Research Article Vol. 35, No. 3 / March 2018 / Journal of the Optical Society of America B 533

Journal of the Optical Society of America B

OPTICAL PHYSICS

Two-stream theory of light propagation in amplifying media

VADIM A. MARKEL

Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA (vmarkel@pennmedicine.upenn.edu)

Received 6 November 2017; revised 7 January 2018; accepted 8 January 2018; posted 9 January 2018 (Doc. ID 312781); published 8 February 2018

The two-stream approximation to the radiative transport equation (RTE) is a convenient exactly solvable model that allows one to analyze propagation of light in amplifying media. In spite of neglecting the phase and the interference effects, this model describes the same phenomena as Maxwell's equation: electromagnetic resonances, onset of lasing, and onset of instabilities. An important added bonus of the RTE description is that it provides for a simple and unambiguous test of physicality of stationary solutions. In the case of Maxwell's equations, it is not always obvious or easy to determine whether certain stationary (in particular, monochromatic) solutions are physical. In the case of RTE, the specific intensity of unphysical stationary solutions becomes negative for some subset of its arguments. In the paper, stationary and time-dependent solutions to the two-stream model are analyzed. It is shown that the conditions for stationary lasing and for emergence of instabilities depend only on the geometry of the sample and the strength of amplification but not on the intensity of incident light.

OCIS codes: (110.6960) Tomography; (290.5855) Scattering, polarization; (110.5405) Polarimetric imaging.

https://doi.org/10.1364/JOSAB.35.000533

1. INTRODUCTION

Recently, propagation of light in amplifying media has attracted considerable interest [1–10], and even some controversy [11–14]. The problem has, in fact, a long history [15,16], and a comprehensive exposition of the subject has been given in [17]. The difficulty seems to be rooted in the attempts to describe amplifying media phenomenologically by a linear permittivity ϵ with a negative imaginary part. This description is sometimes valid and sometimes not, but it is easy to make a mistake by assuming the existence of certain stationary (in particular, monochromatic) solutions to Maxwell's equations that are not physically realizable. That is, these solutions—even though they appear to be perfectly valid—can never be reached if one starts from a physically reasonable initial condition. In addition, amplifying media are known to possess instabilities. This means that solutions can depend dramatically on small variations in the initial conditions or the form of external radiation.

In the case of Maxwell's equations, determining which stationary solutions are physical and which are not is not always straightforward. However, if we describe the light propagation in the medium by the radiative transport equation (RTE), the determination becomes much simpler. The reason for this simplification is that the specific intensity I, which is the physical quantity described by the RTE [18], is point-wise nonnegative by definition. We will see that the stationary solutions purportedly yielding I can be, in fact, negative, at least for some subsets of the arguments of I, if these solutions are unphysical in the above sense. This is a red flag; such solutions are mathematical artifacts of an imprecise or incomplete model; they do not correspond to the physical reality and should not be used. A related point is that, for the corresponding parameters of the medium, stationary solutions do not exist, and one must consider time dependence.

A similar consideration cannot be applied to the electric field, which is a vector and can have a projection of arbitrary sign on any given axis. However, we will show that many important features of light propagation in amplifying slabs as described by Maxwell's equations [17] are also present in the transport theory. This is counterintuitive, because the RTE disregards the phase of light and, correspondingly, it disregards the interference effects. Nevertheless, the transport theory predicts that a sample of amplifying medium has resonances and can support stable lasing or exponentially growing runaway solutions, just as is the case for Maxwell's equations.

Thus, the main goal of this paper is to provide an alternative and a somewhat simpler (albeit an approximate) theoretical framework in which propagation through amplifying media can be considered. It will be shown that, although the mathematical model employed below does not involve the phase, it captures, at least qualitatively, certain physical phenomena such as resonances and the onset of lasing that are generally believed

to be closely related to phase and interference. The advantage of using the formalism of this paper is simplicity and transparency. It will become clear, for example, that, in the recent controversy between Baranov *et al.* and Wang *et al.* [13,14], the former are right: the onset of lasing is a condition that involves the size of the sample and the magnitude of the imaginary part of the permittivity but not the field strength.

To avoid unnecessary complications, we will consider a simple exactly solvable model that involves two oppositely directed *streams* of radiation. The corresponding *two-stream approximation* was originally introduced by Kubelka and Munk in 1931 [19], well before the modern radiative transport theory was developed. The goal of the original paper by Kubelka and Munk was to compute the reflectance of a layer of paint. However, the model with just two streams, although never precise, is surprisingly rich and describes radiation transfer in many physical systems qualitatively. For example, the two-stream approximation can be used to explain the varying colors of clouds. Also, the two-stream theories are applicable to all one-dimensional systems in which forward and backward propagation is possible, such as waveguides and transmission lines. More broadly, the theory of this paper can be applicable to random lasing in micro- or nanopowders wherein the description in terms of Maxwell's equations is too complex and detailed to be practical and the approximate description based on the RTE must be used instead.

The point of departure for this paper is the RTE written in the form

$$\left(\frac{1}{c}\frac{\partial}{\partial t}+\hat{\mathbf{s}}\cdot\nabla+\mu_t\right)I(\mathbf{r},\hat{\mathbf{s}},t)=\mu_s\int A(\hat{\mathbf{s}},\hat{\mathbf{s}}')I(\mathbf{r},\hat{\mathbf{s}}',t)\mathrm{d}^2s'. \tag{1}$$

Here $I(\mathbf{r},\hat{\mathbf{s}},t)$ is the specific intensity at the position $\mathbf{r}$ in the direction of the unit vector $\hat{\mathbf{s}}$ and at the time t, c is the average speed of light in the medium, $\mu_t=\mu_a+\mu_s$ is the total attenuation coefficient of the medium, with μ_a and μ_s being the absorption and the scattering coefficients, and, finally, $A(\hat{\mathbf{s}},\hat{\mathbf{s}}')$ is the single-scattering phase function. We assume that external radiation enters the medium through its boundary, which can be described mathematically by inhomogeneous boundary conditions (stated below).

The two-stream approximation grows from the minimalistic assumption that only forward and backward scattering is possible. In this case, single scattering in the medium is governed by the delta-Eddington phase function:

$$A(\mathbf{s},\hat{\mathbf{s}}')=p\delta_2(\mathbf{s},\hat{\mathbf{s}}')+q\delta_2(\mathbf{s},-\hat{\mathbf{s}}'), \tag{2}$$

where p and q are the probabilities of forward and backward scattering ($p+q=1$). The scattering asymmetry parameter of the medium is $g=p-q=1-2q=2p-1$. The transport mean free path ℓ^* is given for this medium by

$$\ell^*=\frac{1}{\mu_a+(1-g)\mu_s}=\frac{1}{\mu_a+2q\mu_s}. \tag{3}$$

Note that $\delta_2(\mathbf{s},\hat{\mathbf{s}}')$ in Eq. (2) is the two-dimensional angular delta function with the property $\int\delta_2(\mathbf{s},\hat{\mathbf{s}}')f(\hat{\mathbf{s}}')\mathrm{d}^2s'=f(\hat{\mathbf{s}})$ [20]. Upon substitution of Eq. (2) into Eq. (1), the RTE becomes

$$\left(\frac{1}{c}\frac{\partial}{\partial t}+\hat{\mathbf{s}}\cdot\nabla+\mu_a+q\mu_s\right)I(\mathbf{r},\hat{\mathbf{s}},t)=q\mu_sI(\mathbf{r},-\hat{\mathbf{s}},t). \tag{4}$$

In what follows, we will consider the case when the vector $\hat{\mathbf{s}}$ is perpendicular to an infinite slab of a material. In this case, propagation is described by a one-dimensional set of equations.

2. STATIONARY TWO-STREAM EQUATIONS

Consider the case when a sufficiently wide front of parallel rays of stationary intensity I_0 (incoming energy per unit surface per unit time) is normally incident onto a layer contained between the planes $z=0$ and $z=L$. The slab is characterized by spatially uniform coefficients μ_s and μ_a and by the scattering probabilities p and q. Under certain conditions, which are explored in detail below (and certainly in the case $\mu_a>0$), the specific intensity is also stationary and can be written in the form

$$I(\mathbf{r},\hat{\mathbf{s}},t)=I_0[i_1(z)\delta_2(\hat{\mathbf{s}},\hat{\mathbf{z}})+i_2(z)\delta_2(\hat{\mathbf{s}},-\hat{\mathbf{z}})], \tag{5}$$

where the two dimensionless *streams*, $i_1(z)$ and $i_2(z)$, satisfy the pair of ordinary differential equations

$$(+d/\mathrm{d}z+\mu_a+q\mu_s)i_1(z)=q\mu_si_2(z), \tag{6a}$$

$$(-d/\mathrm{d}z+\mu_a+q\mu_s)i_2(z)=q\mu_si_1(z), \tag{6b}$$

and the boundary condition

$$i_1(0)=1,\qquad i_2(L)=0. \tag{7}$$

Note that the above boundary condition does not account for Fresnel reflections of the streams at the slab interfaces (i.e., due to an index mismatch). A more general case of nonnegligible Fresnel reflections is considered in Section 7 below. We thus see that the parameter $1/q\mu_s$ sets the characteristic length scale of the problem. All possible solutions can be expressed in terms of three dimensionless variables: $q\mu_sz$, $q\mu_sL$, and $\mu_a/q\mu_s$. Without loss of generality, we can assume that $q\mu_s$ is fixed, while μ_a and L and z can vary. This point of view is adopted everywhere below except in Section 7, where, in one of the figures, we assume that μ_a is fixed while $q\mu_s$ and L can vary.

Note that the stationary two-stream equations (6) are equivalent to a one-dimensional stationary diffusion equation for the energy density $u(z)$ (this equivalence does not hold in the time-dependent case; see below). Indeed, let us define the density and current of energy as $u(z)=(I_0/c)[i_1(z)+i_2(z)]$ and $J(z)=I_0[i_1(z)-i_2(z)]$. The two functions satisfy

$$dJ/dz+\alpha u(z)=0,\qquad Ddu/dz+J=0, \tag{8}$$

where $D=c\ell^*$ and $\alpha=c\mu_a$ are the diffusion coefficient and the rate of absorption. Thus, Fick's law (second equation above) holds in the two-stream model exactly. We can further transform Eq. (8) into a second-order diffusion equation for u, viz.,

$$(-Dd^2/dz^2+\alpha)u=0. \tag{9}$$

The boundary condition for the diffusion equation follows from Eq. (7) and is of the form

$$u(0)-\ell^*u'(0)=2I_0,\qquad u(L)+\ell^*u'(L)=0. \tag{10}$$

This is a special case of the more general boundary condition $(u+\ell\hat{\mathbf{n}}\cdot\nabla u)|_{\mathbf{r}\in\partial\Omega}$, where $\hat{\mathbf{n}}$ is the outward unit normal to the boundary $\partial\Omega$ of the domain Ω occupied by the medium, and the parameter ℓ is known as the extrapolation distance [21]. Generally, the value of ℓ depends on the medium parameters

and, in three dimensions, the typical value of ℓ is $\ell \sim 0.71\ell^*$ [22] (in the absence of Fresnel reflections at the interfaces). The above result $\ell = \ell^*$, as well as the expression $D = c\ell^*$ for the diffusion coefficient, are exact only in the one-dimensional two-stream model.

Note that it is possible to establish an exact equivalence between the one-dimensional diffusion equation (9) with the boundary condition (10) and the two-stream equations (6) with the boundary condition (7). However, there is no such correspondence between the one-dimensional diffusion equation and a more general one-dimensional RTE. If the phase function is not of the delta-Eddington form (2), then the diffusion equation (9) is only an approximation to the RTE and, moreover, one should use in this case different values for the parameters D and ℓ.

For our purposes, it is more convenient to work directly with the streams i_1 and i_2, since both functions are physically required to be nonnegative. The solution to Eqs. (7), (6) is

$$i_1(z) = \frac{a^- e^{\lambda(z-L)} - a^+ e^{\lambda(L-z)}}{a^- e^{-\lambda L} - a^+ e^{\lambda L}}, \tag{11a}$$

$$i_2(z) = q\mu_s \frac{e^{\lambda(z-L)} - e^{\lambda(L-z)}}{a^- e^{-\lambda L} - a^+ e^{\lambda L}}, \tag{11b}$$

where

$$a^{\pm} = \mu_a + q\mu_s \pm \lambda, \tag{12a}$$

$$\lambda = \sqrt{\mu_a(\mu_a + 2q\mu_s)} = \sqrt{\mu_a/\ell^*}. \tag{12b}$$

As one could expect, the expressions for $i_1(z)$, $i_2(z)$ are invariant with respect to the substitution $\lambda \to -\lambda$. Although nothing depends on the choice of the square root branch in Eq. (12b), we will, for the sake of clarity, fix the branch by applying the condition $0 \le \arg(\lambda) < \pi$. Then we can distinguish the following cases:

(i) If $\mu_a > 0$, then $\lambda > 0$.
(ii) If $\mu_a = 0$, then $\lambda = 0$.
(iii) If $-2q\mu_s < \mu_a < 0$, then $\lambda = i|\lambda| = i\sqrt{|\mu_a|(2q\mu_s - |\mu_a|)}$. In this case, $\max_{\mu_a}|\lambda| = q\mu_s$ is achieved at $\mu_a = -q\mu_s$.
(iv) If $\mu_a = -2q\mu_s$, then $\lambda = 0$.
(v) If $\mu_a < -2q\mu_s$, then $\lambda = \sqrt{|\mu_a|(|\mu_a| - 2q\mu_s)} > 0$.

To summarize, λ is either purely real and positive, or purely imaginary with a positive imaginary part, or zero.

3. STATIONARY TRANSMISSION AND REFLECTION BY A SLAB

We can define the transmission and reflection coefficients of the slab as $T = i_1(L)$ and $R = i_2(0)$. The reflection coefficient R is called the *albedo*; it would yield, for example, the reflectivity of a plane-parallel atmosphere. From the result (11), we can find that

$$T = \frac{\lambda}{(\mu_a + q\mu_s)\sinh(\lambda L) + \lambda\cosh(\lambda L)}, \tag{13a}$$

$$R = \frac{q\mu_s \sinh(\lambda L)}{(\mu_a + q\mu_s)\sinh(\lambda L) + \lambda\cosh(\lambda L)}. \tag{13b}$$

We start the consideration of amplifying media with a few numerical examples. In Fig. 1(a), we plot T and R as functions of the slab width L for a fixed negative value of μ_a, which corresponds to amplification (negative absorptive losses). In Fig. 1(b), we plot T and R as functions of μ_a spanning negative and positive values for a fixed value of L.

In the case of fixed μ_a, the functions $T(L)$, $R(L)$ have resonances (divergences) at the values of L that satisfy the equation

$$e^{2\lambda L} = \frac{\mu_a + q\mu_s - \lambda}{\mu_a + q\mu_s + \lambda}. \tag{14}$$

When $\mu_a > 0$, the above equation does not have positive roots L. Indeed, for $\mu_a, L, \lambda > 0$, the left-hand side of Eq. (14) is greater than unity, while the right-hand side is smaller, and the equality is impossible.

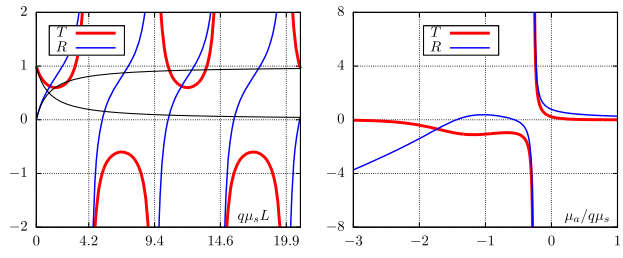
However, if μ_a lies in the interval $-2q\mu_s < \mu_a < 0$, then $\lambda = i|\lambda|$ is imaginary, and Eq. (14) has infinitely many positive roots L_n, which are all of the form

$$L_n(\mu_a) = \frac{1}{|\lambda(\mu_a)|}\left[n\pi - \arctan\left(\frac{|\lambda(\mu_a)|}{\mu_a + q\mu_s}\right)\right], \quad \text{if } -2q\mu_s < \mu_a < 0. \tag{15}$$

In the above expression, λ is viewed as a function of μ_a as is given by Eq. (12b), and n is an integer index labeling the roots. We should select only such indices n that make the expression in the right-hand side of Eq. (15) positive. Let the arctangent be defined by the condition $0 \le \arctan(z) < \pi$ for any real number z. Although $\arctan(z)$ is discontinuous in this case, all functions $L_n(\mu_a)$ defined by Eq. (15) turn out to be continuous in the interval $-2q\mu_s < \mu_a < 0$. Then the consideration of indices is simple: $n = 1, 2, 3, \ldots$ yield all positive roots L_n. In Fig. 1(a), the ticks on the horizontal axis and the dashed vertical lines are shown at the positions of the first four resonance values L_n.

In the case $\mu_a < -2q\mu_s$, Eq. (14) has exactly one positive root:

$$L_1(\mu_a) = \frac{1}{2\lambda(\mu_a)} \ln\frac{|\mu_a + q\mu_s| + \lambda(\mu_a)}{|\mu_a + q\mu_s| - \lambda(\mu_a)}, \quad \text{if } \mu_a < -2q\mu_s. \tag{16}$$



(a) $\mu_a = -0.2q\mu_s$ (b) $L = 3.5/q\mu_s$

Fig. 1. Two-stream approximation predictions for the transmission and reflection coefficients of a slab. Panel (a) shows T and R as functions of L for a fixed negative value of μ_a. Panel (b) shows these coefficients as functions of μ_a spanning negative and positive values for a fixed L. For comparison purposes, the thin lines in panel (a) show the respective results computed at $\mu_a = 0$. The ticks on the horizontal axis and the dashed vertical lines of panel (a) correspond to the first four resonance values L_n as defined by Eq. (15) (the numerical values shown are approximate).

Note that the expression under the logarithm and the function $\lambda(\mu_a)$ are positive under the conditions of applicability of Eq. (16).

The function $L_1(\mu_a)$ can be defined using Eq. (15) and Eq. (16) and is positive and continuous in the whole open interval $\mu_a < 0$, while the functions $L_n(\mu_a)$ with $n > 1$ are defined only in the interval $-2q\mu_s < \mu_a < 0$ [23]. The first several functions $L_n(\mu_a)$ are plotted in Fig. 2.

For a fixed value of L, we can similarly find the resonance values of μ_a as the roots $\mu_{an}(L)$ of Eq. (14). The two obvious roots $\mu_a = 0$ and $\mu_a = -2q\mu_s$ should be considered separately. These values of μ_a satisfy Eq. (14) irrespectively of L. However, the root $\mu_a = 0$ does not correspond to a resonance and should always be excluded from consideration. In fact, the transmission and reflection coefficients at $\mu_a = 0$ are always finite and positive. The root $\mu_a = -2q\mu_s$ corresponds to a resonance only if, in addition, $L = 1/q\mu_s$. For a general L, the transmission and reflection coefficients at $\mu_a = -2q\mu_s$ are given by $T = 1/(1 - q\mu_s L)$ and $R = q\mu_s/(1 - q\mu_s L)$. Obviously, these results are physical and correct if $q\mu_s L < 1$ and unphysical and incorrect otherwise; the divergence occurs at $L = 1/q\mu_s$.

The number of *real-valued* roots $\mu_{an}(L)$ is finite and almost always odd, except for a discrete set of values of L. At least one resonance value of the absorption coefficient exists for any $L > 0$. For example, only one such resonance exists for the parameters of Fig. 1(b); the divergence of T and R at this single resonance value of μ_a is clearly visible in the figure.

The resonance values of μ_a cannot be expressed in terms of elementary functions. However, it is easy to find these roots numerically or visualize graphically. Indeed, let us plot all the functions $L_n(\mu_a)$ defined in Eqs. (15) and (16), as is done in Fig. 2. The roots $\mu_{an}(L)$ can be visualized as the intersections of a horizontal line $L = \text{const}$ with all the curves shown in the figure. It can indeed be seen that the number of roots is odd except when the horizontal line touches the minimum of one of these curves [24].

Let us now fix some $\mu_a < 0$, and let $L_1(\mu_a)$ be the first (the smallest) resonance value of L for this particular μ_a. Then the coefficients T and R are positive for any L in the interval $0 < L < L_1(\mu_a)$. This is illustrated in Fig. 1(a). In fact, not

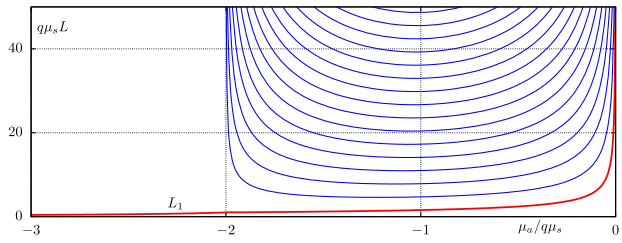


Fig. 2. Loci of the resonances of the two-stream model in the real (μ_a, L) plane. A point on any of the curves corresponds to a root of Eq. (14). The lowest resonance curve $L_1(\mu_a)$ is emphasized by color and linewidth. All points to the right and below the curve $L_1(\mu_a)$ correspond to physically meaningful unique solutions to the stationary two-stream equations (6). Any point above and to the left of $L_1(\mu_a)$, but not on any of the other curves, $L_n(\mu_a)$ corresponds to a unique but unphysical solution. Finally, for any point on any of the curves $L_n(\mu_a)$, solutions, even unphysical ones, do not exist for the inhomogeneous boundary condition (7).

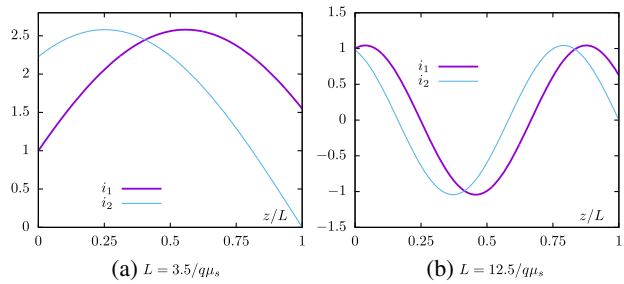


Fig. 3. Two streams $i_1(z)$ and $i_2(z)$ for $\mu_a = -0.2q\mu_s$ and $L = 3.5/q\mu_s$ (a), $L = 12.5/q\mu_s$ (b). The streams are shown as functions of the dimensionless variable z/L. At $L = 12.5/q\mu_s$, both the transmission and reflection coefficients are positive (see Fig. 1), but both functions $i_1(z)$ and $i_2(z)$ change sign; such solutions are unphysical.

only the transmission and reflection coefficients, but the specific intensity is also positive for these values of L. This is illustrated in Fig. 3(a), where the functions $i_1(z)$ and $i_2(z)$ are plotted for $\mu_a = -0.2q\mu_s$ [same value as in Fig. 1(a)] and $q\mu_s L = 3.5$. This value of L satisfies the inequality $0 < L < L_1(\mu_a)$. Indeed, we have $q\mu_s L_1(-0.2q\mu_s) \approx 4.15$.

We can generalize that, for any $\mu_a < 0$, the two-stream approximation yields physically meaningful solutions as long as $L < L_1(\mu_a)$, that is, in sufficiently thin slabs, up to the first resonance value of L. Similarly, for any fixed value of L, the solutions are physical for $\mu_a > \mu_{a1}(L)$, where $\mu_{a1}(L)$ is the first (negative) resonance value of μ_a for the given L. The two-dimensional region in the plane (μ_a, L) where physically meaningful solutions exist is illustrated in Fig. 2.

It may seem that, in some cases, physically meaningful stationary solutions can exist even beyond the first resonance, that is, above and to the left of the curve $L_1(\mu_a)$ in Fig. 2. For example, both coefficients T and R are positive for $\mu_a = -0.2q\mu_s$ and $q\mu_s L = 12.5$, as can be seen in Fig. 1(a). This point (μ_a, L) is above the curve $L_1(\mu_a)$. However, the individual streams inside the medium change sign and are unphysical at these values of parameters; this is illustrated in Fig. 3(b). Generally, the stationary solutions are always unphysical for $L > L_1(\mu_a)$ or $\mu_a < \mu_{a1}(L)$. Exactly at the resonances, stationary solutions simply do not exist. Past the first resonance and between higher-order resonances, solutions exist, but the specific intensity becomes negative at least for some positions and directions.

4. STATIONARY LASING

We have defined the resonances of the two-stream model as the set of parameters μ_a and L, for which the stationary two-stream equations (6) with the inhomogeneous boundary condition (7) has no solutions. This set is defined by all possible solutions to Eq. (14). All *real* resonance values of the above two parameters lie on the curves shown in Fig. 2.

We can as well say that the two-stream equations (6) with the *homogeneous* boundary condition $i_1(0) = i_2(L) = 0$ have nontrivial solutions if and only if the parameters μ_a and L take the resonance values. Let us look in more detail at these nontrivial stationary solutions, which are obtained in the absence of any external sources. Equations (6) with the above homogeneous boundary condition are satisfied by the functions

$$i_1(z) = A(e^{\lambda z} - e^{-\lambda z}), \tag{17a}$$

$$i_2(z) = \frac{A}{q\mu_s}[(\mu_a + q\mu_s + \lambda)e^{\lambda z} - (\mu_a + q\mu_s - \lambda)e^{-\lambda z}], \tag{17b}$$

provided that Eq. (14) holds. Otherwise, the problem has only the trivial solution. Here A is an arbitrary constant, and we should exclude the root $\mu_a = 0$ (which satisfies Eq. (14) for an arbitrary L) from consideration. As was explained in Section 3, $\mu_a = 0$ is not a resonance of the system.

As for the resonance at the point $\mu_a = -2q\mu_s$, $L = 1/q\mu_s$, it should be considered separately. The right-hand sides in the equations (17) turn to zero in this case. However, there exists a nontrivial linear solution at this resonance point; it is of the form $i_1(z) = Az$, $i_2(z) = A(1 - z/L)$. This solution can be, in fact, obtained from Eq. (17) if we make the substitution $A \to A/2\lambda$.

An important consideration is symmetry. Since, in the absence of any incoming external radiation, the two possible directions of propagation are physically equivalent, we expect the expressions for the two streams to satisfy $i_{1,2}(z) = \pm i_{2,1}(L - z)$. The possibility of the minus sign in the above symmetry relation is somewhat counterintuitive. However, Eqs. (17) allow for both possibilities. The physical requirement that each stream transforms into itself after we apply the operation of coordinate inversion twice holds in both cases. Moreover, the solution (17) satisfies the symmetry relation with the plus sign if the point (μ_a, L) lies on one of the curves $L_n(\mu_a)$ with an odd index n and with the minus sign if the index is even. Indeed, let $L = L_n(\mu_a)$. We can use Eq. (17a) to write

$$i_1(L_n - z) = A(e^{\lambda(L_n - z)} - e^{-\lambda(L_n - z)}). \tag{18}$$

We now use Eqs. (15) and (16), which define the functions $L_n(\mu_a)$ for all possible indices and values of the argument, to find that

$$e^{\pm\lambda L_n} = (-1)^n \frac{\mu_a + q\mu_s \mp \lambda}{q\mu_s}. \tag{19}$$

To derive Eq. (19), one should be careful to use the correct branch of the arctangent in Eq. (15) (recall that the arctangent in this formula is defined so that $0 \le \arctan(z) < \pi$ for any real number z). Note that, if we square Eq. (19), we would obtain the resonance condition (14). This result follows immediately if we account for the identity

$$(\mu_a + q\mu_s - \lambda)(\mu_a + q\mu_s + \lambda) = (q\mu_s)^2. \tag{20}$$

Finally, we substitute Eq. (19) into Eq. (18), compare the result to Eq. (17b), and find that the solution (17) satisfies $i_1(L_n - z) = (-1)^{n+1} i_2(z)$.

We can now see that the solution (17) is unphysical for all even n. Indeed, the two streams cannot be simultaneously positive in this case. In fact, an even stronger condition holds: the solution is unphysical for all $n > 1$. Indeed, for $n > 1$, λ is imaginary, and we have $i_1(z) = 2iA \sin(|\lambda|z)$. We can select $A = A'/2i$, where $A' > 0$. Then the overall coefficient is positive. However, we also have $|\lambda|L_n > \pi$ for $n > 1$. Therefore, the function $\sin(|\lambda|z)$ necessarily changes sign in this case, and the solution cannot be physically valid.

We can conclude that the stationary solution given by Eq. (17) is physical only if the point (μ_a, L) lies on the curve $L_1(\mu_a)$. In this case, we can write $i_1(z) = A \sin(|\lambda|z)$ if $-2q\mu_s < \mu_a < 0$ and $i_1(z) = A \sinh(\lambda z)$ if $\mu_a < -2q\mu_s$. Here A is an arbitrary positive constant, and the total power generated by the system, e.g., $i_1(L) + i_2(0)$, is also arbitrary.

The stationary solution described above can be understood as the state of continuous-wave lasing. This interpretation may seem counterintuitive because lasers, generally, require a resonator. In the two-stream model discussed so far, there appears to be no resonator. Moreover, there is no electromagnetic phase and no constructive interference. However, as we have seen, there are *resonances*. The backward reflections and mutual conversion of the two streams work so as to create a semblance of a resonator. We will consider the role of the resonator and the more general condition of stationary lasing with reflection of the streams at the slab interfaces in Section 7 below.

5. TIME EVOLUTION IN AN AMPLIFYING MEDIUM

So far, we have considered only stationary solutions. We have seen that such solutions can be physical or unphysical. Apparently, the unphysical stationary solutions cannot be achieved by considering any transient process in which the initial state of both streams is zero. Questions arise about how a physical state of stable lasing can be excited, or what would happen if we apply external radiation to a medium with parameters (μ_a, L) for which stationary solutions do not exist or are unphysical. We are therefore motivated to look at the time evolution in the two-stream model. We note that time-dependent solutions to a general RTE with any physically reasonable source and initial conditions are unique and nonnegative, even for arbitrary negative μ_a [22]. This mathematical result applies to the two-stream theory of this paper.

The time-dependent two-stream equations are of the form

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial z} + \mu_a + q\mu_s\right) i_1(z,t) = q\mu_s i_2(z,t), \tag{21a}$$

$$\left(\frac{1}{c}\frac{\partial}{\partial t} - \frac{\partial}{\partial z} + \mu_a + q\mu_s\right) i_2(z,t) = q\mu_s i_1(z,t). \tag{21b}$$

We also assume that the streams satisfy the time-dependent boundary conditions

$$i_1(0,t) = S(t), \qquad i_2(L,t) = 0, \tag{22}$$

and the initial conditions $i_k(z,t) = S(t) = 0$ for $t \to -\infty$. Thus, the intensity of the external radiation incident onto the $z = 0$ face of the slab is described by the function $S(t)$, which tends to zero in the sufficiently distant past.

Unlike the stationary equations, the set [Eq. (21)] cannot be reduced to a one-dimensional diffusion equation. Rather, it is equivalent to the telegraph equation, which contains both first and second time derivatives [25].

It is tempting to try to solve Eq. (21) by temporal Fourier transform. This approach works fine if the medium is absorbing (i.e., $\mu_a > 0$), and even under some less restrictive conditions. However, for sufficiently negative values of μ_a or for sufficiently large width L for a given $\mu_a < 0$, the streams $i_k(z,t)$ can increase exponentially in time and, in this case, their temporal Fourier transforms do not exist, even in the sense of

generalized functions. One can still apply the Fourier transform technique formally to Eq. (21), but the result would be incorrect and, in particular, it will violate causality [1]. The mathematical difficulty here is subtle and may not be obvious when the actual calculations are done; it is the *a priori* assumption that certain Fourier transforms exist when, in fact, they might not, that is causing the trouble. We will therefore use the Fourier–Laplace transform instead, as was done previously in the case of Maxwell's equations [1,7,17].

Let us search for solutions to Eq. (21) in the form

$$i_{1,2}(z,t) = e^{-c\mu_a t} f_{1,2}(z,t). \tag{23}$$

The effect of this substitution is to eliminate μ_a from the differential equations. Indeed, the functions $f_{1,2}$ satisfy

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial z} + q\mu_s\right) f_1(z,t) = q\mu_s f_2(z,t), \tag{24a}$$

$$\left(\frac{1}{c}\frac{\partial}{\partial t} - \frac{\partial}{\partial z} + q\mu_s\right) f_2(z,t) = q\mu_s f_1(z,t). \tag{24b}$$

However, μ_a is now present in the boundary conditions

$$f_1(0,t) = e^{c\mu_a t} S(t), \qquad f_2(L,t) = 0. \tag{25}$$

It can be seen that the functions $f_{1,2}$ are solutions to the original two-stream equations (21) in a medium with $\mu_a = 0$ and with the modified source $S'(t) = e^{c\mu_a t} S(t)$. Let us assume that this modified source injects a finite energy into the medium so that $\int S'(t)\mathrm{d}t < \infty$. If $\mu_a = 0$, this energy is neither absorbed nor amplified and will eventually be radiated into free space through the slab surfaces without net loss or gain. In this case (and accounting for the positivity of $f_{1,2}$) we have $\int |f_{1,2}(z,t)|\mathrm{d}t < \infty$. Consequently, the functions can be Fourier-transformed with respect to t. Keep in mind, however, that, if $S(t)$ contains generalized functions, the same can be expected of $f_{1,2}(z,t)$.

We can now write the boundary condition in Fourier representation as

$$\tilde{f}_1(0,\omega) = F(\omega) \equiv \int_{-\infty}^{\infty} S(t) e^{(c\mu_a + i\omega)t}\mathrm{d}t, \qquad \tilde{f}_2(L,\omega) = 0. \tag{26}$$

Here we have used the Fourier transform convention,

$$\tilde{x}(\omega) = \int_{-\infty}^{\infty} x(t) e^{i\omega t}\mathrm{d}t, \; x(t) = \int_{-\infty}^{\infty} \tilde{x}(\omega) e^{-i\omega t}\frac{\mathrm{d}\omega}{2\pi}, \tag{27}$$

for all time-dependent functions. We can now solve Eq. (24) by Fourier transform. The result is

$$f_{1,2}(z,t) = \int_{-\infty}^{\infty} g_{1,2}(z,\omega) F(\omega) e^{-i\omega t}\frac{\mathrm{d}\omega}{2\pi}, \tag{28}$$

where $g_{1,2}(z,\omega)$ are the linear transfer functions. At every value of the Fourier variable ω, the problem of computing the transfer functions is formally equivalent to the stationary problem considered in Section 3, with the only difference being that μ_a should now be replaced by $-i\omega/c$. Therefore, we obtain the following expressions:

$$g_1(z,\omega) = \frac{a^-(\omega)e^{\lambda(\omega)(z-L)} - a^+(\omega)e^{\lambda(\omega)(L-z)}}{a^-(\omega)e^{-\lambda(\omega)L} - a^+(\omega)e^{\lambda(\omega)L}}, \tag{29a}$$

$$g_2(z,\omega) = q\mu_s \frac{e^{\lambda(\omega)(z-L)} - e^{\lambda(\omega)(L-z)}}{a^-(\omega)e^{-\lambda(\omega)L} - a^+(\omega)e^{\lambda(\omega)L}}, \tag{29b}$$

where

$$\lambda(\omega) = \sqrt{-i(\omega/c)(-i\omega/c + 2q\mu_s)}, \tag{30a}$$

$$a^{\pm}(\omega) = -i\omega/c + q\mu_s \pm \lambda(\omega). \tag{30b}$$

Let us further assume that the source is a delta-function pulse of the form $S(t) = \delta(t - t_0)$ and, correspondingly, $F(\omega) = e^{(c\mu_a + i\omega)t_0}$. Solutions to Eq. (21) with this delta source are the Green's functions of the time-dependent two-stream equations, which we denote by $G_{1,2}(z, t - t_0)$. In terms of the response functions, we have

$$G_{1,2}(z,\tau) = e^{-c\mu_a\tau}\int_{-\infty}^{\infty} g_{1,2}(z,\omega) e^{-i\omega\tau}\frac{\mathrm{d}\omega}{2\pi}. \tag{31}$$

Solutions for more general sources can be found by superposition, viz.,

$$i_{1,2}(z,t) = \int G_{1,2}(z,\tau) S(t-\tau)\mathrm{d}\tau. \tag{32}$$

Note the following properties of the transfer functions:

$$g_1(z,0) = 1 - \frac{q\mu_s z}{1 + q\mu_s L}, \qquad g_2(z,0) = \frac{q\mu_s(L-z)}{1 + q\mu_s L}. \tag{33}$$

From this, we conclude that $0 \le \int G_{1,2}(z,\tau) e^{c\mu_a\tau}\mathrm{d}\tau < 1$. The first inequality can be expected on physical grounds, since the streams are nonnegative by definition. The second inequality places a bound on how fast the solutions can grow with time in amplifying media. Note that $g_1(L,0) + g_2(0,0) = 1$. This equality means that all energy injected into a medium with $\mu_a = 0$ leaves eventually through the surfaces. This can be expected, since Eqs. (24) do not contain the absorption or amplification; the latter are accounted for by the exponential factor $e^{-c\mu_a\tau}$ in Eq. (31).

It is also useful to write out the first few terms in the ballistic expansion of the Green's functions. This expansion is of the form

$$G_{1,2}(z,\tau) = e^{-c(\mu_a + q\mu_s)\tau}\sum_{k=0}^{\infty} (q\mu_s)^k K_{1,2}^{(k)}(z,\tau), \tag{34}$$

and it can be derived by starting with the initial guess $i_1 = i_2 = 0$ in the right-hand side of Eq. (24) and iterating the resulting equations. The first two terms in this expansion are easy to compute, and they are given by the expressions

$$K_1^{(0)}(z,\tau) = \delta(\tau - z/c), \tag{35a}$$

$$K_2^{(1)}(z,\tau) = \frac{c}{2}\theta(\tau - z/c), \tag{35b}$$

$$K_1^{(2)}(z,\tau) = \frac{c}{2} z\theta(\tau - z/c), \tag{35c}$$

$$K_1^{(2k+1)}(z,\tau) = K_2^{(2k)}(z,\tau) = 0, \qquad k = 0, 1, 2\ldots, \tag{35d}$$

where $\theta(x)$ is the unit step function. The term $K_1^{(0)}$ describes the propagation of the nonscattered light (the precursor). We will refer to this term as to the *ballistic* component of the specific intensity. Note that $K_1^{(0)}(z,\tau)$ turns to zero identically when $\tau > L/c$. In the case of a more general excitation, the

ballistic component of the forward stream can be found from Eqs. (32) and (35a):

$$i_1^{(0)}(z,t) = e^{-(\mu_a+q\mu_s)z} S(t - z/c). \tag{36}$$

The second-order term $K_1^{(2)}(z,\tau)$ does not contain a singularity but is discontinuous at $\tau = z/c$. All higher-order terms $K_1^{(2k)}(z,\tau)$ are continuous. For this reason, Eq. (35c) yields the exact result for the jump of the Green's function $G_1(z,\tau)$ at $\tau = z/c$: the jump is equal to $\frac{1}{2}c(q\mu_s)^2 z e^{-(\mu_a+q\mu_s)z}$.

The Green's function $G_2(z,\tau)$ and the reverse stream $i_2(z,t)$ do not have ballistic components and, correspondingly, $K_2^{(0)}(z,\tau) = 0$. This is, of course, due to the assumption that the exciting pulse enters the medium only through the surface $z = 0$. However, $G_2(z,\tau)$ still has a discontinuity at $\tau = z/c$. Again, Eq. (35b) predicts the jump correctly, since all higher-order terms are continuous.

We could easily derive several higher-order terms in Eq. (34); all these terms are causal; that is, they turn identically to zero for $\tau < z/c$. The ballistic expansion is useful at small times with the characteristic time scale being $1/q\mu_s c$. We therefore have shown that, at small times, the response of the system is causal. We are however more interested in the long-time behavior. To this end, it is instructive to evaluate the Fourier transform (31) by contour integration. Although the poles of the transfer functions cannot be computed analytically, the resulting semianalytical formulas will provide an important insight.

The first thing to note is that $g_{1,2}(z,\omega)$ are single-valued functions of ω in spite of the presence of square roots in the definitions (29) and (30). This is similar to the case of transmission and reflection coefficients of homogeneous slabs computed from Maxwell's equations [1,2]. In particular, when $\omega \to \pm\infty$ (on the real axis), we have

$$g_1(z,\omega) \underset{\omega\to\pm\infty}{\longrightarrow} e^{\left(i\frac{\omega}{c}-q\mu_s\right)z}. \tag{37}$$

Therefore, $G_1(z,\omega)$ contains a singularity, which must be extracted before computing the Fourier transform (31). This singularity is exactly equal to the ballistic term given in Eq. (35a). Let us adopt the following notations:

$$g_1(z,\omega) = g_1^{(0)}(z,\omega) + g_1^{(d)}(z,\omega), \tag{38a}$$

$$g_1^{(0)}(z,\omega) = e^{\left(i\frac{\omega}{c}-q\mu_s\right)z}. \tag{38b}$$

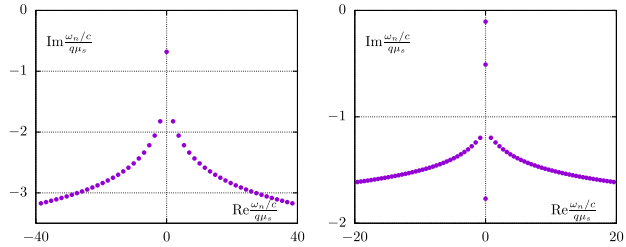
The corresponding *ballistic* and *diffuse* parts of the Green's function $G_1^{(0)}$ and $G_1^{(d)}$ are obtained by computing the Fourier transforms of $g_1^{(0)}$ and $g_1^{(d)}$ according to Eq. (31). Referring to expansion (35), we can identify

$$G_1^{(d)}(z,\tau) = G_1(z,\tau) - e^{-c(\mu_a+q\mu_s)\tau}\delta(\tau - z/c). \tag{39}$$

Viewed as functions of ω, $g_1^{(d)}(z,\omega)$ and $g_2(z,\omega)$ are casual linear transfer functions. Both have Fourier transforms and no singularities on or above the real axis. In the lower open half-plane, the functions are analytic everywhere except for a countable infinite set (same for both functions) of simple poles $\omega_n(L)$, which accumulate at infinity. The poles can be found by solving Eq. (14) for a fixed value of L treating μ_a as the unknown. The equation has infinitely many solutions $\mu_{an}(L)$ but only a finite number of them are real-valued. These real roots have been discussed in Sections 3 and 4. Now we require all solutions to Eq. (14), including the complex ones. The frequencies $\omega_n(L)$ are related to $\mu_{an}(L)$ by $\omega_n(L) = ic\mu_{an}(L)$. The argument L in $\omega_n(L)$ and $\mu_{an}(L)$ will be omitted below for simplicity. The poles are shown in Fig. 4 for two different slab widths. It can be seen that, for $L = 2/q\mu_s L$, only one of the roots μ_{an} is purely real. Correspondingly, only one of the frequencies ω_n is purely imaginary; the rest are complex. For $L = 6/q\mu_s$ there are three purely real roots μ_{an} and three purely imaginary frequencies ω_n.

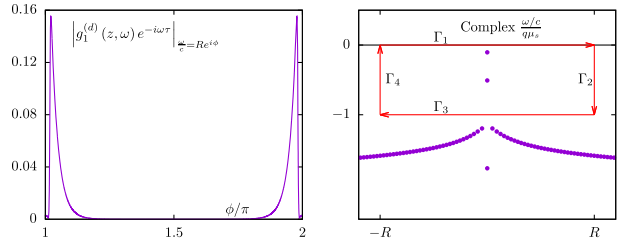
When performing contour integration according to Eq. (31), we can close the loop in the upper half-plane if $\tau < z/c$, which obviously yields zero. This guarantees causality of the response. However, when $\tau > z/c$, the situation is not so simple. If we close the integration loop in the lower half-plane using the semicircle $\omega = cq\mu_s Re^{i\phi}$, $\pi \le \phi \le 2\pi$, we would enclose a finite number of poles inside the integration contour. However, the integral over the semicircle does not tend to zero when $R \to \infty$, because the integrand has sharp spikes when the integration path crosses the "trajectories" of the complex poles, as is illustrated in Fig. 5(a). These spikes have nonzero integral weight. Related to the above, the sums of residues computed at the poles enclosed inside the above integration path does not converge when $R \to \infty$. It is true that the nonconverging part of the sum and the integral over the circular arc tend to zero exponentially with τ. However, for any nonzero τ, these quantities are nonzero.

To avoid dealing with diverging series, we can adopt the following approximation, which yields exponentially accurate



(a) $L = 2/q\mu_s$ (b) $L = 6/q\mu_s$

Fig. 4. Complex frequencies ω_n for $L = 2/q\mu_s$ (a) and $L = 6/q\mu_s$ (b). The patterns seen in the plots are infinite and can be continued indefinitely beyond the figure frames.



(a) Integrand on lower semicircle (b) Rectangular integration path

Fig. 5. (a) Plot of $|g_1^{(d)}(z,\omega)e^{-i\omega\tau}|$ on the lower semicircle $\frac{\omega/c}{q\mu_s} = Re^{i\phi}$ as a function of ϕ for the following parameters: $R = 60$, $L = 2/q\mu_s$, $z = 0.5L$, $\tau = 0.6L/c$. (b) Rectangular integration path discussed in the text for $\beta = 1$. Poles are shown for $L = 6/q\mu_s$.

results when $\tau \to \infty$. Let $\tau > z/c$. We then close the integration contour, as is shown in Fig. 5(b). The integration path is the rectangle with the vertices at $(-R, 0)$ to $(R, 0)$, $(R, -\beta)$, $(-R, -\beta)$ in the complex $\frac{\omega/c}{q\mu_s}$ plane. It can be seen that the segment Γ_3 passes clear of any poles. This is different from the case of a semicircle, which necessarily crosses the regions where the poles are located very close to each other and form an almost continuous "trajectory." As a result, the integral over Γ_3 is not zero but small and can definitely be neglected when $\tau \to \infty$ (integrals over Γ_2 and Γ_4 tend to zero when $R \to \infty$ and can always be neglected). The optimal choice of the parameter β depends on L. In relatively thin slabs, one might need to take $\beta \geq 2$ so that at least one pole is inside the rectangle. For $L > 2/q\mu_s$, one can safely use $\beta = 1$, as is shown in the figure. This choice yields an especially simple integral over the segment Γ_2. For our purposes, computing this integral is unimportant; we just assume that it is negligible at large times.

We finally need to compute the residues of $g_1^{(d)}(z, \omega)$, which is a cumbersome but straightforward task. The result is

$$\lim_{\omega \to \omega_n} (\omega - \omega_n) g_1^{(d)}(z, \omega) = \frac{c\lambda_n}{2i} \frac{e^{\lambda_n z} - e^{-\lambda_n z}}{1 + (q\mu_s + \mu_{an})L}. \tag{40}$$

We do not need to compute the residues of $g_2(z, \omega)$, as they can easily be found by applying Eq. (24a) to Eq. (40). Putting everything together, we obtain the following result:

$$G_1^{(d)}(z, \tau) \approx \frac{c}{2} \theta\left(\tau - \frac{z}{c}\right) e^{-c\mu_a \tau} \times \sum_n \theta\left(\frac{\mu_{an}}{q\mu_s} + \beta\right) e^{c\mu_{an}\tau} \frac{\lambda_n (e^{-\lambda_n z} - e^{\lambda_n z})}{1 + (q\mu_s + \mu_{an})L}, \tag{41a}$$

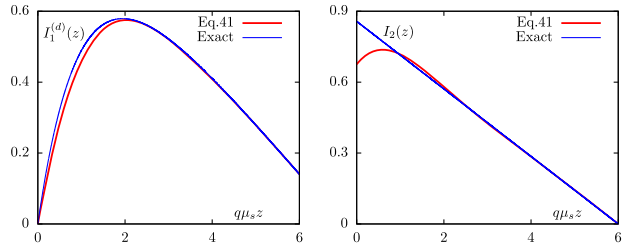
$$G_2(z, \tau) \approx \frac{c}{2} \theta\left(\tau - \frac{z}{c}\right) e^{-c\mu_a \tau} \times \sum_n \theta\left(\frac{\mu_{an}}{q\mu_s} + \beta\right) e^{c\mu_{an}\tau} \frac{\lambda_n (a_n^- e^{-\lambda_n z} - a_n^+ e^{\lambda_n z})}{q\mu_s [1 + (q\mu_s + \mu_{an})L]}. \tag{41b}$$

Thus, the nature of the approximation is quite simple: truncate the summation over the infinite set of poles so that the cutoff is drawn in the widest gap separating one group of poles from another. The line Γ_3 in Fig. 5 is such a cutoff. It separates the poles with relatively small imaginary parts, which dominate the large-time behavior from the rest of the poles.

The quality of the approximation given by Eq. (41) is illustrated in Fig. 6, where we plot the integrals

$$I_1^{(d)}(z) = \int e^{\alpha\tau} G_1^{(d)}(z, \tau) d\tau, \qquad I_2(z) = \int e^{\alpha\tau} G_2(z, \tau) d\tau \tag{42}$$

as functions of z in a slab of the width $L = 6/q\mu_s$. Exact results inferred from Eq. (33) are compared to the approximation (41). For the parameters considered, the sum (41) contains only two terms, which correspond to the two poles inside the integration loop in Fig. 5(b). However, the agreement is excellent at relatively large values of z, which correspond to sufficiently large propagation times. In fact, the approximation for $I_1^{(d)}(z)$ is



(a) $I_1^{(d)}(z)$ (b) $I_2(z)$

Fig. 6. Integrals (42) for $I_1^{(d)}(z)$ (a) and $I_2(z)$ (b) in a slab of the width $L = 6/q\mu_s$ computed approximately according to Eq. (41) and exactly according to Eq. (33). Note that the exact result for $I_1^{(d)}(z)$ is obtained from the first equation in Eq. (33) by subtracting the additional term $e^{-q\mu_s z}$.

quite good even at small-to-moderate values of z. This is so because we have already extracted analytically the singular part of the Green's function, which cannot be captured accurately by an expansion of the form (41). For $I_2(z)$, there is no singular part, and the approximation is not as accurate at small distances. However, note that Eq. (35b) predicts the value of $I(0)$ exactly.

Approximation (41) is instructive for understanding the long-time evolution of the system. It can be seen that the response to a delta pulse is always transient if $\mu_a > 0$. That is, both $G_1(z, \tau)$ and $G_2(z, \tau)$ tend to zero exponentially when $\tau \to \infty$. If $\mu_a < 0$, time evolution can still be transient as long as $\mu_a > \max_n[\mathrm{Re}(\mu_{an})] = \max_n[\mathrm{Im}(\omega_n/c)]$. We note that the maximum is always achieved for the real-valued root $\mu_{a1}(L)$, which is discussed in detail in Section 3. In other words, the equality $\mu_a = \max_n[\mathrm{Re}(\mu_{an})]$ holds if and only if the two parameters μ_a and L lie on the lowest (emphasized) curve $L = L_1(\mu_a)$ in Fig. 2. If the parameters are above and to the left of this curve, then $\mu_a < \max_n[\mathrm{Re}(\mu_{an})]$. In this case, the solution (41a) is not transient; rather, it exhibits unbounded exponential growth with τ. One can say that the system is in this case unstable, as an arbitrarily small excitation in a distant past can create an arbitrarily large intensity at the time of observation.

To summarize, the response to a delta-pulse excitation is transient and stable if and only if the parameters (μ_a, L) are to the right of and below the emphasized curve $L_1(\mu_a)$ in Fig. 2. If the parameters are above and to the left of this curve, the response is not transient and experiences unbounded exponential growth. However, if this pair of parameters is exactly on the curve $L_1(\mu_a)$, the response become stationary at large times, which corresponds to a state of stable lasing.

Let us consider the latter occurrence in more detail. Let $\alpha = \max_n[\mathrm{Re}(\alpha_n)] = c\mu_{a1}(L) < 0$. Then, at large times, the only term that will survive in Eq. (41a) is $n = 1$. We then obtain

$$G_1(z, \tau) \underset{\tau \to \infty}{\to} \frac{c}{2} \frac{\lambda_1 (e^{-\lambda_1 z} - e^{\lambda_1 z})}{1 + (q\mu_s + \mu_{a1})L}. \tag{43}$$

The above result applies to the total Green's function, not only to its diffuse component, since the ballistic term (the precursor) is zero at large times. Assuming that the slab is sufficiently thick so that $\mu_{a1} > -2q\mu_s$, we have $\lambda_1 = i|\lambda_1|$ so that the limit becomes

$$G_1(z,\tau) \underset{\tau\to\infty}{\to} \frac{c|\lambda_1|\sin(|\lambda_1|z)}{1+(q\mu_s+\mu_{a1})L}. \tag{44}$$

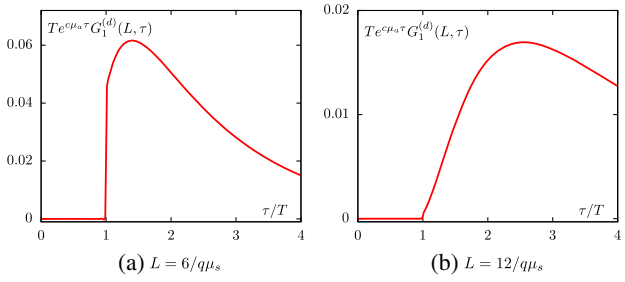
The denominator in this expression is positive. This is the sinusoidal solution corresponding to stable lasing that was found previously in Section 4. The positive z-independent coefficient in Eq. (44) quantifies the achieved lasing power due to a unit power injected into the system in the distant past. Also, as was discussed in Section 4, the sine in Eq. (44) does not change sign, so that the solution is positive for all z, as expected.

Finally, if the slab is sufficiently thin so that $\mu_{1a} < -2q\mu_s$, the exponent λ_1 is real and positive so that the hyperbolic stationary lasing solution of the form $G_1 \propto \sinh(\lambda_1 z)$, which was also discussed in Section 4, is obtained.

6. NUMERICAL EXAMPLES

A practical approach to accurate and efficient evaluation of the integral (31) that is valid at all times is simply by the trapezoidal rule. In Fig. 7, we plot the *exponentially scaled* time-dependent diffuse transmission $e^{c\mu_a\tau}G_1^{(d)}(L,\tau)$ for different widths of the slab L. The diffuse component is shown because the ballistic contribution is a delta pulse, which is difficult to represent graphically. One should, however, keep in mind that the ballistic component is always present, although, in sufficiently thick slabs, it becomes negligible. The ratio of the integral weights of the ballistic component and diffuse components is (for transmission) $(1+q\mu_sL)/(e^{q\mu_sL}-1-q\mu_sL)$. For $L=6/q\mu_sL$, this ratio is ≈ 0.02 and for $L=12/q\mu_sL$, it is $\approx 8\cdot 10^{-5}$. In time-resolved experimental measurements, the precursor can appear as a sharp spike of relatively small integral weight centered at $\tau=T$, where $T=L/c$ is the time of flight through the slab [26].

The scaling by the function $e^{c\mu_a\tau}$ is used in the figure to make the displayed functions independent of μ_a. More precisely, the functions that are shown correspond to a medium with $\mu_a=0$. To obtain the corresponding functions in a medium with $\mu_a\neq 0$, all one has to do is multiply the scaled functions by $e^{-c\mu_a\tau}$. If μ_a is negative, the exponential growth of $e^{-c\mu_a\tau}$ can overcome the exponential decay of the scaled functions. In this case, the transmission will experience an unbounded exponential growth with τ. If the rate of increase of $e^{c\mu_a\tau}$ matches exactly the rate of decay of the scaled functions, the system will evolve over time into a state of stationary lasing. Exact compensation occurs when the parameters (μ_a, L) lie on the curve $L_1(\mu_a)$ in Fig. 2. It can be seen that the rate of exponential decay of the scaled transmission is slower in a wider slab. Correspondingly, smaller gain is required to compensate for this decay.



(a) $L=6/q\mu_s$ (b) $L=12/q\mu_s$

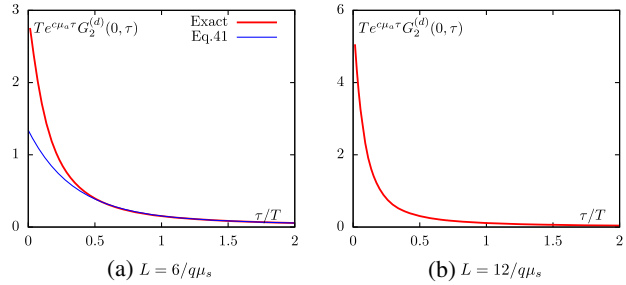
Fig. 7. Time dependence of the exponentially scaled diffuse transmission, $e^{c\mu_a\tau}G_1^{(d)}(L,\tau)$, for $L=6/q\mu_s$ (a) and $L=12/q\mu_s$ (b). The normalization factor $T=L/c$ is the time of ballistic propagation of light through the slab (time of flight). The ballistic component of the transmission (the precursor) is not shown and should be accounted for separately.

Both curves shown in Fig. 7 experience a discontinuity at $\tau=T$. The jump magnitudes are given exactly by Eq. (35c), where we should use $z=L$. As expected, the jump goes to zero with L, and the discontinuity is barely visible in the case $L=12/q\mu_sL$.

The data points of Fig. 7 were obtained numerically by using the trapezoidal rule. The data points match with very high accuracy the approximate result given by Eq. (41a) with only two terms in the summation, which correspond to the two poles inside the integration loop of Fig. 5(b). The solutions match almost perfectly and cannot be visually distinguished. This means that, at $z=6/q\mu_s$, the solution is exponentially dominated by the two first terms in Eq. (41a). One, however, needs to be careful and not include contributions from the poles that lie below the cutoff line Γ_3 in Fig. 5(b). Including such contributions will worsen the precision and eventually yield a wrong result.

Reflection is less interesting than transmission (see Fig. 8). Since the Green's function $G_2(0,\tau)$ is evaluated in this case at $z=0$, there are no discontinuities in time domain, or, rather, the discontinuity occurs exactly at $\tau=0$. The magnitudes of the displayed functions at $\tau=0$ can be found analytically from Eq. (35b), which approaches the exact result at $\tau\to 0$. Specifically, we have $TG_2(0,0)=\frac{1}{2}q\mu_sL$, which is, indeed, the case, as can be seen from the figure. However, the numerical integration becomes unstable very close to $\tau=0$, and we have suppressed the left-most data points in the plots of Fig. 8. Trying to compute these points numerically is similar to trying to compute (e.g., by some integral or expansion) the value of a step function exactly at the point where it is discontinuous.

The reflection data computed numerically by the trapezoidal rule are noticeably different from the approximate result (41b), as is shown in Fig. 8(a). The difference is, however, visible only at the relatively small times. As noted above, the approximation (41) is less accurate for the reverse than for the forward stream. The reason is that we were able to subtract the ballistic component of the forward stream analytically from the solution and then treat it separately. There is no analogous manipulation that is applicable to the reverse stream.



(a) $L=6/q\mu_s$ (b) $L=12/q\mu_s$

Fig. 8. Same as in Fig. 7 but for the exponentially scaled reflection, $e^{c\mu_a\tau}G_2(0,\tau)$.

One final remark about the time dependence of reflection needs to be made. It may appear that the scaled reflection decays exponentially with τ on the time scale of T. However, this is not so. The functions have wide, slowly decaying tails, which match the transmission functions shown in Fig. 8 in the decay rate and amplitude. As can be expected, the specific intensity inside the slab becomes symmetric at large times, and both directions of propagation become equivalent. This implies that $G_2(0,\tau) \to G_1(L,\tau)$ when $\tau \to \infty$ (at large times, there is no difference between G_1 and $G_1^{(d)}$). The stationary lasing and instability conditions are, of course, the same for transmission and reflection.

7. ROLE OF THE RESONATOR

As was noted above, the boundary condition (7) does not account for the possibility of Fresnel reflections from the slab interfaces. A more general boundary condition (for the stationary equations) that takes the reflections into account is of the form

$$i_1(0) = 1 + |r_0|^2 i_2(0), \qquad i_2(L) = |r_L|^2 i_1(L), \tag{45}$$

where r_0 and r_L are the *internal* Fresnel reflection coefficients at the $z = 0$ and $z = L$ interfaces for the field amplitudes (not to be confused with the power reflection coefficient R of the slab as a whole). If one imagines that the regions $z < 0$, $0 < z < L$, and $z > L$ have the average refractive indices of n_1, n, and n_2, respectively, then $r_0 = (n - n_1)/(n + n_1)$ and $r_L = (n - n_2)/(n + n_2)$. It is sufficient to assume for our purposes that n_1 and n_2 are real-valued, while $n = n' + in''$ can have a positive or negative imaginary part n'' corresponding to either absorption or amplification. In this case, one can write $\mu_a = 2k_0 n''$, where $k_0 = \omega_0/c$ is the wavenumber at the central frequency of quasi-monochromatic radiation. We note that, in Eq. (45), the coefficients r_0 and r_L are squared because the streams $i_1(z)$, $i_2(z)$ describe the propagation of energy; for the same reason, the relation between μ_a and n'' contains the factor of 2.

It is instructive to start with the simple case when $q\mu_s = 0$ so that there is no interaction of the streams inside the medium and the only process that can result in energy exchange between the two streams is reflection at the interfaces. Physically, this case corresponds to a perfectly homogeneous medium in which there is no scattering. The solution to the stationary two-stream equations (6) with the boundary condition (45) is

$$i_1(z) = \frac{e^{\mu_a(2L-z)}}{e^{2\mu_a L} - |r_0 r_L|^2}, \qquad i_2(z) = \frac{|r_L|^2 e^{\mu_a z}}{e^{2\mu_a L} - |r_0 r_L|^2}. \tag{46}$$

It can be seen that the resonance condition becomes in this case $e^{2\mu_a L} = |r_0 r_L|^2$ or, accounting for the relation $\mu_a = 2k_0 n''$,

$$e^{2k_0 n'' L} = |r_0 r_L|. \tag{47}$$

This is the easily recognizable condition of stationary lasing: the energy gain on double passage of a resonator should be equal to the energy loss due to transmission through (and absorption) at its mirrors. Obviously, the condition (47) can hold only if $n'' < 0$ and then only if $0 < |r_0 r_L| < 1$. It can further be seen that, if $e^{2k_0 n'' L} > |r_0 r_L|$, then both streams are positive. In this case, physically meaningful stationary solutions exist. If $e^{2k_0 n'' L} < |r_0 r_L|$, then Eq. (46) predicts negative values of $i_1(z)$, $i_2(z)$. This means that, in this case, stationary solutions do not exist, and the system is unstable.

In regard to the various possible values of the reflection coefficients, three special cases can be mentioned. The case $r_0 = r_L = 0$ (the resonator is effectively absent) was considered in all previous sections. The condition (47) tells us that lasing in the absence of a resonator and with no internal scattering (recall that $q\mu_s = 0$ at the moment) can occur only in the limit $L \to \infty$, which is not physical. The second special case, $|r_0 r_L| = 1$, corresponds to perfectly reflecting mirrors and an *ideal* resonator. In this case, an infinitesimally small amount of amplification will give rise to exponentially growing instabilities. Of course, we know that there are no perfect mirrors in nature, so that this case is also unphysical. Finally, the third special case one can consider is $r_0 = 0$ and $r_L \neq 0$. In this case, the resonator is effectively open, and there is no lasing in any finite system. We note that, on physical grounds, the lasing condition should be invariant with respect to the interchange $r_0 \longleftrightarrow r_L$ and Eq. (47) indeed satisfies this condition. We will see below that the permutation invariance holds in a more general case as well.

We thus have rederived by a somewhat unconventional method the well-known result that, according to the condition (47), a resonator is required for lasing. In contrast, the previously derived lasing condition (14) does not require any resonator. We reiterate that there is no paradox or contradiction here. The condition (47) was obtained under the assumption that there is no internal scattering in the medium and that stream conversion can occur only at the interfaces $z = 0$ and $z = L$ due to Fresnel reflections. The condition (14) was obtained under the assumption that stream conversion does not occur at the interfaces (there are no Fresnel reflections) but can occur anywhere inside the medium due to scattering. Thus, we come to the conclusion that lasing requires, in general, some physical mechanism for stream conversion; it does not matter whether it occurs at the interfaces as in conventional lasers or inside the medium (as one can interpret the action of random lasers).

We will now consider the case when the stream conversion can occur both at the interfaces and inside the medium. In other words, we abandon the simplifying assumption $q\mu_s = 0$ but still use the more general boundary condition (45). Omitting the intermediate details, the condition of stationary lasing in this case is of the form (we use μ_a instead of n'' everywhere below)

$$e^{2\lambda L} = \frac{\mu_a + (1 - |r_0|^{-2})q\mu_s + \lambda}{\mu_a + (1 - |r_0|^{-2})q\mu_s - \lambda} \times \frac{\mu_a + (1 - |r_L|^2)q\mu_s - \lambda}{\mu_a + (1 - |r_L|^2)q\mu_s + \lambda}. \tag{48}$$

This expression is invariant with respect to the permutation $r_0 \longleftrightarrow r_L$ but to see this, we need to account for the specific form of λ given in Eq. (12b). In the limit $r_0, r_L \to 0$, Eq. (48) is equivalent to Eq. (14) and in the limit $q\mu_s \to 0$, it is equivalent to Eq. (47). To obtain the latter result, one needs to expand λ as $\lambda = \pm(\mu_a + q\mu_s) + O((q\mu_s)^2)$.

Just as was done in Section 3, we can view Eq. (48) as an equation defining an infinite number of curves $L_n(\mu_a)$, $n = 1, 2, \ldots$ in the real (μ_a, L) plane. The lowest of these curves,

$L_1(\mu_a)$, is the locus of all points for which stationary laser generation is possible. For all points below and to the right of this curve, stationary solutions with a nonzero external source are possible, and for all points above and to the left, stationary solutions are not possible even without external excitation. In Fig. 9, we illustrate the condition of stationary lasing for various values of the reflection coefficients r_0 and r_L, including the case $r_0 = r_L = 0$. It can be seen that allowing reflections to occur at the slab interfaces tends to relax the lasing condition. In other words, lasing is ignited for smaller values of L for a given $\mu_a < 0$ and for larger values of μ_a for a given L when $|r_0|^2$ and $|r_L|^2$ are equal to each other and monotonically increased towards the limiting value of unity. If $|r_0|^2$ and $|r_L|^2$ are different, the situation is somewhat more complicated, as can be seen from the intersection of the curves labeled 2 and 4 in Fig. 9. However, increasing one of the reflection coefficients while keeping the other unchanged always relaxes the lasing condition.

The data of Fig. 9 do not allow one, at least easily, to visualize the dependence of the stationary lasing condition on $q\mu_s$. In particular, it is not possible to infer the limit $q\mu_s \to 0$ from this figure. This is so because, as in all previous figures, we have used $(q\mu_s)^{-1}$ as the unit of length in Fig. 9. To visualize the dependence on $q\mu_s$, we have displayed in Fig. 10 the condition of stationary lasing as curves in the $(q\mu_s, L)$ plane for a fixed negative $\mu_a < 0$. Again, it can be seen that allowing reflections at the interfaces to occur relaxes the lasing condition.

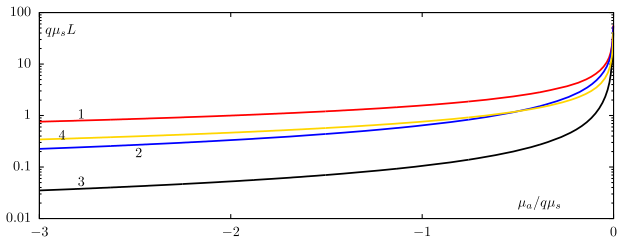


Fig. 9. Illustration of the condition of stationary lasing (48). The lowest-lying solutions of Eq. (48) are shown for varying values of the reflection coefficients $|r_0|^2$ and $|r_L|^2$. The curve labeled as 1 corresponds to $r_0 = r_L = 0$ and is identical to the curve $L_1(\mu_a)$ shown in Fig. 2. Other curves have the following parameters: $|r_0|^2 = |r_L|^2 = 0.5$ (Curve 2); $|r_0|^2 = |r_L|^2 = 0.9$ (Curve 3); and $|r_0|^2 = 0.05$, $|r_L|^2 = 0.95$ (Curve 4).

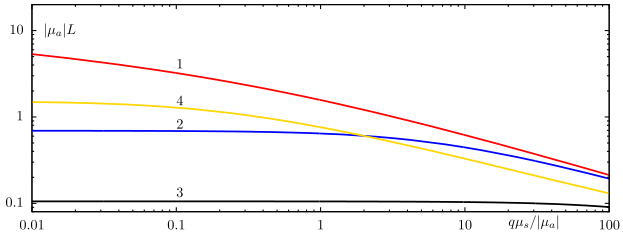


Fig. 10. Same as in Fig. 9 but the stationary lasing condition is shown in the $(q\mu_s, L)$ plane for a fixed negative value $\mu_a < 0$. Same curve labels as in Fig. 9. Note that the curve labeled 1, which corresponds to $r_0 = r_L = 0$, exhibits a logarithmic divergence at $q\mu_s = 0$.

8. DISCUSSION

The simple two-stream approximation to the RTE is a convenient model for studying propagation of light in amplifying media. An added bonus of this approach is that the unphysical stationary solutions, which are not always easy to spot in the case of Maxwell's equations, become apparent in the case of the RTE because the specific intensity in all such instances changes sign, which is clearly unphysical.

Three-dimensional generalizations of the theory developed above can be applicable to random lasers. It should be also kept in mind that the RTE is equally applicable to neutron transport. The critical thickness of an amplifying slab beyond which the system response is unstable is directly analogous to the critical mass in nuclear fusion, beyond which an uncontrolled chain reaction occurs. Although this phenomenon has been very well understood in nuclear physics, the analogy with optics may not be obvious. The main reason is that the RTE has no phase, whereas interference effects are often viewed as conceptually important in the functioning of lasers.

Of course, the instabilities and the unbounded growth of intensity are artifacts of an incomplete mathematical model. In real systems, the growth of intensity is always limited by the effects of saturation and other kinds of negative feedback. For example, the competition of gain and losses and the effect of saturation play an important role in conventional lasers. However, the laser medium is not described in this case by a linear RTE; one must use a nonlinear equation where the gain is saturated at high energy density. For example, we can use the model expression

$$\mu_a(u) = a - \frac{b}{1 + u/u_0}, u(z) = i_1(z) + i_2(z). \tag{49}$$

Here a and b are two positive constants, and u is the density of energy. It can be seen that the coefficient μ_a becomes in this case position-dependent, since $u(z)$ will change with position. It can be concluded that, in the region of parameters where the medium response is unstable, linear or weak-field approximation is not applicable in principle.

In light of the above, we can analyze the arguments appearing in Wang *et al.*'s reply [14] to the comment by Baranov *et al.* [13]. In particular, Wang *et al.* write "The main argument of Baranov et al. that 'the movement of SPs [singular points-V.M.] … indicates the onset of lasing' implies that the onset of lasing does not depend on the field strength but relies on the slab's thickness. This incorrect argument is different from the conventional view on lasing …." In the foregoing text, Wang *et al.* also argue that the slab width at which the onset of lasing occurs, denoted by d_{las} in the discussed references, depends on the intensity of incident light: "The lower the intensity of the incident light is, the larger the width d_{las} will be."

The results of this paper show clearly that the above arguments are incorrect. The critical width of the slab beyond which the system is unstable *does not* depend on the intensity of incident light. In fact, the instabilities will develop in the absence of any incident light; a tiny thermal fluctuation will get exponentially amplified over time. This phenomenon is physically similar to the uncontrolled chain fusion reaction and is qualitatively similar to all other instabilities in systems

with positive feedback. Even though we did not use Maxwell's equations to arrive at this conclusion, the mathematics in both cases is exactly the same: it is the motion of the poles of the transfer function (SPs in the terminology of the discussed references) that governs the transition from a linear regime in which a weak-field approximation is indeed applicable to the regime of stationary lasing and further to the regime of unstable generation.

Finally, the question of "resolving the index of refraction" in amplifying media [1,3,5] [determining the "correct" branch of the square root in Eq. (12b) in the context of the theory developed above] is of philosophical rather than of physical interest. In any finite sample, the solutions are independent of this choice, and therefore, no prediction of a physical observable depends on it. The question can be relevant to a semi-infinite medium. In conventional passive media, a semi-infinite sample is a useful idealization of a very thick slab; the question of determining the refractive index correctly in this case makes physical sense. However, in the case of even arbitrarily small gain, the model of semi-infinite medium is incompatible with the assumption of linearity. Linear equations (either Maxwell's or the RTE) do not have physically valid stationary solutions in semi-infinite amplifying samples. The idealization is in this case not useful or physically applicable, and the notion of index of refraction is equally inapplicable. Of course, it should always be kept in mind that the instabilities can realistically propagate not only in the direction *perpendicular* to the slab but also *along* the slab. Therefore, the linear (or weak-field) approximation is applicable only to sufficiently small samples that have finite size in all three spatial dimensions. An infinite slab (even of finite width) with any amount of gain is intrinsically unstable, and linear weak-field approximation is not applicable to it in principle.

Funding. Agence Nationale de la Recherche (ANR) (ANR-11-IDEX-0001-02).

REFERENCES AND NOTES

1. J. Skaar, "On resolving the refractive index and the wave vector," Opt. Lett. **31**, 3372–3374 (2006).
2. J. Skaar, "Fresnel equations and the refractive index of active media," Phys. Rev. E **73**, 026605 (2006).
3. V. U. Nazarov and Y.-C. Chang, "Resolving the wave vector and the refractive index from the coefficient of reflectance," Opt. Lett. **32**, 2939–2941 (2007).
4. J. O. Grepstad and J. Skaar, "Total internal reflection and evanescent gain," Opt. Express **19**, 21404–21418 (2011).
5. S. A. Afanas'ev, D. G. Sannikov, and D. I. Sementsov, "The refractive index sign chosen for amplifying and lossy metamaterials," J. Commun. Technol. Electron. **58**, 5–15 (2013).
6. L.-G. Wang, L. Wang, M. Al-Amri, S.-Y. Zhu, and M. S. Zubairy, "Counterintuitive dispersion violating Kramers-Kronig relations in gain slabs," Phys. Rev. Lett. **112**, 233601 (2014).
7. H. O. Hagenvik and J. Skaar, "Fourier-Laplace analysis and instabilities of a gainy slab," J. Opt. Soc. Am. B **32**, 1947–1953 (2015).
8. H. O. Hagenvik, M. E. Malema, and J. Skaar, "Fourier theory of linear gain media," Phys. Rev. A **91**, 043826 (2015).
9. A. K. Jahromi and A. F. Abouraddy, "Observation of Poynting's vector reversal in an active photonic cavity," Optica **3**, 1194–1200 (2016).
10. A. K. Jahromi, S. Shabahang, H. E. Kondakci, S. Orsila, P. Melanen, and A. F. Abouraddy, "Transparent perfect mirror," ACS Photon. **4**, 1026–1032 (2017).
11. M. Perez-Molina and L. Carratero, "Comment on "Resolving the wave vector and the refractive index from the coefficient of reflectance"," Opt. Lett. **33**, 1828 (2008).
12. V. U. Nazarov and Y.-C. Chang, "Reply to Comment on "Resolving the wave vector and the refractive index from the coefficient of reflectance"," Opt. Lett. **33**, 1829 (2008).
13. D. G. Baranov, A. A. Zyablovsky, A. V. Dorofeenko, A. P. Vinogradov, and A. A. Lisyansky, "Comment on "Counterintuitive dispersion violating Kramers-Kronig relations in gain slabs"," Phys. Rev. Lett. **114**, 089301 (2015).
14. L.-G. Wang, L. Wang, M. Al-Amri, S.-Y. Zhu, and M. S. Zubairy, "Wang *et al.* reply," Phys. Rev. Lett. **114**, 089302 (2015).
15. A. A. Kolokolov, "Reflection of plane waves from an amplifying medium," Pis'ma v ZhETF **21**, 660–662 (1975).
16. L. A. Vainshtein, "Propagation of pulses," Uspekhi Fiz. Nauk **118**, 339–367 (1976).
17. A. V. Dorofeenko, A. A. Zyablovsky, A. A. Pukhov, A. A. Lisyansky, and A. P. Vinogradov, "Light propagation in composite materials with gain layers," Phys. Usp. **55**, 1080–1097 (2012).
18. Here we discuss only the scalar RTE that describes the specific intensity of depolarized light. Vector RTE (vRTE) describes propagation of the Stokes vector (I, Q, U, V) where some of the components Q, U, V can become negative while I is always positive.
19. P. Kubelka and F. Munk, "Ein Beitrag zur Optik der Farbanstriche," Z. Tech. Phys. **12**, 593–601 (1931).
20. It should be kept in mind that $\delta_2(\hat{\mathbf{s}}, \hat{\mathbf{s}}')$ is always a function of two arguments. For example, the expression $\delta_2(\hat{\mathbf{s}})$ does not make mathematical sense. In this respect, δ_2 is different from the ordinary delta function.
21. R. Aronson, "Boundary conditions for diffuse light," J. Opt. Soc. Am. A **12**, 2532–2539 (1995).
22. K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, 1967).
23. The function $L_1(\mu_a)$ is infinitely differentiable everywhere on the open half-axis $\mu_a < 0$ except for the point $\mu_a = -2q\mu_s$. Expressions (15) with $n = 1$ and (16) have different first derivatives at this point. The relevant expansions are $L_1(\mu_a) \approx q\mu_s + (5/3)(\mu_a + 2q\mu_s)$ for $\mu_a/q\mu_s > -2$ and $L_1(\mu_a) \approx q\mu_s + (1/3)(\mu_a + 2q\mu_s)$ for $\mu_a/q\mu_s < -2$.
24. Let $\xi_n = q\mu_s \min_{\mu_a}[L_n(\mu_a)]$ for $n = 2, 3, \ldots$. Then $\xi_2 \approx 4.6$ and, for large n, $\xi_n \to (n - 1/2)\pi$. If $q\mu_s L < \xi_2$, there is one and only one real-valued root $\mu_{a1}(L)$ because the function $L_1(\mu_a)$ does not have a minimum. For this reason, the constant ξ_1 is not defined.
25. D. J. Durian and J. Rudnick, "Photon migration at short times and distances and in cases of strong absorption," J. Opt. Soc. Am. A **14**, 235–245 (1997).
26. S. Andersson-Engels, R. Berg, S. Svanberg, and O. Jarlman, "Time-resolved transillumination for medical diagnostics," Opt. Lett. **15**, 1179–1181 (1990).