is just the familiar Brillouin formula (4) for homogeneous media and, for periodic media, it

is the generalization using the inner product $(\ldots, \ldots)$ defined in Eq. (A4). In particular, if we had a time-harmonic Bloch wave in the form (58), then the formula (67) just becomes

$$\left\langle \lim_{\eta \downarrow 0} u_{\text{real}}(t) \right\rangle = \int_V U(\mathbf{r}) \, d\mathbf{r}, \tag{68}$$

where U is defined in Eq. (A22).

VI. CONCLUSION

In this paper, we have generalized the notion of linear media and have shown that periodic passive linear media imply the following: the polarization response obeys causality, i.e., the electromagnetic susceptibility $\chi(t)$ (up to an additive constant which we argue is zero on physical grounds) satisfies $\chi(t) = 0$ for $t < 0$ (in the distributional sense), existence of its Fourier transform $\widehat{\chi}(\omega)$ in the upper half-plane (i.e., Im $\omega > 0$), the function $h(\omega) = \omega \widehat{\chi}(\omega)$ is a bounded operator-valued Herglotz function (i.e., an analytic bounded-operator-valued function in the upper half-plane with a positive-semidefinite imaginary part), and electromagnetic losses can be quantified in terms of a non-negative operator-valued measure Ω on the bounded Borel subsets of $\mathbb{R}$ defined in terms of a limit (as we approach the real axis from the upper half-plane) of an integral involving the imaginary part Im $h(\omega)$. We have also shown that in a transparency window $(\omega_1, \omega_2) \subseteq \mathbb{R}$ (i.e., frequency window of negligible losses), which we defined in terms of that measure as $\Omega((\omega_1, \omega_2)) = 0$, the Herglotz function $h(\omega)$ has an analytic continuation into the lower half-plane through the transparency window such that $h(\omega)$ is a monotonically increasing [i.e., $h'(\omega) \geqslant 0$] bounded-operator-valued function which is self-adjoint [i.e., Im $h(\omega) = 0$] in the window. From this monotonicity property, we showed that the energy density of any planewave or Bloch wave is always positive and the energy velocity $\mathbf{v}_e$ of such a wave is always bounded above by c (i.e., $\|\mathbf{v}_e\| \leqslant c$), the speed of light in a vacuum. Furthermore, we showed that almost never is the upper bound c achieved (i.e., $\|\mathbf{v}_e\| = c$), except in vacuum or in a few special cases, but if it is achieved then the dispersion relation $\widehat{\chi}(\omega)$ must be of a very specific form which is proportional to $1/\omega$, the latter of which is a very strong and unrealistic constraint upon the material dispersion since it typically decays faster then $1/\omega$ as $\omega \to \infty$. For lossy media, we generalized the result of Glasgow *et al.* [8] on the dynamical energy velocity $\mathbf{v}_e(t)$ being subluminal in passive media to include susceptibilities that are distributions as well as periodic and nonlocal media. Moreover, we have shown that dynamical energy density $U(t)$ and velocity $\mathbf{v}_e(t)$ reduce, in the limit to a transparency window, to the familiar Brillouin formula for energy density and velocity.

Although much has been said in this paper on bounding the energy velocity by c, there are still some major open questions and problems. First, how should one handle spatial nonlocality across unit-cell boundaries? This is clearly going to involve imposing some constraints on the type of nonlocality that would be allowed. Second, how should one determine which of the many different types of well-known velocities for light [2,62,63], [64, Sec 2.9] such as the front velocity [2], [23, Sec. 5.18], [64, Sec. 2.1] (sometimes referred to as the signal velocity, see [62]) are also bounded by c in passive linear media? Certain progress has been made on the front velocity;

for instance, it is known that the total dynamical energy density in passive anisotropic local media has a front velocity bound by c [8]. Moreover, since the work of Sommerfeld and Brillouin [2,58–60] on signal velocity (for a modern definition of a "signal" see, for instance, [64, Secs. 2.9 and 4.5], [65, Sec. 8]), there has been significant mathematical development in the theory of hyperbolic partial differential equations (PDEs) relating to the wave-front velocity, propagation of singularites, and finite speed of propagation (see, for instance, [66]) which may be effective in the study of subliminal velocities of light in dissipative and dispersive linear media. Finally, how should one derive a similar type of velocity bound, if possible, in nonlinear or time-varying media?

APPENDIX A: ENERGY VELOCITY AND GROUP VELOCITY OF ELECTROMAGNETIC BLOCH WAVES IN PERIODIC MEDIA

In this Appendix, we study the relationship between the energy velocity and group velocity of electromagnetic Bloch waves in periodic passive linear media in a transparency window, and show they are the same in local media.

The equality of energy velocity and group velocity in media with negligible loss was previously shown for homogeneous dispersive media [24], [21, Chap. VI.31], [2, Sec. IV.5], [22, Sec. 7.1], [7, Sec. 83], [25]; for periodic nondispersive media [26, Appendix B], [22, Sec. 8.5], [11, p. 41]; for periodic dispersive media [27–31]; and for spatial-temporal periodic media [32].

In this Appendix, we combine and generalize these results to include some cases that were not handled previously, including periodic bianisotropic media and susceptibilities that are generalized functions. We also discuss the case of nonlocal media (within the unit cell), for which energy and group velocity are not in general equal. Our extension of previous results is straightforward, essentially using the variational approach of [28,29], [26, Appendix B], [11, p. 41] based on a generalization of the Hellmann-Feynman theorem of quantum mechanics [67,68], [69, Sec. 11, Eq. (11.16)], [70, Theorem 7.3.1] which in essence is just a result of first-order perturbation theory for self-adjoint linear operators [71–74].

In the analysis that follows, we assume the periodic media is passive in a periodic unit cell V with a transparency window $(\omega_1, \omega_2) \subseteq \mathbb{R}$ and restrict ourselves to frequencies ω in this range. For a unified treatment, the unit cell V is a measurable set which may be unbounded in directions that are not periodic (as in the case of a periodic waveguide), and may be lower dimensional in cases of continuous translational invariance. For instance, if there are certain directions of continuous translational invariance (such as for homogeneous media),

we can treat these as having "period" zero (reducing the dimensionality of V).

We consider time-harmonic electromagnetic Bloch waves

$$\mathbf{E} = \mathbf{E_k(r)}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{H} = \mathbf{H_k(r)}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad (A1)$$

where $\mathbf{E_k(r)}$, $\mathbf{H_k(r)}$ are periodic functions on the lattice, with real frequency $\omega \in (\omega_1, \omega_2)$ and real wavevector $\mathbf{k} \in \mathbb{R}^3$ (for $1d$ or $2d$ periodic media, the components of $\mathbf{k}$ are zero in nonperiodic directions) such that $[\mathbf{E_k} \quad \mathbf{H_k}]^{\mathrm{T}} \in [L^2(V)]^6$ (or $\mathbb{C}^6$ in the case of homogeneous media). For certain nonzero frequencies $\omega = \omega(\mathbf{k})$, the Bloch waves are solutions of Maxwell's equations in the frequency domain

$$\begin{bmatrix} 0 & \nabla\times \\ -\nabla\times & 0 \end{bmatrix}\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix} = -\frac{i\omega}{c}\begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix} + 4\pi\begin{bmatrix} \mathbf{P} \\ \mathbf{M} \end{bmatrix}, \quad (A2)$$

where $\mathbf{P}$ and $\mathbf{M}$ are given in terms of the susceptibility by

$$\begin{bmatrix} \mathbf{P} \\ \mathbf{M} \end{bmatrix} = \widehat{\chi}(\omega)\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix} \quad (A3)$$

(the conditions $\nabla\cdot\mathbf{D} = 0$, $\nabla\cdot\mathbf{B} = 0$ are automatically satisfied since $\nabla\cdot\nabla\times = 0$ should hold for the electromagnetic fields being considered.)

On the Hilbert space $\mathfrak{H} = [L^2(V)]^6$ with inner product $(\ldots,\ldots)$ given by

$$(\psi,\varphi) = \int_V \psi(\mathbf{r})^{\dagger}\varphi(\mathbf{r})d\mathbf{r}, \quad \psi,\varphi \in \mathfrak{H} \quad (A4)$$

[or, in the case of homogeneous media, $\mathfrak{H} = \mathbb{C}^6$ with inner product $(\ldots,\ldots)$ given by $(\psi,\varphi) = \psi^{\dagger}\varphi$], the operator $h(\omega) = \omega\widehat{\chi}(\omega)$, because of the hypotheses of passivity and transparency window, has the properties that it is a bounded-operator-valued function which is analytic in a region containing (ω_1,ω_2) and self-adjoint for any $\omega \in (\omega_1,\omega_2)$ satisfying $h'(\omega) \geqslant 0$ for every $\omega \in (\omega_1,\omega_2)$. We introduce the operator $\mathcal{T}(\mathbf{k})$ defined on $\mathfrak{H}$ by multiplication

$$[\mathcal{T}(\mathbf{k})\psi](\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\psi(\mathbf{r}), \quad \psi \in \mathfrak{H}. \quad (A5)$$

For this type of geometry and these types of solutions, the curl operator satisfies

$$\nabla\times = \mathcal{T}(\mathbf{k})\mathcal{A}(\mathbf{k})\mathcal{T}(-\mathbf{k}), \quad \text{where} \quad \mathcal{A}(\mathbf{k}) = \nabla\times + i\mathbf{k}\times. \quad (A6)$$

Thus, Eqs. (A2) and (A3) become

$$\mathcal{M}(\omega,\mathbf{k})\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix} = \mathbf{0}, \quad \text{where} \quad \mathcal{M}(\omega,\mathbf{k}) = \mathcal{M}_0 + \mathcal{M}_1(\omega,\mathbf{k}), \quad (A7)$$

$$\mathcal{M}_0 = \begin{bmatrix} 0 & -\frac{ic}{16\pi}\nabla\times \\ \frac{ic}{16\pi}\nabla\times & 0 \end{bmatrix},$$

$$\mathcal{M}_1(\omega,\mathbf{k}) = \begin{bmatrix} 0 & \frac{c}{16\pi}\mathbf{k}\times \\ -\frac{c}{16\pi}\mathbf{k}\times & 0 \end{bmatrix}$$

$$+ \frac{\omega}{16\pi}I + \frac{1}{4\pi}\mathcal{T}(-\mathbf{k})h(\omega)\mathcal{T}(\mathbf{k}). \quad (A8)$$

Suppose now that at some real point $(\omega_0,\mathbf{k}_0)$ in $(\omega_1,\omega_2)\times\mathbb{R}^3$ and for variations $\mathbf{k} = \mathbf{k}_0 + \delta\mathbf{k}$, $\|\delta\mathbf{k}\| \ll 1$, along the periodic directions that there exists a continuous function $\omega = \omega(\mathbf{k})$ and family of Bloch wave solutions to (A7) satisfying $\omega = \omega_0 + o(1)$, $[\mathbf{E_k} \quad \mathbf{H_k}]^{\mathrm{T}} = [\mathbf{E_{k_0}} \quad \mathbf{H_{k_0}}]^{\mathrm{T}} + o(1)$ (in the norm topology of $\mathfrak{H}$) as $\|\delta\mathbf{k}\| \to 0$. We will now prove that the total differential $d\omega$ exists and derive a formula for the gradient

$$\nabla_{\mathbf{k}}\omega(\mathbf{k}_0) = \sum_j \frac{\partial\omega}{\partial k_j}(\mathbf{k}_0)\,\mathbf{e}_j, \quad (A9)$$

where $\mathbf{e}_i$, $i = 1,2,3$, denote the standard basis vectors in $\mathbb{R}^3$ and the sum is over the j in the periodic directions. For typical boundary conditions, the usual integration by parts formula for these types of periodic solutions to (A7) hold, i.e.,

$$\left(\mathcal{M}_0\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right) - \left(\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \mathcal{M}_0\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right) = 0, \quad (A10)$$

and so it follows by our assumptions that

$$\left(\mathcal{M}(\omega,\mathbf{k})\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right) - \left(\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \mathcal{M}(\omega_0,\mathbf{k}_0)\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)$$

$$= \left(\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \delta\mathcal{M}_1\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right) = 0, \quad (A11)$$

where

$$\delta\mathcal{M}_1 = \mathcal{M}_1(\omega,\mathbf{k}) - \mathcal{M}_1(\omega_0,\mathbf{k}_0)$$

$$= \delta\omega\frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0) + \sum_j \delta k_j\frac{\partial\mathcal{M}_1}{\partial k_j}(\omega_0,\mathbf{k}_0)$$

$$+ o(\delta\omega) + o(\|\delta\mathbf{k}\|) \quad (A12)$$

(in the operator norm topology of $[\mathfrak{H};\mathfrak{H}]$, the space of continuous linear operators on $\mathfrak{H}$) as $\|\delta\mathbf{k}\| \to 0$. From this, our variation assumptions, and the Schwarz inequality [39] for the inner product $(\ldots,\ldots)$ it follows that

$$0 = \left(\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \delta\mathcal{M}_1\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)$$

$$= \left(\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}, \delta\mathcal{M}_1\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right) + \left(\delta\begin{bmatrix} \mathbf{E_k} \\ \mathbf{H_k} \end{bmatrix}, \delta\mathcal{M}_1\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)$$

$$= \delta\omega\left(\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0)\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)$$

$$+ \sum_j \delta k_j\left(\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial k_j}(\omega_0,\mathbf{k}_0)\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)$$

$$+ o(\delta\omega) + o(\|\delta\mathbf{k}\|) \quad (A13)$$

as $\|\delta\mathbf{k}\| \to 0$. It follows from this and the fact

$$\frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0) = \frac{1}{16\pi}I + \frac{1}{4\pi}\mathcal{T}(-\mathbf{k}_0)h'(\omega_0)\mathcal{T}(\mathbf{k}_0) \geqslant \frac{1}{16\pi}I \quad (A14)$$

that the total differential $d\omega$ exists and is given by the formula

$$d\omega = \sum_j \frac{-\left(\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial k_j}(\omega_0,\mathbf{k}_0)\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)}{\left(\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0)\begin{bmatrix} \mathbf{E_{k_0}} \\ \mathbf{H_{k_0}} \end{bmatrix}\right)}dk_j. \quad (A15)$$

From the vector identity

$$
\mathrm{Re}\left(\mathbf{E}\times\mathbf{H}^*\right) = -\frac{1}{2}\sum_{j=1}^{3}\begin{bmatrix}\mathbf{E}\\\mathbf{H}\end{bmatrix}^{\dagger}\begin{bmatrix}0 & \mathbf{e}_j\times\\-\mathbf{e}_j\times & 0\end{bmatrix}\begin{bmatrix}\mathbf{E}\\\mathbf{H}\end{bmatrix}\mathbf{e}_j,
$$

(A16)

it follows that the numerator is given by the formula

$$
-\left(\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial k_j}(\omega_0,\mathbf{k}_0)\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}\right)
$$
$$
= \int_V \mathrm{Re}\,\mathbf{S}(\mathbf{r})\cdot\mathbf{e}_j d\mathbf{r} - 4\pi\left(\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}, \frac{\partial h}{\partial k_j}(\omega_0,\mathbf{k}_0)\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}\right),
$$

(A17)

where $\mathbf{S} = \frac{c}{8\pi}\mathbf{E}_{\mathbf{k_0}}\times\mathbf{H}_{\mathbf{k_0}}^*$ is the spatially dependent complex Poynting vector and

$$
h(\omega,\mathbf{k}) = \mathcal{T}(-\mathbf{k})h(\omega)\mathcal{T}(\mathbf{k}).
$$

(A18)

By the formula for the total differential $d\omega$ it follows immediately that

$$
\frac{\partial\omega}{\partial k_j}(\mathbf{k}_0) = \frac{-\left(\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial k_j}(\omega_0,\mathbf{k}_0)\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}\right)}{\left(\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}, \frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0)\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}\right)}.
$$

(A19)

For local media, $\mathcal{T}(\mathbf{k})$ and $h(\omega)$ commute so that the second term on the right-hand side of (A17) is zero (for generic nonlocal media this will not be true) and hence it follows from the derived equalities that the group velocity is given by the formulas

$$
\frac{\partial\omega}{\partial k_j}(\mathbf{k}_0) = \frac{\int_V \mathrm{Re}\,\mathbf{S}(\mathbf{r})\cdot\mathbf{e_j}d\mathbf{r}}{\int_V U(\mathbf{r})d\mathbf{r}},
$$

(A20)

where

$$
\mathbf{S} = \frac{c}{8\pi}\mathbf{E}_{\mathbf{k_0}}\times\mathbf{H}_{\mathbf{k_0}}^*,
$$

(A21)

$$
U = \begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix}^{\dagger}\frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0)\begin{bmatrix}\mathbf{E}_{\mathbf{k_0}}\\\mathbf{H}_{\mathbf{k_0}}\end{bmatrix},
$$

(A22)

where $\frac{\partial\mathcal{M}_1}{\partial\omega}(\omega_0,\mathbf{k}_0)$ is given by the formulas (A14) and (A5).

APPENDIX B: PROOFS

In this Appendix, we prove the rest of statements in this paper from Secs. II B, II C, III B, III C, and IV that were left to be proved here.

This section is organized as follows. First, in Sec. B 1 we give our notations, conventions, and definitions that will be needed in our proofs. Next, in Sec. B 2 we state and prove two results on antiderivatives of certain classes of distributions which will be needed in the next section. Finally, Sec. B 3 contains the main body of this appendix, namely, the proofs of the statements in this paper from the sections mentioned above.

1. Notations, conventions, and definitions

Unless otherwise indicated, we will adhere to the notation, conventions, and definitions in Ref. [16]. The following

modifications and additional definitions and notations should be noted.

First, any Hilbert space will be denoted by $\mathfrak{H}$ with the convention that the inner product $(\ldots,\ldots)$ is antilinear in first vector coordinate rather then the second. As in Ref. [16, p. 3], $[\mathfrak{H};\mathfrak{H}]$ denotes the space of all continuous linear operators on $\mathfrak{H}$.

Second, we will use the notation $\mathbb{R}$ to denote the set of real numbers as opposed to R. Thus, as in Ref. [16, pp. 50 and 52], $\mathscr{D}(\mathfrak{H})$ denotes the space of all infinitely differentiable functions $\phi : \mathbb{R} \to \mathfrak{H}$ with compact support and $[\mathscr{D}(\mathfrak{H});\mathfrak{H}]$ denotes the $[\mathfrak{H};\mathfrak{H}]$-valued distributions (or generalized functions), that is, the set of all continuous linear functions $f : \mathscr{D}(\mathfrak{H}) \to \mathfrak{H}$.

Third, as opposed to the definition in Ref. [16, p. 66 and Sec. 8.4], the Fourier transform $\mathscr{F}$ of any $\phi \in \mathscr{S}(\mathfrak{H})$ (the $\mathfrak{H}$-valued testing functions of rapid descent) will be denoted by $\widehat{\phi} = \mathscr{F}\phi$ and defined by

$$
\widehat{f}(\omega) = (\mathscr{F}\phi)(\omega) = \int_{-\infty}^{\infty}\phi(t)e^{i\omega t}dt, \quad \omega \in \mathbb{R}.
$$

(B1)

Fourth, using the definition in Ref. [16, Sec. 6.2] of the Laplace transform on distributions belonging to the generalized function space $[\mathscr{L}(0,\infty;\mathfrak{H});\mathfrak{H}]$, the Fourier transform on such distributions can be defined. In particular, for a Laplace-transformable distribution $y \in [\mathscr{D}(\mathfrak{H});\mathfrak{H}]$ with strip of definition for the Laplace transform of y containing $C_+ = \{\zeta \in \mathbb{C} : \mathrm{Re}\,\zeta > 0\}$ then this means $y \in [\mathscr{L}(0,\infty;\mathfrak{H});\mathfrak{H}]$ and it has a Laplace transform $\mathfrak{L}y : C_+ \to [\mathfrak{H};\mathfrak{H}]$ as defined in Ref. [16, Sec. 6.2, Eq. (2)]. From the Laplace transform, the Fourier transform of y denoted by $\widehat{y} = \mathscr{F}y$ is defined as the function

$$
\widehat{y}(\omega) = (\mathscr{F}y)(\omega) = (\mathfrak{L}y)(-i\omega), \quad \omega \in \mathbb{C}_+
$$

(B2)

where $\mathbb{C}_+ = \{\omega \in \mathbb{C} : \mathrm{Im}\,\omega > 0\}$.

Finally, a function $h : \mathbb{C}_+ \to [\mathfrak{H};\mathfrak{H}]$ is said to be a Herglotz function if it is analytic on $\mathbb{C}_+$ such that $\mathrm{Im}\,h(\omega) \geqslant 0$ for every $\omega \in \mathbb{C}_+$, where Im denotes the imaginary part of an operator from $[\mathfrak{H};\mathfrak{H}]$ (i.e., for any $A \in [\mathfrak{H};\mathfrak{H}]$, $\mathrm{Im}\,A = \frac{1}{2i}(A - A^{\dagger})$, where $A^{\dagger}$ denotes the adjoint of A).

2. On antiderivatives of certain distributions

In this section, we state and prove certain results on antiderivatives of distributions that will be needed in Sec. B 3. Although the results here seem to be known, we have not been able to find references to them in the literature and hence for completeness we prove them here. The two key results below are (i) and (ii). The result (i) says that for a given $[\mathfrak{H};\mathfrak{H}]$-valued distribution f with support in $[0,\infty)$, i.e., $f(t) = 0$ for $t < 0$ in the distributional sense [16, p. 55], there exists a unique antiderivative of f with the same property. Result (ii) is almost identical to (i) except the hypothesis and conclusion that f and its antiderivative are in $[\mathscr{D}(\mathfrak{H});\mathfrak{H}]$ is instead replaced by the subspace $[\mathscr{L}(0,\infty;\mathfrak{H});\mathfrak{H}]$.

To begin, recall that the generalized derivative $D : [\mathscr{D}(\mathfrak{H});\mathfrak{H}] \to [\mathscr{D}(\mathfrak{H});\mathfrak{H}]$ on distributions is defined as in

Ref. [16, p. 54] by

$$\langle Df, \phi \rangle = -\langle f, D\phi \rangle, \quad \phi \in \mathscr{D}(\mathfrak{H}). \tag{B3}$$

Now, an antiderivative (or primitive) of a distribution $f \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$ is any $g \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$ such that $Dg = f$. Following virtually the same argument from [16, Sec. 5.6], one can easily show that every distribution has an antiderivative and any two antiderivatives differ by a constant distribution [i.e., if $Dg_1 = Dg_2$ for some $g_1, g_2 \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$, then there exists an $A \in [\mathfrak{H}; \mathfrak{H}]$ such that $g_1 - g_2 = A$, where $\langle A, \phi \rangle = \int_{-\infty}^{\infty} A\phi(t)dt$ for all $\phi \in \mathscr{D}(\mathfrak{H})$].

We will need the following two facts in this paper on antiderivatives of certain distributions which we will prove below: (i) if $f \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$ and supp $f \subseteq [0, \infty)$ then there exists a unique antiderivative $f^{(-1)}$ of f such that supp $f^{(-1)} \subseteq [0, \infty)$; (ii) if f is a distribution in $[\mathscr{L}(0, \infty; \mathfrak{H}); \mathfrak{H}]$ and supp $f \subseteq [0, \infty)$, then there exists a unique antiderivative $f^{(-1)}$ of f such that $f^{(-1)} \in [\mathscr{L}(0, \infty; \mathfrak{H}); \mathfrak{H}]$ and supp $f^{(-1)} \subseteq [0, \infty)$.

We prove these two statements now. We begin with the proof of statement (i) following the argument in Ref. [16, Sec. 5.6] on the existence of an antiderivative. Suppose $f \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$ and supp $f \subseteq [0, \infty)$. Let H denote the subspace of $\mathscr{D}(\mathfrak{H})$ whose elements φ have the form $\varphi = D\psi$, where $\psi \in \mathscr{D}(\mathfrak{H})$. Let $\phi_0 \in \mathscr{D}(\mathbb{C})$ be such that $\int_{-\infty}^{\infty} \phi_0(t)dt = 1$ and supp $\phi_0 \subseteq (-\infty, 0)$. Then, every $\phi \in \mathscr{D}(\mathfrak{H})$ has the unique decomposition

$$\phi = \phi_0 w + \varphi, \tag{B4}$$

where $w = \int_{-\infty}^{\infty} \phi(t)dt$ and $\varphi \in H$. In particular, $\varphi = D\psi$, where $\psi(t) = \int_{-\infty}^{t} \varphi(t)dt$ and $\psi \in \mathscr{D}(\mathfrak{H})$. Similar to [16, Sec. 5.6, Eq. (2)], one defines a distribution $f^{(-1)} \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$ in terms of the decomposition (B4) of $\phi \in \mathscr{D}(\mathfrak{H})$ by

$$\langle f^{(-1)}, \phi \rangle = -\langle f, \psi \rangle. \tag{B5}$$

From this definition and (B3), it follows immediately that $f^{(-1)}$ is an antiderivative of f and if $\phi \in \mathscr{D}(\mathfrak{H})$ and supp $\phi \subseteq (-\infty, 0)$, then supp $\psi \subseteq (-\infty, 0)$ so that since supp $f \subseteq [0, \infty)$ this implies $0 = -\langle f, \psi \rangle = \langle f^{(-1)}, \phi \rangle$. This proves that supp $f^{(-1)} \subseteq [0, \infty)$. The proof of uniqueness now just follows immediately from the facts two antiderivatives of a distribution differ by a constant distribution and the only constant distribution A with supp $A \subseteq [0, \infty)$ is $A = 0$. This completes the proof of statement (i).

We now prove statement (ii). Suppose f is a distribution in $[\mathscr{L}(0, \infty; \mathfrak{H}); \mathfrak{H}]$ and supp $f \subseteq [0, \infty)$. Then, it follows from [16, Theorem 6.4-2; 6.5-1] and [16, Appendix E, E2] that there exists a unique Laplace-transformable distribution $f^{(-1)}$ in $[\mathscr{L}(0, \infty; \mathfrak{H}); \mathfrak{H}]$ with supp $f^{(-1)} \subseteq [0, \infty)$ such that its Laplace transform $\mathfrak{L}f^{(-1)} : C_+ \to [\mathfrak{H}; \mathfrak{H}]$ is given by

$$(\mathfrak{L}f^{(-1)})(\zeta) = \frac{1}{\zeta}(\mathfrak{L}f)(\zeta), \quad \zeta \in C_+. \tag{B6}$$

It follows immediately from this and [16, Prob. 6.3-2] that $(\mathfrak{L}Df^{(-1)})(\zeta) = (\mathfrak{L}f)(\zeta)$ for every $\zeta \in C_+$ which, by the uniqueness of the Laplace transform [16, Theorem 6.4-2], implies $f^{(-1)}$ is an antiderivative of f. This proves the existence portion of statement (ii). The uniqueness portion follows immediately from statement (i). This completes the proof of statement (ii).

3. Proofs of main results

This section contains the main body of this appendix. Here, we prove the statements from Secs. II B, II C, III B, III C, and IV that were left to be proved. Our proofs are essentially based on a translation of the electromagnetic assumptions of passive linear media and transparency window into the distributional approach of Zemanian on the theory of passive linear systems [16, Chap. 8] and connecting all of this to the theory of Herglotz functions (see, for instance, [17,19,20] and [37, Appendix A], and references within).

This section is organized as follows. We begin in Sec. B 3 a by recasting the polarization response (3) and passivity, either (5) or (23), in terms of the theory of distributions and passive linear systems from [16, Chap. 8] using the susceptibility χ, a generalized function. Then, in Sec. B3b we prove the statements from Secs. II B and III B, in particular, that linearity (3) and passivity (5) and (23) implies the three main consequences of causality, analyticity, and positivity. Our proof though essentially follows from [16, Theorem 8.12-1] and [16, Lemma 7.2-1], the existence of antiderivatives of distributions as discussed in Sec. B 2, the relationship between the Laplace and Fourier transforms as discussed in Sec. B 1, and the relationship between positive* mappings [16, Def. 8.11-1] and Herglotz functions (defined in Sec. B 1). Finally, in Sec. B3c we prove the rest of statements that were left to be proved here from Secs. II C, III C, and IV. In particular, we give a precise and rigorous definition of a transparency window using the Herglotz function $h(\omega) = \omega\widehat{\chi}(\omega)$ from (6) and prove the main consequences of a transparency window (ω_1, ω_2), namely, analytic continuation of $h(\omega)$ across (ω_1, ω_2) to a bounded-operator-valued function satisfying the self-adjoint and monotonicity conditions (15). We also prove in Sec. B3c the fact mentioned in Sec. IV on achieving the upper bound $\|\mathbf{v}_e\| = c$ that the condition (31) being true for h at a single frequency in a transparency window implies that (31) and (32) are true for all frequencies. Our proofs of all these results are essentially just based on the connection between the well-known integral representation of a Herglotz function from [19, Theorem 5.4], [20, Eqs. (1.1)–(1.3)], and [37, Theorem A.4] to the Schwindt's representation of a positive* mapping [16, Theorem 8.11-2] and the characterization of passive convolution operators in terms of positive* mappings [16, Theorem 8.12-1].

a. Susceptibility and passive linear media

In this section, we show how the electromagnetic assumptions of passive linear media are related to the theory of distributions and the theory of passive linear systems [16, Chap. 8] through the susceptibility χ, a generalized function, and its time derivative $\frac{d\chi}{dt}$ (in the distributional sense as defined in Ref. [16, p. 54] or as in Secs. B 2 and B 3 of this paper).

To begin with, the Hilbert spaces we are dealing with are either $\mathfrak{H} = \mathbb{C}^6$ with inner product $(\psi, \varphi) = \psi^\dagger \varphi$ or $\mathfrak{H} = [L^2(V)]^6$ for some measurable set $V \subseteq \mathbb{R}^d$, $d = 1, 2,$ or 3, with inner product $(\ldots, \ldots)$ given by

$$(\psi, \varphi) = \int_V \psi(\mathbf{r})^\dagger \varphi(\mathbf{r})d\mathbf{r}, \quad \psi, \varphi \in \mathfrak{H}. \tag{B7}$$

Also, for $\mathfrak{H} = \mathbb{C}^6$, we may replace in the following sections any limit in the strong operator topology (s-$\lim$) with the limit in the operator norm topology ($\lim$) since in any finite-dimensional Hilbert space the operator norm topology and strong operator topology (and weak operator topology) are equivalent [75, Chap. IV], [39, Sec. VI.1].

Now, the susceptibility χ is a generalized function (distribution), i.e., $\chi \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]$, and polarization response (3) is given in terms of the convolution operator $\chi* : \mathscr{D}(\mathfrak{H}) \to \mathscr{E}(\mathfrak{H})$ [as defined in Ref. [16, Lemma 5.5-1, Eq. (1)] by $(\chi * v)(t) = \langle \chi, v(t - \cdot) \rangle$. Thus, since $D(\chi * v) = (D\chi) * v$ for every $v \in \mathscr{D}(\mathfrak{H})$ (where $D = \frac{d}{dt}$ denotes the generalized derivative) which follows from [16, Lemma 5.5-1, Eq. (2)] and the definition (B3), it follows that the main assumption in our paper of passivity, either (5) or (23), is exactly the condition

$$0 \leqslant \operatorname{Re} \int_{-\infty}^{t} (v(\tau), (D\chi) * v(\tau)) d\tau \qquad (B8)$$

for every $v \in \mathscr{D}(\mathfrak{H})$ and every $t \in \mathbb{R}$. Hence, by definition [16, Def. 8.2-2], the condition (B8) means that $D\chi*$ is a passive convolution operator on $\mathscr{D}(\mathfrak{H})$.

b. Proofs of results in Secs. II B and III B

In this section, we prove the statements from Secs. II B and III B, namely, that linearity (3) and passivity (5), (23) implies causality, analyticity, and positivity as described in Secs. II B and III B. As we will show, these results just follow from the facts that $D\chi*$ is a passive convolution operator on $\mathscr{D}(\mathfrak{H})$, χ is an antiderivative of $D\chi$, and our existence and uniqueness results in Sec. B 2 on antiderivatives of distributions with support in $[0, \infty)$. Then, we conclude this section with two results that completely characterize $D\chi*$ as a passive convolution operator in terms of the Herglotz function $h(\omega) = \omega \widehat{\chi}(\omega)$ from (6).

Now, to begin, we have already shown in Sec. B 3 a that linearity (3) and passivity (5), (23) imply $D\chi*$ is a passive convolution operator on $\mathscr{D}(\mathfrak{H})$. This implies by [16, Theorem 8.2-1] that $D\chi*$ is a causal convolution operator on $\mathscr{D}(\mathfrak{H})$ [16, Def. 4.6-1] and hence [16, Theorem 5.11-1] implies that $(D\chi)(t) = 0$ for all $t < 0$, that is, $\operatorname{supp}(D\chi) \subseteq [0, \infty)$. From this property it follows from our results in Sec. B 2 that there exists a unique antiderivative of $D\chi$, say X, such that $\operatorname{supp} X \subseteq [0, \infty)$. It then follows from [16, Theorem 5.11-1] that $X*$ is a causal convolution operator on $\mathscr{D}(\mathfrak{H})$. As we discussed in Sec. B 2, since χ is also an antiderivative of $D\chi$, this implies $\chi - X$ must be a constant distribution. Our proofs following show that, in fact, we do not require this constant to be zero, but in order to simplify the exposition we will assume in the rest of this paper that $\chi = X$.

The next two results characterize $D\chi*$ as a passive convolution operator and these results follow immediately from [16, Theorem 8.12-1], [16, Lemma 7.2-1], the existence of antiderivatives of distributions as discussed in Sec. B 2, the relationship between the Laplace and Fourier transforms as discussed in Sec. B 1, and the relationship between positive* mappings [16, Def. 8.11-1] and Herglotz functions as defined in Sec. B 1 [namely, if $Y : C_+ \to [\mathfrak{H}; \mathfrak{H}]$ is a positive* mapping then $h : \mathbb{C}_+ \to [\mathfrak{H}; \mathfrak{H}]$ defined by $h(\omega) = iY(-i\omega)$ is a Herglotz function]. First, the following three statements are

equivalent: (i) $D\chi*$ is a passive convolution operator on $\mathscr{D}(\mathfrak{H})$; (ii) $D\chi$ is Laplace transformable with a strip of definition C_+ and its Laplace transform $Y = \mathfrak{L}(D\chi) : C_+ \to [\mathfrak{H}; \mathfrak{H}]$ is a positive* mapping; (iii) $D\chi$ has a Fourier transform $\widehat{D\chi}$ in $\mathbb{C}_+$ and $h = i\widehat{D\chi} : \mathbb{C}_+ \to [\mathfrak{H}; \mathfrak{H}]$ is a Herglotz function. In particular, the Laplace transform Y in (ii) and the Herglotz function h in (iii) are related by $h(\omega) = iY(-i\omega)$. Second, to every Herglotz function $h : \mathbb{C}_+ \to [\mathfrak{H}; \mathfrak{H}]$ there exists a unique distribution $D\chi$ [with $\chi \in [\mathscr{D}(\mathfrak{H}); \mathfrak{H}]]$ such that it has a Fourier transform $\widehat{D\chi}$ in $\mathbb{C}_+$ with $h = i\widehat{D\chi}$.

Now, since $D\chi$ has a Fourier transform $\widehat{D\chi}$ in $\mathbb{C}_+$ and the antiderivative χ has the property $\operatorname{supp} \chi \subseteq [0, \infty)$ then this implies by statement (ii) in Sec. B 2 and its proof that χ has a Fourier transform $\widehat{\chi}$ in $\mathbb{C}_+$ and the Herglotz function $h = i\widehat{D\chi} : \mathbb{C}_+ \to [\mathfrak{H}; \mathfrak{H}]$ satisfies

$$h(\omega) = i\widehat{D\chi}(\omega) = \omega\widehat{\chi}(\omega), \quad \omega \in \mathbb{C}_+. \qquad (B9)$$

This completes the proof of the results in Secs. II B and III B.

c. Proofs of results in Secs. II C, III C, and IV

In this section, we prove the statements from Secs. II C, III C, and IV that were left to be proved here. Our proofs of all these results are essentially based on the connection between the well-known integral representation of a Herglotz function from [19, Theorem 5.4], [20, Eqs. (1.1)–(1.3)], and [37, Theorem A.4] to the Schwindt's representation of a positive* mapping [16, Theorem 8.11-2] and the characterization of passive convolution operators in terms of positive* mappings [16, Theorem 8.12-1].

More specifically, we first use the integral representation (B11) and (B12) for the Herglotz function $h(\omega) = \omega\widehat{\chi}(\omega)$ in (B9) in terms of an operator-valued measure Ω that quantifies losses via the identity (B15). In particular, this identity proves the equality between (10) and (11). Next, we prove the formula (B14) which gives the correspondence between losses $\operatorname{Im} h(\omega)$ and the measure Ω and hence proves the identity (12). Finally, we prove the five main consequences [labeled as (i)–(v) in the subsection below entitled "On transparency windows"] of a transparency window defined in (14) as a frequency interval $(\omega_1, \omega_2) \subseteq \mathbb{R}$ in which $\Omega((\omega_1, \omega_2)) = 0$. In particular, using the integral representations (B11) and (B12) for $h(\omega)$ in (B9) we prove the main consequences of a transparency window (ω_1, ω_2), namely, analytic continuation of $h(\omega)$ across (ω_1, ω_2) to a bounded-operator-valued function satisfying the self-adjoint and monotonicity conditions (15). And, moreover, we also prove the fact mentioned in Sec. IV on achieving the upper bound $\|\mathbf{v}_e\| = c$ that the condition (31) being true for h at a single frequency in a transparency window implies that (31) and (32) are true for all frequencies.

Integral representations. As we proved in Sec. B 3 b, $h = i\widehat{D\chi} : \mathbb{C}_+ \to [\mathfrak{H}; \mathfrak{H}]$ is a Herglotz function. Hence, it follows from the well-known results on the integral representation of a Herglotz function (see, for instance, [17,19,20], [37, Appendix A], and references within) that there exists a unique tempered positive-operator-valued (PO) measure $\Omega : \mathfrak{B}_{\mathbb{R}, \infty} \to [\mathfrak{H}; \mathfrak{H}]_+$ (see [16, Def. 8.7-1], where $\mathfrak{B}_{\mathbb{R}, \infty}$ denotes the bounded Borel subsets of $\mathbb{R}$) such that

$$g(\lambda) = (1 + \lambda^2)^{-1} \qquad (B10)$$

is integrable with respect to the measure Ω and

$$h(\omega) = h_0 + h_1\omega + \int_{\mathbb{R}} d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right) \quad \text{(B11)}$$

for every $\omega \in \mathbb{C}_+$, where

$$h_0 = \operatorname{Re} h(i), \quad 0 \leqslant h_1 = s\text{-}\lim_{\eta\uparrow\infty} \frac{1}{i\eta} h(i\eta), \quad \text{(B12)}$$

where $\operatorname{Re} h(i)$ denotes the real part of the operator $h(i)$ and $s\text{-}\lim_{\eta\uparrow\infty}$ means the limit in the strong operator topology on $[\mathfrak{H};\mathfrak{H}]$ as real $\eta \to +\infty$. And, moreover, the Stieltjes inversion formula for the measure Ω holds in the strong operator topology, i.e.,

$$\frac{1}{2}\Omega((\omega_1,\omega_2)) + \frac{1}{2}\Omega([\omega_1,\omega_2]) = s\text{-}\lim_{\eta\downarrow 0} \int_{\omega_1}^{\omega_2} \frac{1}{\pi} \operatorname{Im} h(\lambda + i\eta) d\lambda$$
$$\text{(B13)}$$

for any $\omega_1, \omega_2 \in \mathbb{R}$ with $\omega_1 < \omega_2$, where the integral in Eq. (B13) is the Bochner integral (see [16, Appendix G] for definition) of the function $\lambda \mapsto \operatorname{Im} h(\lambda + i\eta)$ with respect to the Lebesgue measure on $\mathfrak{B}_{\mathbb{R}}$ (the Borel subsets of $\mathbb{R}$) and $s\text{-}\lim_{\eta\downarrow 0}$ means the limit in the strong operator topology on $[\mathfrak{H};\mathfrak{H}]$ as positive real $\eta \to 0$. Hence, from formula (B13) and [16, Lemma 2.2-1], it follows that

$$\Omega((\omega_1,\omega_2)) = s\text{-}\lim_{\delta\downarrow 0} s\text{-}\lim_{\eta\downarrow 0} \int_{\omega_1+\delta}^{\omega_2-\delta} \frac{1}{\pi} \operatorname{Im} h(\lambda + i\eta) d\lambda \quad \text{(B14)}$$

for any $\omega_1, \omega_2 \in \mathbb{R}$ with $\omega_1 < \omega_2$.

We will now prove that the following equality holds:

$$\operatorname{Re} \int_{-\infty}^{\infty} (v(t), D\chi * v(t)) dt = \frac{1}{2} \int_{\mathbb{R}} d(\Omega_\omega \widehat{v}(\omega), \widehat{v}(\omega))$$
$$\text{(B15)}$$

for every $v \in \mathscr{D}(\mathfrak{H})$, where the integral on the right in (B15) is defined as in [16, Secs. 2.6 and 8.7] for the Fourier transform defined by (B1). In fact, it follows from [16, Theorem 8.7-3] and the polarization identity for sesquilinear forms [16, Appendix A6], that the tempered PO measure Ω is uniquely determined by the positive quadratic form $v \mapsto \operatorname{Re} \int_{-\infty}^{\infty} (v(t), D\chi * v(t)) dt$ on $\mathscr{D}(\mathfrak{H})$.

To prove (B15), we begin by using [16, Theorem 2.3-3] to define a PO measure $Q : \mathfrak{B}_{\mathbb{R}} \to [\mathfrak{H};\mathfrak{H}]_+$ (where $\mathfrak{B}_{\mathbb{R}}$ denotes the Borel subsets of $\mathbb{R}$) by

$$Q(B) = \int_B d\Omega_\lambda g(\lambda), \quad B \in \mathfrak{B}_{\mathbb{R}}. \quad \text{(B16)}$$

Thus, since

$$\frac{1 + \lambda\omega}{\lambda - \omega} = g(\lambda)^{-1} \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right), \quad \text{(B17)}$$

it follows from [16, Theorem 2.3-4] that

$$h(\omega) = h_0 + h_1\omega + \int_{\mathbb{R}} dQ_\lambda \frac{1 + \lambda\omega}{\lambda - \omega} \quad \text{(B18)}$$

for every $\omega \in \mathbb{C}_+$.

Next, using the change of variable $\zeta = -i\omega$, it follows that the function $Y : C_+ \to [\mathfrak{H};\mathfrak{H}]$ defined by $Y(\zeta) = -ih(i\zeta)$ is a positive* mapping and is the Laplace transform $Y =$

$\mathfrak{L}(D\chi) : C_+ \to [\mathfrak{H};\mathfrak{H}]$ with the Schwindt's representation (cf. [16, Theorem 8.11-2])

$$Y(\zeta) = P_0 + P_1\zeta + \int_{\mathbb{R}} dP_\eta \frac{1 - i\eta\zeta}{\zeta - i\eta}, \quad \text{(B19)}$$

where P_0, P_1, and the PO measure $P : \mathfrak{B}_{\mathbb{R}} \to [\mathfrak{H};\mathfrak{H}]_+$ are related to h_0, h_1, and Q by the formulas

$$P_0 = -ih_0, \quad P_1 = h_1, \quad dP_\eta = dQ_{-\eta}, \quad \text{(B20)}$$

i.e., $P(B) = Q(-B)$ for every $B \in \mathfrak{B}_{\mathbb{R}}$, and, in particular, $\int_{\mathbb{R}} dP_\eta f(\eta) = \int_{\mathbb{R}} dQ_\eta f(-\eta)$ for every $f \in \mathscr{G}_{\mathbb{R}}$ (where $\mathscr{G}_{\mathbb{R}}$ denotes the set of all bounded Borel functions on $\mathbb{R}$). Finally, the proof of the equality (B15) now follows immediately from these facts and the proof of [16, Theorem 8.12-1].

On transparency windows. In the rest of this section, we will assume that $(\omega_1,\omega_2) \subseteq \mathbb{R}$, $\omega_1 < \omega_2$, is a transparency window for the medium with susceptibility χ, i.e.,

$$\Omega((\omega_1,\omega_2)) = 0. \quad \text{(B21)}$$

From the linear systems theory perspective, the condition (B21) interpreted in terms of the equality (B15) just means the linear system characterized by the passive convolution operator $D\chi*$ on $\mathscr{D}(\mathfrak{H})$ is lossless in the interval (ω_1,ω_2).

Now, we denote below the complement of the set (ω_1,ω_2) in $\mathbb{R}$ by E, i.e.,

$$E = (-\infty,\omega_1] \cup [\omega_2,\infty). \quad \text{(B22)}$$

Then, condition (B21) implies the Herglotz function $h = i\widehat{D\chi} : \mathbb{C}_+ \to [\mathfrak{H};\mathfrak{H}]$ has the following properties (i)–(v) which we first state as follows and then prove after.

(i) First, the domain of definition of h can be extended to $\mathbb{C} \setminus E$ by defining it in terms of the integral

$$h(\omega) = h_0 + h_1\omega + \int_E d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right) \quad \text{(B23)}$$

for every $\omega \in \mathbb{C} \setminus E$.

(ii) Next, the function $h : \mathbb{C} \setminus E \to [\mathfrak{H};\mathfrak{H}]$ in Eq. (B23) is an analytic $[\mathfrak{H};\mathfrak{H}]$-valued function in its domain. In particular, this function is just the analytical continuation of the Herglotz function $h : \mathbb{C}_+ \to [\mathfrak{H};\mathfrak{H}]$ into the open lower-half of the complex plane through the interval (ω_1,ω_2).

(iii) Next, differentiation under the integral in Eq. (B23) is allowed and the derivative of h has the integral representation

$$h'(\omega) = h_1 + \int_E d\Omega_\lambda \frac{1}{(\lambda - \omega)^2} \quad \text{(B24)}$$

for every $\omega \in \mathbb{C} \setminus E$.

(iv) Next, the function h is self-adjoint and monotonic in (ω_1,ω_2), i.e.,

$$\operatorname{Im} h(\omega) = 0, \quad h'(\omega) \geqslant 0, \quad \omega \in (\omega_1,\omega_2). \quad \text{(B25)}$$

(v) Finally, if $\psi \in \mathfrak{H}$ and $h'(\omega_0)\psi = 0$ for some $\omega_0 \in (\omega_1,\omega_2)$ then

$$h(\omega)\psi = h_0\psi, \quad \operatorname{Im} h_0 = 0 \quad \text{(B26)}$$

for every $\omega \in \mathbb{C} \setminus E$. Hence, extending this identity by continuity to E implies $h(\omega)\psi = h_0\psi$ for all $\omega \in \mathbb{C}$.

We will now prove these statements (i)–(v). We begin with the proof of statement (i). First, it follows immediately

from the assumption (B21), the identity (B17), and the elementary properties of σ-finite PO measures (namely, [16, Theorem 2.3-4], [16, Theorem 2.2-1], and the identity [16, p. 25, Eq. (4)]) that for every $\omega \in \mathbb{C}_+$, we have

$$\int_{\mathbb{R}} d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right) = \int_E d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right), \tag{B27}$$

where the integral on the right is with respect to the restriction of the measure Ω to $\mathfrak{B}_{E,\infty}$ (where $\mathfrak{B}_{E,\infty}$ denotes the bounded Borel subsets of E). Furthermore, by the identity (B17) and [16, Theorem 2.3-4] it follows that

$$\int_E d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right) = \int_E dQ_\lambda \frac{1 + \lambda\omega}{\lambda - \omega}, \tag{B28}$$

where $Q : \mathfrak{B}_E \to [\mathfrak{H}; \mathfrak{H}]_+$ is the PO measure defined above in (B16) restricted to $\mathfrak{B}_E$ (where $\mathfrak{B}_E$ denotes the Borel subsets of E). Moreover, it follows from the identity (B17) and [16, Theorem 2.3-4] that the integrals in (B28) are well defined and equal for all $\omega \in \mathbb{C} \setminus E$. From these facts, the proof of statement (i) now follows immediately.

We will now prove statements (ii) and (iii). Consider the function

$$f(\omega)(\lambda) = \frac{1 + \lambda\omega}{\lambda - \omega}, \quad \omega \in \mathbb{C} \setminus E, \quad \lambda \in E. \tag{B29}$$

Then, $f(\omega) \in \mathscr{G}_E$ for every $\omega \in \mathbb{C} \setminus E$ (where $\mathscr{G}_E$ denotes the set of all bounded Borel functions on E) and the $\mathscr{G}_E$-valued function $f : \mathbb{C} \setminus E \to \mathscr{G}_E$ is analytic on the open set $\mathbb{C} \setminus E$ into the Banach space $\mathscr{G}_E$ with the sup norm $\| \ldots \|_\infty$. Also, the integral map $\mathcal{I}_Q : \mathscr{G}_E \to [\mathfrak{H}; \mathfrak{H}]$ on $\mathscr{G}_E$ defined by

$$\mathcal{I}_Q(u) = \int_E dQ_\lambda u(\lambda), \quad u \in \mathscr{G}_E \tag{B30}$$

is a continuous linear operator which follows from [16, Theorem 2.2-3] and the inequality [16, p. 25, In Eq. (2)]. This implies the composition of the two maps $\mathcal{I}_Q \circ f : \mathbb{C} \setminus E \to$

$[\mathfrak{H}; \mathfrak{H}]$ is an analytic $[\mathfrak{H}; \mathfrak{H}]$-valued function which satisfies the equality $\frac{d}{d\omega}(\mathcal{I}_Q \circ f) = \mathcal{I}_Q \circ \frac{df}{d\omega}$ on $\mathbb{C} \setminus E$, where

$$\frac{df}{d\omega}(\omega)(\lambda) = \frac{g(\lambda)^{-1}}{(\lambda - \omega)^2}, \quad \omega \in \mathbb{C} \setminus E, \quad \lambda \in E. \tag{B31}$$

It follows now immediately from these facts, [16, Theorem 2.3-4], and the equality of the integrals in (B28) that

$$\frac{d}{d\omega} \int_E d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right) = \int_E d\Omega_\lambda \frac{1}{(\lambda - \omega)^2} \tag{B32}$$

for $\omega \in \mathbb{C} \setminus E$, where $\frac{d}{d\omega}$ here denotes the derivative in the operator norm topology of $[\mathfrak{H}; \mathfrak{H}]$. From these facts and statement (i), the proof of statements (ii) and (iii) now follows immediately.

We will now prove statements (iv) and (v). First, it follows from the identity [16, p. 25, Eq. (4)], [16, Theorem 2.2-1], and (B28) that

$$\begin{aligned}
&\left(\psi, \int_E d\Omega_\lambda \left(\frac{1}{\lambda - \omega} - \frac{\lambda}{1 + \lambda^2} \right) \psi \right) \\
&= \int_E d(Q_\lambda \psi, \psi) \frac{1 + \lambda\omega}{\lambda - \omega} \left(\psi, \int_E d\Omega_\lambda \frac{1}{(\lambda - \omega)^2} \psi \right) \\
&= \int_E d(Q_\lambda \psi, \psi) \frac{g(\lambda)^{-1}}{(\lambda - \omega)^2}
\end{aligned} \tag{B33}$$

for any $\psi \in \mathfrak{H}$ and any $\omega \in \mathbb{C} \setminus E$, where the integrals in (B33) on the right are Lebesgue integrals with respect to the positive finite measure $(Q(\ldots)\psi, \psi) : \mathfrak{B}_E \to [0, \infty)$. Now, the statements (iv) and (v) immediately follow from the representation (B33), the fact that for all $\omega \in (\omega_1, \omega_2)$ and $\lambda \in E$ the integrands in Eq. (B33) are real with $\frac{g(\lambda)^{-1}}{(\lambda - \omega)^2} > 0$, the fact that $h_0 = \text{Re}\, h(i)$ is self-adjoint, $h_1 \geqslant 0$, and from statements (i) and (iii). This completes the proof of statements (i)–(v). These statements together with the equality (B9) prove our results in Secs. II C, III C, and IV.

---

[1] L. Brillouin, in *Congrès International d'Electricité, Paris, 1932* (Gauthier-Villars, Paris, 1933), Vol. 2, pp. 739–788.

[2] L. Brillouin, *Wave Propagation and Group Velocity*, Pure and Applied Physics, Vol. 8 (Academic, New York, 1960).

[3] R. Loudon, *The Propagation of Electromagnetic Energy Through an Absorbing Dielectric*, Memorandum 2155 (Royal Radar Establishment, Malvern, Worcestershire, England, 1965).

[4] R. Loudon, J. Phys. A: Gen. Phys. **3**, 233 (1970).

[5] E. O. Schulz-DuBois, Proc. IEEE **57**, 1748 (1969).

[6] K. E. Oughstun and S. Shen, J. Opt. Soc. Am. B **5**, 2395 (1988).

[7] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media*, 2nd ed. (Pergamon, New York, 1984).

[8] S. Glasgow, M. Ware, and J. Peatross, Phys. Rev. E **64**, 046610 (2001).

[9] A. D. Yaghjian, IEEE Trans. Antennas Propag. **55**, 1495 (2007).

[10] A. D. Yaghjian, IEEE Trans. Antennas Propag. **55**, 3748 (2007).

[11] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, Princeton, NJ, 2008).

[12] D. Geppert, Proc. IEEE **53**, 1790 (1965).

[13] A. H. Zemanian, *Distribution Theory and Transform Analysis* (Dover, New York, 1965).

[14] A. H. Zemanian, IEEE Trans. Circuit Theory **CT–10**, 265 (1963).

[15] A. H. Zemanian, SIAM J. Appl. Math. **18**, 98 (1970).

[16] A. H. Zemanian, *Realizability Theory for Continuous Linear Systems* (Academic, New York, 1972).

[17] N. Aronszajn and W. F. Donoghue, J. Anal. Math. **5**, 321 (1956/57).

[18] S. D. Greenstein, J. Math. Anal. Appl. **1**, 355 (1960).

[19] F. Gesztesy and E. Tsekanovskii, Math. Nachr. **218**, 61 (2000).

[20] F. Gesztesy, N. J. Kalton, K. A. Makarov, and E. Tsekanovskii, in *Operator Theory, System Theory and Related Topics (Beer-Sheva/Rehovot, 1997)*, Operator Theory Advanced Applications, Vol. 123 (Birkhäuser, Basel, 2001), pp. 271–321.

[21] T. Havelock, *The Propagation of Disturbances in Dispersive Media*, Cambridge Tracts in Mathematics and Mathematical Physics (Cambridge University Press, Cambridge, UK, 1914).

[22] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).

[23] J. Stratton, *Electromagnetic Theory*, IEEE Press classic reissue (Wiley, Hoboken, NJ, 2007).

[24] M. Abraham, Nuovo Cimento **I**, 443 (1911).

[25] A. D. Yaghjian, S. Maci, and E. Martini, New J. Phys. **11**, 113011 (2009).

[26] P. Yeh, J. Opt. Soc. Am. **69**, 742 (1979).

[27] M. A. Biot, Phys. Rev. **105**, 1129 (1957).

[28] E. H. Wagner, Acta Crystallogr. **12**, 345 (1959).

[29] E. Wagner, Z. Phys. **154**, 352 (1959).

[30] H. Bertoni and A. Hessel, IEEE Trans. Antennas Propag. **14**, 344 (1966).

[31] P. Y. Chen, R. C. McPhedran, C. M. de Sterke, C. G. Poulton, A. A. Asatryan, L. C. Botten, and M. J. Steel, Phys. Rev. A **82**, 053825 (2010).

[32] R. Chu and T. Tamir, Electron. Lett. **7**, 410 (1971).

[33] A. Yaghjian and S. Best, IEEE Trans. Antennas Propag. **53**, 1298 (2005).

[34] A. Figotin and J. Schenker, J. Comput. Appl. Math. **204**, 199 (2007).

[35] A. Figotin and J. Schenker, J. Stat. Phys. **128**, 969 (2007).

[36] L. Brillouin, *Wave Propagation in Periodic Structures: Electric Filters and Crystal Lattices*, International Series in Pure and Applied Physics (McGraw-Hill, New York, 1946).

[37] F. Gesztesy, R. Weikard, and M. Zinchenko, Oper. Matrices **7**, 241 (2013).

[38] J. Yeh, *Real Analysis: Theory of Measure and Integration*, 2nd ed. (World Scientific, Hackensack, NJ, 2006).

[39] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, London, 1980).

[40] T. M. Apostol, *Mathematical Analysis*, 2nd ed. (Addison-Wesley, Reading, MA, 1974).

[41] P. Kuchment, *The Mathematics of Photonic Crystals* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001), pp. 207–272, Chap. 7.

[42] J. M. McMahon, S. K. Gray, and G. C. Schatz, Phys. Rev. Lett. **103**, 097403 (2009).

[43] F. J. García de Abajo, Rev. Mod. Phys. **82**, 209 (2010).

[44] C. David and F. J. Garca de Abajo, J. Phys. Chem. C **115**, 19470 (2011).

[45] R. Ruppin, Phys. Rev. B **11**, 2871 (1975).

[46] X. Ben and H. S. Park, J. Phys. Chem. C **116**, 18944 (2012).

[47] V. Yannopapas, J. Phys.: Condens. Matter **20**, 325211 (2008).

[48] C. Tserkezis, G. Gantzounis, and N. Stefanou, J. Phys.: Condens. Matter **20**, 075232 (2008).

[49] J. Li, G. Sun, and C. T. Chan, Phys. Rev. B **73**, 075117 (2006).

[50] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Pergamon, New York, 1996), Vol. 5.

[51] I. V. Lindell and A. H. Sihvola, J. Electromagn. Waves Appl. **19**, 861 (2005).

[52] A. Sihvola, Metamaterials **1**, 2 (2007).

[53] A. Sihvola and I. V. Lindell, Ann. Phys. (Berlin) **17**, 787 (2008).

[54] H. A. Lorentz, Versl. K. Akad. W. (Amsterdam) **4**, 176 (1896).

[55] J. A. Kong, Proc. IEEE **58**, 1966 (1970).

[56] R. J. Potton, Rep. Prog. Phys. **67**, 717 (2004).

[57] W. C. Chew, IEEE Trans. Antennas Propag. **56**, 970 (2008).

[58] A. Sommerfeld, Phys. Z. **8**, 841 (1907).

[59] A. Sommerfeld, Ann. Phys. (Berlin) **44**, 177 (1914).

[60] L. Brillouin, Ann. Phys. (Berlin) **44**, 203 (1914).

[61] L. C. Evans, *Partial Differential Equations*, 2nd ed., Graduate Studies in Mathematics, Vol. 19 (American Mathematical Society, Providence, RI, 2010).

[62] R. L. Smith, Am. J. Phys. **38**, 978 (1970).

[63] S. C. Bloch, Am. J. Phys. **45**, 538 (1977).

[64] P. W. Milonni, *Fast Light, Slow Light and Left-Handed Light*, Series in Optics and Optoelectronics (IOP, Bristol, 2005).

[65] R. Y. Chiao and A. M. Steinberg, in *Tunneling Times and Superluminality* (Elsevier, Amsterdam, 1997), pp. 347–406, Chap. VI.

[66] J. Rauch, *Hyperbolic Partial Differential Equations and Geometric Optics*, Graduate Studies in Mathematics (American Mathematical Society, Providence, RI, 2012).

[67] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[68] H. Hellmann, *Einführung in die Quantenchemie (Introduction to Quantum Chemistry, in German)* (Franz Deuticke, Leipzig Germany, 1937).

[69] L. Landau and E. Lifshits, *Quantum Mechanics: Non-relativistic Theory*, 3rd ed. (Pergamon, New York, 1991).

[70] M. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*, No. 13 (Cambridge University Press, Cambridge, UK, 2005).

[71] S. T. Epstein, Am. J. Phys. **22**, 613 (1954).

[72] J. I. Musher, Am. J. Phys. **34**, 267 (1966).

[73] R. D. Levine, Proc. R. Soc. London, Ser. A **294**, 467 (1966).

[74] V. I. Pupyshev, Russian J. Phys. Chem. **74**, S267 (2000).

[75] P. R. Halmos, *Finite-Dimensional Vector Spaces*, 2nd ed. (Springer, New York, 1974).