


Eigenmodes of a lamellar optical grating: Profile, propagation, reflection, transmission, and nonadiabatic mode coupling

V. V. Kocharovskiy ^{1,2}, C. B. Reynolds,¹ and V. V. Kocharovskiy^{2,3}

¹*Department of Physics and Astronomy and Institute for Quantum Studies and Engineering, Texas A&M University, College Station, Texas 77843-4242, USA*

²*Institute of Applied Physics, Russian Academy of Science, Nizhny Novgorod 603950, Russia*

³*Lobachevsky State University, Gagarin Avenue 23, Nizhny Novgorod 603950, Russia*



(Received 31 December 2018; published 25 November 2019)

We present the explicit analytic formulas for the electromagnetic (optical) eigenmodes propagating in a lamellar grating. Then, we analytically calculate a mutual transformation of the wave's spatial harmonics due to diffraction in the course of propagation through the grating as well as the matrices of the eigenmode reflectance and transmittance at the grating boundaries. Finally, we develop a theory of the nonadiabatic transformation of the eigenmodes in the inhomogeneous gratings. It allows one to disclose the effects which go beyond the geometrical optics of eigenmodes. We show that in a transition layer of a thick nonlamellar grating such effects occur due to mode coupling of just a few lower-band eigenmodes. This approach allows one to find the overall optical response of the nonlamellar gratings fully and efficiently. The results are given for both transverse electric and transverse magnetic polarizations, as well as for any absorptive permittivity.

DOI: [10.1103/PhysRevA.100.053854](https://doi.org/10.1103/PhysRevA.100.053854)

I. INTRODUCTION: THE EIGENMODES AND THE GRATINGS

A concept of the normal modes, or the eigenmodes, has been widely used in the theory of wave propagation in electrodynamics and condensed matter physics since the first decades of the 20th century due to the works by Born, Bloch, and many others. This is especially true for the cases of the periodically inhomogeneous media and light or electronic waves in crystals [1–3]. In fact, the gratings constitute one of the most important components of the numerous modern high-tech devices, and diffraction by grating has been studied in the fields of optics (see, e.g., [4–14]), photonic crystals (see, e.g., [15–20]), lasers and optoelectronics (see, e.g., [21–23]), etc., for decades. For the purpose of the gratings' design, various numerical methods and codes have been developed, including the one based on the concept of eigenmodes [4,5,20].

The analytic description of the eigenmodes propagating in a lamellar grating constitutes one of the most basic and important element of the diffraction-grating theory. An excellent recent review on the exact modal method for various lamellar structures can be found in Ref. [4] (chapters 7 and 10), where a general formalism is presented and applied for the numerical algorithms, in particular, the case of a stack of lamellar layers, each of which could have the unit cell made of many rods of different widths and permittivities. That review contains a general scheme of how to determine the eigenmodes associated with a lamellar layer in a very general case and a note that all of the related matrices can be computed analytically in order to eliminate the numerical instabilities. However, the corresponding solutions were not provided.

In the present paper, we do not use that formalism, but derive the transverse electric (TE) and transverse magnetic

(TM) eigenmodes directly from Maxwell's equations. One of the purposes of the present paper is to give the analytic, concise formulas for the propagation, transformation to or from the plane waves, reflection, and transmission of the eigenmodes in a simple lamellar grating in the form that would be the most suitable for the development of a theory of nonadiabatic mode coupling [24] for the gratings with varying parameters. We formulate the geometrical optics of the eigenmodes and clarify a remarkable fact that just a very few lower-band eigenmodes perfectly describe the diffraction in the relatively thick gratings. Finally, we present the theory of linear mode coupling of the co- and/or counterpropagating eigenmodes in gratings with inhomogeneous permittivity and groove profiles and illustrate it by two generic examples of mode transformation in a trapezoidal grating, shown in Fig. 1 (see Sec. VIII).

In this respect, the present paper intends to advance the papers [10–13] that have been pioneering on the subject. Simultaneously, the present paper aims to set a stage for the theory's application to a new, grating-outcoupled surface-emitting laser (GOSEL) design of the semiconductor diode lasers generating mid- or far-IR radiation via the intracavity nonlinear mixing [25–36].

As is illustrated below, the exact analytic solution for eigenmodes is especially advantageous and efficient for practical calculations when the gratings are relatively deep and/or the permittivity contrast is high (for instance, in the case of a metal-dielectric grating), which is exactly when the standard Fourier, plane-wave expansion method encounters considerable numerical difficulties. The analysis of the simple analytic solution could facilitate finding the eigenmodes for the 2D and 3D photonic crystals, a very important problem that still remains unsolved. The point is that the

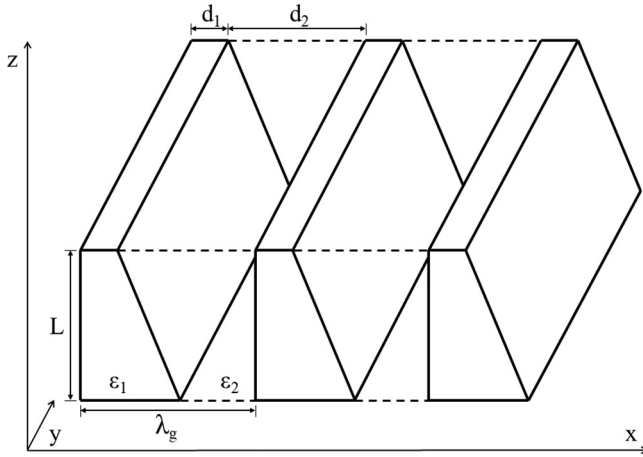


FIG. 1. An inhomogeneous, nonlamellar grating with trapezoidal grooves of a thickness L . Two alternating dielectric layers are made of the dielectric media of constant permittivities ϵ_1, ϵ_2 and have the widths $d_1(z), d_2(z)$ varying along the z axis. A grating period, $\lambda_g = d_1 + d_2$, is a constant.

photonic-crystal structures can be viewed as the gratings with a very large thickness of the order of several tens of wavelengths.

Originally, the most important applications of the diffraction gratings in optics were related to various spectrally selective filters, monochromators, beam splitters, reflectors, polarizers, spectrographs, and other devices for spectroscopy and astronomy. In the last four decades, the range of applications of the grating-like structures has been greatly expanded due to novel applications in the photonic crystal technologies, integrated optics, optical communication systems, and optoelectronics. The examples include engineering effective-medium properties in the periodic metamaterials, control of light propagation and light-matter interaction in photonic nanostructures, engineering photonic-band-gap materials, omnidirectional mirrors which reflect light in a certain frequency range for all directions of incidence and all polarizations [37–39], fiber Bragg gratings [40], 1D lattices of rib waveguides [41–43], chirped mirrors used for dispersion compensation of ultrashort laser pulses inside and outside laser cavities, a wavelength multiplexer based on a superprism (an unusually dispersive prism) [44,45], wavelength-division multiplexing in telecommunications by means of an add-drop filter based on a high-finesse microcavity coupled to a waveguide in a 2D photonic-crystal slab waveguide [46], fabricating the 3D photonic crystals with photonic band gaps at telecommunication wavelengths by means of a woodpile stacking of the gratings made of the III–V direct-gap semiconductors GaAs or InP [47], quantum information processing based on the quantum dots or vacancies embedded into the photonic crystals, etc.

The application of gratings for making new laser sources is particularly interesting. An analysis of surface-emitting semiconductor lasers was started in the 1980s and is given, e.g., in [21,74–81]. Recently, the concept of the GOSELS saw an amazing burst of interest. In particular, it has led to creation of the high-power surface-emitting quantum cascade lasers

(QCLs) [22,23,48–53] and photonic-crystal surface-emitting lasers (PCSELS) capable of a vertical single-mode emission over a large area [54–60] and demonstrating other remarkable properties [61–67]. Note that in the near-IR frequency range, the high-power semiconductor lasers are mostly the vertical-cavity surface-emitting lasers (VCSELs) [68], which do not involve the grating-outcoupled design. In the mid- and far-IR and THz frequency ranges, the QCLs cannot be based on the VCSEL design since their intersubband transitions generate the IR fields of the TM polarization. So, the grating-outcoupled design is used for surface emission [69–73]. The GOSEL design has numerous important advantages over the standard edge-emitting design in power scaling, radiation beam shaping (low divergence) and steering, optical alignment, preventing facet damaging, frequency stability, packaging, assembling the laser arrays, etc.

It is worth noting that sometimes the similar theoretical techniques for solving one and the same diffraction problem have been independently developed by different research communities. For instance, the rigorous coupled wave analysis (or the Fourier modal method) of the diffractive optics community [4–9,14] is more often known as the scattering-matrix approach in the photonic-crystal community [15–19].

Currently, the theoretical analysis and system design in all these fields of the diffraction-grating applications are mostly focused on the development and implementation of various numerical methods [4,5,15–17,20]. This is also true for the GOSEL electrostatics and optics: The *ad hoc* numerical simulations, for example, based on the codes for the layered periodic structures [82], the finite-difference time-domain (FDTD) method [83], or various finite-element modeling softwares like the commercially available software COMSOL [84] dominate the analysis and design of the surface-emitting semiconductor lasers. Many efforts of the photonic-crystal community are devoted to developing the new design-optimization numerical schemes and algorithms. Remarkably, the examples of successfully optimized structures look somewhat disordered and could possibly never have been guessed on the basis of the existing analytic results and intuition [16,85–87]. The latter fact points to a necessity of a further progress in the analytic theory of the photonic-crystal structures.

The present paper is devoted to the analytic theory. In Sec. II, we derive the explicit formulas for the spatial profiles and the spectrum of the TE and TM eigenmodes. A calculation of the diffraction conversion of the wave's spatial harmonics propagating through the grating via an expansion over the eigenmodes is given in Sec. III. In Sec. IV, we calculate the reflectance and transmittance coefficients of the TE and TM eigenmodes at a grating boundary in the form that explicitly generalizes the well-known formulas for the reflection and transmission of a plane wave incident on a boundary between two homogeneous media. In Sec. V, we outline a geometrical optics approximation based on the adiabatic propagation of the eigenmodes. In Secs. VI–IX, we develop a theory of linear mode coupling and transformation in the inhomogeneous gratings. A generalization of the results for the TE and TM eigenmodes to the case of a grating with a complex (absorptive) permittivity is given in Appendix A.

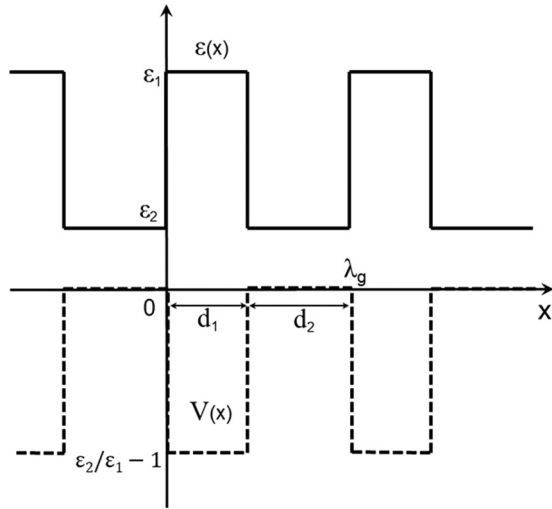


FIG. 2. The profiles of the permittivity $\varepsilon(x)$ and Kronig-Penney potential $V(x) = [\varepsilon_2 - \varepsilon(x)]/\varepsilon_1$ for a lamellar grating made of two alternating dielectric layers of the widths d_1, d_2 and permittivities $\varepsilon_1, \varepsilon_2$. The grating period is $\lambda_g = d_1 + d_2$.

II. EIGENFUNCTIONS AND PROPAGATION CONSTANTS OF THE TRANSVERSE ELECTRIC AND TRANSVERSE MAGNETIC MODES

Let us consider propagation of monochromatic ($e^{-i\omega t}$) electromagnetic waves in a simple lamellar grating made of two alternating layers (with isotropic permittivities) which are homogeneous in the plane perpendicular to the x axis (Fig. 2). The primary question is how the waves are transmitted along the dielectric planes, i.e., along the z axis (see Fig. 1). For simplicity's sake, we will assume that the electromagnetic fields are homogeneous along the third orthogonal y axis. A generalization to the case when the grating has more than two layers per period and/or the fields depend on y as e^{iky} (a conical case of diffraction [13]) is straightforward.

It is sufficient to consider a wave with a definite wave number k_x directed along the x axis and limited to the first Brillouin zone, $k_x \in [-k_g/2, k_g/2]$, along with all of its Bragg harmonics $k_x^{(p)} = k_x + pk_g$ coupled together by diffraction on the grating. Here the grating wave number $k_g = 2\pi/\lambda_g$ is defined by the grating period $\lambda_g = d_1 + d_2$, and $p = 0, \pm 1, \dots$ is an integer. The point is that the waves with different wave numbers k_x are not transformed into each other by the diffraction grating. Thus, an entire field can be always found as a superposition of the k_x waves via a simple integration over k_x .

The wave inside the grating is generated either by a wave incident from outside or by a monochromatic current density $\mathbf{j}(x, z)$ distributed inside the grating. The x dependence of the amplitudes of both sources could be assumed harmonic, i.e., $e^{i(k_x + pk_g)x}$. In the first scenario considered in the pioneering papers [10–13], the wave penetrated inside the grating can be found via the transmittance coefficients (see Sec. IV). In the second scenario, it is enough to consider a thin current sheet with a complex amplitude localized, say, at $z = 0$ along the z direction as a Dirac δ function and represented by one

Fourier harmonic oscillating in the x direction,

$$\mathbf{j}(x, z) = \mathbf{i}_p \delta(z) e^{ik_x^{(p)} x}, \quad k_x^{(p)} = k_x + pk_g, \quad k_x \in \left[-\frac{k_g}{2}, \frac{k_g}{2}\right], \quad (1)$$

where the integer p enumerates various Bragg sidebands.

Contrary to the previous works [10–13], we will follow the second scenario that is relevant, e.g., to the GOSEL lasers with a nonlinear mixing taking place in the bulk of the grating heterostructure. In virtue of a homogeneity of the lamellar grating along the z axis, the dependence of the eigenmode fields on the z coordinate can be factored out. Hence, the contributions of the different δ -current sheets, like in Eq. (1), located at the different positions z' along the z axis can be accounted for by a simple convolution integral over z' with a well-known 1D Green's function

$$G(z, z') = e^{ik_z|z-z'|} / (2ik_z), \quad (2)$$

where k_z is the propagation constant along the z axis.

The electric, \mathbf{E} , and magnetic, \mathbf{H} , fields propagating in the grating can be found from Maxwell's equations

$$\text{curl } \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad \text{curl } \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c} \mathbf{j}, \quad \mathbf{D} = \varepsilon \mathbf{E}, \quad (3)$$

where c is the velocity of light in vacuum. For the sake of a clear physical interpretation, they are written in the Gaussian system of units (assuming that a magnetic permittivity equals unity, $\mu = 1$, so that $\mathbf{B} = \mathbf{H}$) and only for the case of an isotropic permittivity. An analysis that follows can be easily generalized to the case when the permittivity is anisotropic and/or has a spatial dispersion.

For the diffraction grating layout stated above, Maxwell's equations have two independent solutions—with the transverse electric or magnetic polarizations of the propagating fields. For the sake of clarity, each of these solutions is dealt with separately below.

A. Transverse electric eigenmodes

The TE wave is generated by the y component of the current and has only one, y component of the electric field, $E_y(x, z)$. It satisfies the partial differential equation

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} + \frac{\omega^2}{c^2} \varepsilon E_y = -\frac{4\pi i \omega}{c^2} j_y, \quad (4)$$

which follows from Eq. (3). The magnetic field has the x and z components given by the partial derivatives of E_y :

$$H_x = \frac{ic}{\omega} \frac{\partial E_y}{\partial z}, \quad H_z = -\frac{ic}{\omega} \frac{\partial E_y}{\partial x}. \quad (5)$$

In order to find the solution to the Helmholtz equation (4), we first assume that the source current has a form $j_y = \delta(z)f(x)$ with an x profile given by some function $f(x)$ to be determined later on. Then, using the homogeneity of the lamellar grating along the z axis, let us plug in a factorized form of the solution $E_y = E_y(z)f(x)$ with the same x profile $f(x)$ into the Helmholtz equation (4) and, following the

standard method of separation of variables, split it into a system of two equations

$$\hat{L}f(x) = k_z^2 f(x), \quad \hat{L} = \frac{d^2}{dx^2} + \frac{\varepsilon\omega^2}{c^2}, \quad (6)$$

$$\frac{d^2 E_y(z)}{dz^2} + k_z^2 E_y(z) = -\frac{4\pi i\omega}{c^2} \delta(z) \quad (7)$$

with an additional term $k_z^2 E_y(z)f(x)$ being added to the first equation and subtracted from the second equation. Obviously, in order to get a nontrivial solution, one has to equate the newly introduced parameter k_z^2 to one of the eigenvalues of the operator \hat{L} in Eq. (6), $k_z^2 = k_{zn}^2$, and to identify the function $f(x)$ with the corresponding TE eigenmode, $f(x) = f_n(x)$, $n = 1, 2, \dots$. The amplitude of this solution, $E_y(z) = E_{yn}(z)$, is determined by the source current in Eq. (7) via the Green's function (2).

Thus, representing the x profile of the source current j_y as an expansion over a set of the eigenfunctions $f_n(x)$, we can find the solution to the Helmholtz equation (4) in the region free of the source current in Eq. (1) at $z > 0$ as a superposition of the factorized contributions from the TE eigenmodes as follows:

$$E_y(x, z) = \sum_{n=1}^{\infty} E_{yn}(z) f_n(x), \quad E_{yn}(z) = E_{yn}(0) e^{ik_{zn}z}. \quad (8)$$

The amplitude, $E_{yn}(z)$, of each eigenmode is transferred along the z axis independently from the other eigenmodes in accordance with its propagation constant k_{zn} which has a non-negative imaginary part, $\text{Im}k_{zn} \geq 0$, since the grating is made of a nonamplifying medium and the radiation boundary conditions are assumed. The solution in Eq. (8), after a substitution $k_{zn} \rightarrow -k_{zn}$, also describes waves propagating in the opposite, negative direction of the z axis at $z < 0$.

The x profile of the n th eigenmode is a superposition of the counterpropagating sinusoidal or exponential waves,

$$f_n(x) = b_q^+ e^{ik_q(x-d_1)} + b_q^- e^{-ik_q(x-d_1)}, \quad k_q = \left[\frac{\omega^2 \varepsilon_q}{c^2} - k_{zn}^2 \right]^{\frac{1}{2}}, \quad (9)$$

where the index q equals 1 or 2 depending on whether x is inside the grating layer with the permittivity ε_1 or ε_2 , respectively. The amplitudes of these counterpropagating waves are explicit simple functions of the wave numbers

$$\begin{aligned} b_1^{\pm} &= \pm [(-k_2 \pm k_1) e^{ik_2 d_2 - ik_x \lambda_g / 2} \\ &\quad + 2k_2 e^{\pm ik_1 d_1 + ik_x \lambda_g / 2} - (k_2 \pm k_1) e^{-ik_2 d_2 - ik_x \lambda_g / 2}], \\ b_2^{\pm} &= \pm [(k_1 \pm k_2) e^{ik_1 d_1 + ik_x \lambda_g / 2} \\ &\quad - 2k_1 e^{\mp ik_2 d_2 - ik_x \lambda_g / 2} + (k_1 \mp k_2) e^{-ik_1 d_1 + ik_x \lambda_g / 2}]. \end{aligned} \quad (10)$$

The wave amplitudes b_q^{\pm} inside a given layer may be obtained from each other by changing the sign of the wave number $k_q \rightarrow -k_q$ and the amplitude's sign, that is,

$$b_q^- = -b_q^+ [k_q \rightarrow -k_q], \quad q = 1, 2. \quad (11)$$

Equation (10) for the amplitude of the wave in one layer coincides with Eq. (10) for its counterpropagating counterpart in the other layer if one changes the sign of the wave number,

$k_x \rightarrow -k_x$, and interchanges the indexes, $1 \rightarrow 2$ and $2 \rightarrow 1$, in all wave numbers k_1, k_2 and widths d_1, d_2 :

$$b_2^{\pm} = b_1^{\mp} [k_x \rightarrow -k_x \text{ and } 1 \rightarrow 2 \text{ and } 2 \rightarrow 1]. \quad (12)$$

The symmetry relations in Eqs. (11) and (12) immediately follow also from the physical meaning of the amplitudes.

The result in Eq. (10) for the amplitudes is the solution to the system of four linear algebraic equations

$$\begin{aligned} b_1^+ + b_1^- &= b_2^+ + b_2^-, \\ k_1(b_1^+ - b_1^-) &= k_2(b_2^+ - b_2^-), \\ [b_1^+ e^{-ik_1 d_1} + b_1^- e^{ik_1 d_1}] e^{ik_x \lambda_g} &= b_2^+ e^{ik_2 d_2} + b_2^- e^{-ik_2 d_2}, \\ k_1[b_1^+ e^{-ik_1 d_1} - b_1^- e^{ik_1 d_1}] e^{ik_x \lambda_g} &= k_2[b_2^+ e^{ik_2 d_2} - b_2^- e^{-ik_2 d_2}], \end{aligned} \quad (13)$$

which express continuity of the eigenmode x profile and its derivative at the borders between the layers at $x = d_1, d_1 + d_2$ and follow from the well-known property of the tangential components E_y and H_z of electric and magnetic fields to be continuous at the jumps of permittivity.

The nontrivial solution (10) exists only if the determinant of a matrix of Eqs. (13) equals zero. The latter condition constitutes the characteristic equation

$$\begin{aligned} \cos(k_x \lambda_g) &= \cos(k_1 d_1) \cos(k_2 d_2) \\ &\quad - [(k_1^2 + k_2^2) / (2k_1 k_2)] \sin(k_1 d_1) \sin(k_2 d_2) \end{aligned} \quad (14)$$

determining the eigenvalues of the propagation constant k_{zn} , $n = 1, 2, \dots$, for a given x wave number k_x through the dependence of the wave numbers $k_{1,2}$ on k_{zn} , Eq. (9). The eigenmode dispersion law $k_{zn}(k_x)$ has a Bloch-type band-gap structure illustrated in Fig. 3 for the parameters relevant to the grating in a GaAs heterostructure. Usually, only a few (or even none if $n_0 = 0$) lower-band modes $n = 1, \dots, n_0$ propagate along the z axis with the real-valued wave numbers $k_{zn} = \text{Re}(k_{zn})$ while all higher-band eigenmodes are evanescent and exponentially decay due to imaginary wave numbers $k_{zn} = i\text{Im}(k_{zn})$. The higher the band order $n > n_0$ the steeper the decay.

In accordance with the Floquet-Bloch theorem [2], the eigenmode profile in Eq. (9) can be written in a canonical form as a product of an exponential quasiperiodic factor and a periodic function of the variable x as follows:

$$f_n(x, k_x) = e^{ik_x x} \tilde{f}_n(x, k_x), \quad \tilde{f}_n(x + \lambda_g, k_x) = \tilde{f}_n(x, k_x), \quad (15)$$

where ik_x is a Floquet, or characteristic, exponent. Let us introduce an inner product in the space of x profiles,

$$\langle F(x), f(x) \rangle = \frac{1}{2X} \lim_{X \rightarrow \infty} \int_{-X}^X F^*(x) f(x) dx \quad (16)$$

(where a star $*$ means the complex conjugate), for the profiles which are square-integrable on a period λ_g of the grating, that is, have a finite norm defined as follows:

$$\|f(x)\| = \left[\frac{1}{\lambda_g} \int_0^{\lambda_g} |f(x)|^2 dx \right]^{1/2}. \quad (17)$$

Equations (9) and (10) allow us to calculate explicitly the norm squared of the TE-eigenmode x profile as a sum of the

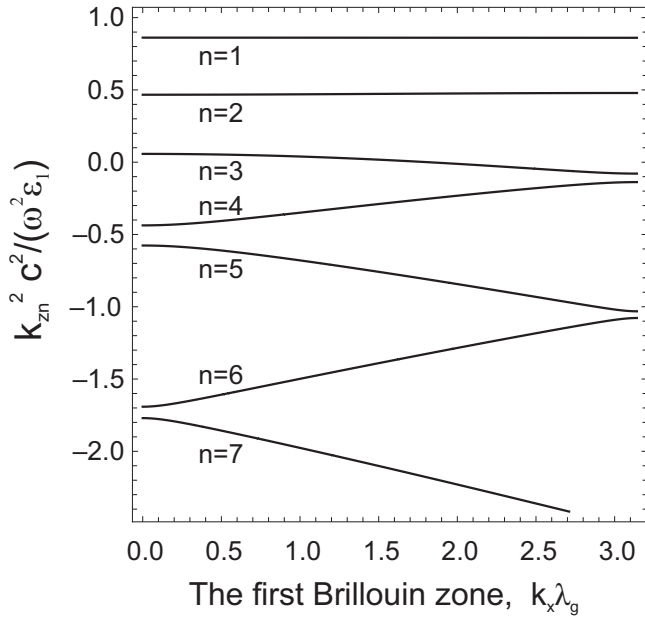


FIG. 3. A typical band-gap structure of the TE-eigenmode dispersion law: The scaled propagation constant squared, $k_{zn}^2 c^2 / (\omega^2 \epsilon_1)$, as a function of the scaled x wave number, $k_x \lambda_g$, within the positive half of the first Brillouin zone. The dielectric layers have equal widths but different permittivities $\epsilon_1 = 12$, $\epsilon_2 = 1$; the grating period is twice the wavelength in the first layer, namely, $d_1 = d_2 = \lambda_g/2 = 2\pi c \sqrt{\epsilon_1} / \omega$.

contributions from the two grating layers per period:

$$\begin{aligned} \|f_n(x, k_x)\|^2 &= \sum_{q=1,2} \frac{d_q}{\lambda_g} \{ |b_q^+|^2 v_1(-\text{Im}\zeta_q) + |b_q^-|^2 v_1(\text{Im}\zeta_q) \\ &\quad + 2\text{Re}[b_q^+ b_q^{*-}] v_1(i\text{Re}\zeta_q) \}; \\ \zeta_q &= (-1)^q 2d_q k_q, \end{aligned} \quad (18)$$

where the elementary function $v_1(u) = (e^u - 1)/u$ behaves analytically and tends to unity, $v_1(0) = 1$, when $u \rightarrow 0$.

In the case of the real-valued permittivities, it is easy to prove that any two eigenmode x profiles corresponding to the different wave numbers, $k'_x \neq k_x$, or bands, $n' \neq n$, are orthogonal. Moreover, for a given wave number $k_x \in [-k_g/2, k_g/2]$ within the first Brillouin zone, the discrete set of the TE-eigenmode x profiles scaled by the norm (17) and (18) and enumerated by the band index n is a complete orthonormal basis in this function space:

$$\frac{\langle f_n, f_{n'} \rangle}{\|f_n(x, k_x)\|^2} = \delta_{n',n} \quad \forall f_n, f_{n'} \in \{f_n(x, k_x) | n = 1, 2, \dots\}. \quad (19)$$

In a general case of the absorptive (complex-valued) permittivities, it is necessary to use the biorthonormal basis described in Appendix A.

B. Transverse magnetic eigenmodes

The TM wave is generated by the x and z components of current and has only one, y component of the magnetic field,

$H_y(x, z)$. It satisfies the partial differential equation

$$\frac{\partial}{\partial x} \left[\frac{\partial H_y}{\epsilon \partial x} \right] + \frac{\partial}{\partial z} \left[\frac{\partial H_y}{\epsilon \partial z} \right] + \frac{\omega^2}{c^2} H_y = \frac{4\pi}{c} \left[\frac{\partial}{\partial x} \left[\frac{j_z}{\epsilon} \right] - \frac{\partial}{\partial z} \left[\frac{j_x}{\epsilon} \right] \right], \quad (20)$$

which follows from Eq. (3). The electric field has the x and z components given by the partial derivatives of H_y :

$$E_x = -\frac{ic}{\epsilon\omega} \frac{\partial H_y}{\partial z}, \quad E_z = \frac{ic}{\epsilon\omega} \frac{\partial H_y}{\partial x}. \quad (21)$$

The analysis of the TM eigenmodes is very similar to that presented above for the TE eigenmodes. The only main difference is the presence of an additional factor $1/\epsilon$ in Eq. (21) and under the external partial derivative $\partial/\partial x$ in the Helmholtz equation (20); cf. Eqs. (4) and (5). It can be taken care of via introducing a new variable

$$\begin{aligned} \bar{x}(x) &= \int_0^x \epsilon(x) dx, \quad \text{i.e., } \bar{x}(x) = \epsilon_1 x \text{ if } 0 \leq x \leq d_1, \\ \bar{x}(x) &= \epsilon_1 d_1 + \epsilon_2 (x - d_1) \quad \text{if } d_1 \leq x \leq d_2. \end{aligned} \quad (22)$$

Below, we present the TM counterparts of Eqs. (6)–(19).

Again, in order to find the solution to the Helmholtz equation (20), we first assume that the source on the right-hand side has the form $(-4\pi i\omega/c^2)\delta(z)\varphi(x)/\epsilon$ with an x profile defined by some function $\varphi(x)$ to be determined later on. Then, using the grating's homogeneity along the z axis and the method of separation of variables, we plug in a factorized form of the solution $H_y = H_y(z)\varphi(x)$ with the same x profile $\varphi(x)$ into the Helmholtz equation (20) and split it into a system of two equations

$$\begin{aligned} \hat{L}_M \varphi(x) &= k_z^2 \varphi(x), \quad \hat{L}_M = \epsilon \frac{d}{dx} \left(\frac{d}{\epsilon dx} \right) + \frac{\epsilon\omega^2}{c^2}, \quad (23) \\ \frac{d^2 H_y(z)}{dz^2} + k_z^2 H_y(z) &= -\frac{4\pi i\omega}{c^2} \delta(z), \quad (24) \end{aligned}$$

with an additional term $k_z^2 H_y(z)\varphi(x)$ being added to the first equation and subtracted from the second equation. Obviously, in order to get a nontrivial solution, one has to equate the newly introduced parameter k_z^2 to one of the eigenvalues of the operator \hat{L}_M in Eq. (23), $k_z^2 = k_{zn}^2$, and to identify the function $\varphi(x)$ with the corresponding TM eigenmode $\varphi(x) = \varphi_n(x)$, $n = 1, 2, \dots$. The amplitude of this solution, $H_y(z) = H_{yn}(z)$, is determined by the source in Eq. (24) via the Green's function (2).

An expansion of the x profile of the source in Eq. (20) over the set of the eigenfunctions $\varphi_n(x)$ yields the solution to the Helmholtz equation (20) in the region free of the source current in Eq. (1) at $z > 0$ via a superposition of the contributions from TM eigenmodes:

$$H_y(x, z) = \sum_{n=1}^{\infty} H_{yn}(z) \varphi_n(x), \quad H_{yn}(z) = H_{yn}(0) e^{ik_{zn}z}. \quad (25)$$

Here the TM-eigenmode propagation constant k_{zn} has a non-negative imaginary part, $\text{Im}(k_{zn}) \geq 0$, since the grating is made of a nonamplifying medium and radiation boundary conditions are assumed. The solution in Eq. (25), after a substitution $k_{zn} \rightarrow -k_{zn}$, also describes TM waves propagating in the negative direction at $z < 0$.

The x profile of the n th eigenmode is a superposition of the counterpropagating waves

$$\varphi_n(x) = b_q^+ e^{ik_q(x-d_1)} + b_q^- e^{-ik_q(x-d_1)}, \quad k_q = \left[\frac{\omega^2 \varepsilon_q}{c^2} - k_{zn}^2 \right]^{\frac{1}{2}}, \quad (26)$$

where again the index q equals 1 or 2 depending on whether x is inside the grating layer with the permittivity ε_1 or ε_2 , respectively. The explicit formulas for the amplitudes of these counterpropagating waves in the TM eigenmodes can be written in the form that also includes the similar formulas for TE eigenmodes in Eq. (10):

$$\begin{aligned} b_1^\pm &= \pm \left[\left(-\frac{k_2}{\varepsilon_2^s} \pm \frac{k_1}{\varepsilon_1^s} \right) e^{ik_2 d_2 - ik_x \lambda_g / 2} \right. \\ &\quad \left. + \frac{2k_2}{\varepsilon_2^s} e^{\pm ik_1 d_1 + ik_x \lambda_g / 2} - \left(\frac{k_2}{\varepsilon_2^s} \pm \frac{k_1}{\varepsilon_1^s} \right) e^{-ik_2 d_2 - ik_x \lambda_g / 2} \right], \\ b_2^\pm &= \pm \left[\left(\frac{k_1}{\varepsilon_1^s} \pm \frac{k_2}{\varepsilon_2^s} \right) e^{ik_1 d_1 + ik_x \lambda_g / 2} \right. \\ &\quad \left. - \frac{2k_1}{\varepsilon_1^s} e^{\mp ik_2 d_2 - ik_x \lambda_g / 2} + \left(\frac{k_1}{\varepsilon_1^s} \mp \frac{k_2}{\varepsilon_2^s} \right) e^{-ik_1 d_1 + ik_x \lambda_g / 2} \right]. \end{aligned} \quad (27)$$

Namely, one should set $s = 1$ for the TM-eigenmode amplitudes, while setting $s = 0$ reduces these formulas to Eq. (10) for the TE-eigenmode amplitudes b_q^\pm . The symmetry relations analogous to those in Eqs. (11) and (12) are valid for the TM-eigenmode amplitudes as well:

$$b_q^- = -b_q^+ [k_q \rightarrow -k_q], \quad q = 1, 2; \quad (28)$$

$$b_2^\pm = b_1^\mp [k_x \rightarrow -k_x \text{ and } 1 \rightarrow 2 \text{ and } 2 \rightarrow 1]. \quad (29)$$

In the latter equation, the permutation of the indexes, $1 \rightarrow 2$ and $2 \rightarrow 1$, refers to all wave numbers k_1, k_2 , widths d_1, d_2 , and explicitly present permittivities $\varepsilon_1, \varepsilon_2$.

The result in Eq. (27) for the TM amplitudes is the solution to the system of four linear algebraic equations

$$\begin{aligned} b_1^+ + b_1^- &= b_2^+ + b_2^-, \\ \frac{k_1}{\varepsilon_1} b_1^+ - \frac{k_1}{\varepsilon_1} b_1^- &= \frac{k_2}{\varepsilon_2} b_2^+ - \frac{k_2}{\varepsilon_2} b_2^-, \\ [b_1^+ e^{-ik_1 d_1} + b_1^- e^{ik_1 d_1}] e^{ik_x \lambda_g} &= b_2^+ e^{ik_2 d_2} + b_2^- e^{-ik_2 d_2}, \\ \frac{k_1}{\varepsilon_1} [b_1^+ e^{-ik_1 d_1} - b_1^- e^{ik_1 d_1}] e^{ik_x \lambda_g} &= \frac{k_2}{\varepsilon_2} [b_2^+ e^{ik_2 d_2} - b_2^- e^{-ik_2 d_2}], \end{aligned} \quad (30)$$

which express the continuity of the eigenmode x profile, $\varphi_n(x)$, and its derivative divided by the permittivity, $\varepsilon^{-1} \partial \varphi_n / \partial x$, at the borders between the layers at $x = d_1, d_1 + d_2$ and follow from the well-known property of the tangential components H_y and E_z of magnetic and electric fields to be continuous at the jumps of permittivity.

The nontrivial solution (27) exists only if the determinant of a matrix of Eqs. (30) equals zero. Here, again, we write down this condition in the form unified for the cases of the TM and TE polarizations (similarly to [15]) by introducing the index s that should be set $s = 1$ or $s = 0$ in order to

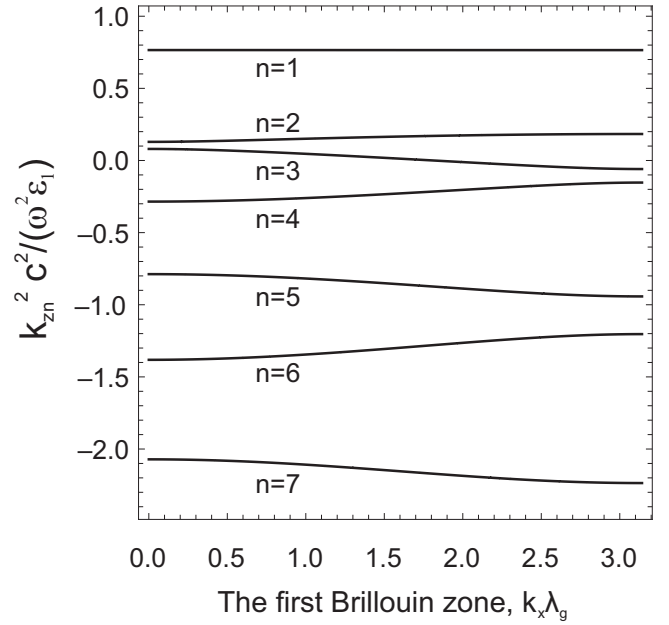


FIG. 4. A typical band-gap structure of the TM-eigenmode dispersion law: The scaled propagation constant squared, $k_{zn}^2 c^2 / (\omega^2 \varepsilon_1)$, as a function of the scaled x wave number, $k_x \lambda_g$, within the positive half of the first Brillouin zone. The parameters of the grating are the same as in Fig. 3.

get the dispersion equation for the TM- or TE-eigenmode propagation constants k_{zn} , respectively:

$$\begin{aligned} \cos(k_x \lambda_g) &= \cos(k_1 d_1) \cos(k_2 d_2) \\ &\quad - \frac{(\varepsilon_2 / \varepsilon_1)^s k_1^2 + (\varepsilon_1 / \varepsilon_2)^s k_2^2}{2k_1 k_2} \sin(k_1 d_1) \sin(k_2 d_2). \end{aligned} \quad (31)$$

The characteristic equation (31) determines the series of eigenmode propagation constants k_{zn} , $n = 1, 2, \dots$, for a given x wave number k_x via the dependence of the wave numbers $k_{1,2}$ on k_{zn} in Eq. (26). The Bloch-type structure of the TM-eigenmode bands and gaps $k_{zn}(k_x)$ is very similar, but quantitatively different from the TE Bloch-type structure as is clearly seen from comparing the TE spectrum in Fig. 3 with the TM spectrum in Fig. 4 plotted for the same grating parameters. In the case of the TM polarization, again only a few (or even none if $n_0 = 0$) lower-band eigenmodes $n = 1, \dots, n_0$ propagate along the z axis with the real-valued wave numbers $k_{zn} = \text{Re}(k_{zn})$ while all higher-band eigenmodes are evanescent and exponentially decay due to their purely imaginary wave numbers $k_{zn} = i \text{Im}(k_{zn})$. The higher the band order $n > n_0$, the steeper the eigenmode decay.

In accordance with the Floquet-Bloch theorem [2], the eigenmode profile in Eq. (26) can be written in a canonical form as a product of an exponential quasiperiodic factor [that, contrary to Eq. (15) for the TE eigenmodes, includes the new variable $\bar{x}(x)$ in Eq. (22), not the original x_0] and a periodic function of the variable x as follows:

$$\varphi_n(x, k_x) = e^{i \bar{k}_x \bar{x}(x)} \tilde{\varphi}_n(x, k_x), \quad (32)$$

$$\tilde{\varphi}_n(x + \lambda_g, k_x) = \tilde{\varphi}_n(x, k_x), \quad \bar{k}_x (\varepsilon_1 d_1 + \varepsilon_2 d_2) = k_x \lambda_g,$$

where $i \bar{k}_x$ is called a Floquet, or characteristic, exponent.

It is convenient to introduce a modified inner product in the space of TM-wave x profiles (see chapter 10 in [4], and [13])

$$\langle \phi(x), \varphi(x) \rangle_M = \frac{1}{2X} \lim_{X \rightarrow \infty} \int_{-X}^X \frac{\phi^*(x)\varphi(x)}{\varepsilon(x)} dx \quad (33)$$

for all profiles which have a finite norm squared

$$\|\varphi(x)\|_M^2 = \frac{1}{\lambda_g} \int_0^{\lambda_g} \frac{|\varphi(x)|^2}{\varepsilon(x)} dx, \quad (34)$$

i.e., are square-integrable on a period λ_g of the grating. Equations (26) and (27) allow us to calculate explicitly the norm squared of the TM-eigenmode x profile as a sum of the contributions from the two grating layers per period:

$$\begin{aligned} \|\varphi_n(x, k_x)\|_M^2 &= \sum_{q=1,2} \frac{d_q}{\lambda_g} \{ |b_q^+|^2 v_1(-\text{Im}\zeta_q) + |b_q^-|^2 v_1(\text{Im}\zeta_q) \\ &\quad + 2\text{Re}[b_q^+ b_q^{-*} v_1(i\text{Re}\zeta_q)] \} / |\varepsilon_q|, \\ \zeta_q &= (-1)^q 2d_q k_q, \end{aligned} \quad (35)$$

where the elementary function $v_1(u) = (e^u - 1)/u$ behaves analytically and tends to unity, $v_1(0) = 1$, when $u \rightarrow 0$.

In the case of the real-valued permittivities $\varepsilon_q > 0$, it is easy to prove that any two eigenmode x profiles corresponding to different wave numbers, $k'_x \neq k_x$, or bands, $n' \neq n$, are orthogonal. Moreover, for a given wave number $k_x \in [-k_g/2, k_g/2]$ within the first Brillouin zone, the operator \hat{L}_M in Eq. (23) is self-adjoint (Hermitian) with respect to the modified inner product (33). The latter is easy to prove, similarly to the self-adjointness of the 1D operator of a particle kinetic energy in quantum mechanics, i.e., $-d^2/dx^2$, by means of rewriting the modified inner product (33) in terms of the new variable $\bar{x}(x)$ in Eq. (22) and employing an integration by parts:

$$\begin{aligned} \left\langle \phi, \varepsilon \frac{d}{dx} \left[\frac{d}{\varepsilon dx} \right] \varphi \right\rangle_M &\equiv \frac{1}{\lambda_g} \int_0^{\bar{x}(\lambda_g)} \phi^* \frac{d^2 \varphi}{d\bar{x}^2} d\bar{x} \\ &= \frac{1}{\lambda_g} \int_0^{\bar{x}(\lambda_g)} \frac{d^2 \phi^*}{d\bar{x}^2} \varphi d\bar{x} \\ &\equiv \left\langle \varepsilon \frac{d}{dx} \left[\frac{d}{\varepsilon dx} \right] \phi, \varphi \right\rangle_M. \end{aligned} \quad (36)$$

Hence, the discrete set of TM-eigenmode x profiles scaled by the norm (34) and (35) and enumerated by the band index n is a complete orthonormal basis in this function space:

$$\frac{\langle \varphi_n, \varphi_{n'} \rangle_M}{\|\varphi_n\|_M \|\varphi_{n'}\|_M} = \delta_{n',n} \quad \forall \varphi_n, \varphi_{n'} \in \{\varphi_n(x, k_x) | n = 1, 2, \dots\}. \quad (37)$$

The latter fact allows us to use the set of TM eigenmodes as the orthonormal basis for an expansion of the TM waves and the source current density as the functions of the coordinate x in the grating instead of or in addition to the standard Fourier expansion over the plane waves. It is the main reason for introducing the modified (by an extra factor $1/\varepsilon$ under the integral) inner product (33).

Note that it is inconsistent to interpret the modified inner product of functions φ and ϕ in Eq. (33) as the standard inner product (16) of functions $\phi/\sqrt{\varepsilon^*}$ and $\varphi/\sqrt{\varepsilon}$,

$$\langle \phi(x), \varphi(x) \rangle_M = \left\langle \frac{\phi(x)}{\sqrt{\varepsilon^*(x)}}, \frac{\varphi(x)}{\sqrt{\varepsilon(x)}} \right\rangle. \quad (38)$$

In particular, a function set $\{\varphi_n(x, k_x)/\sqrt{\varepsilon}\}$ is not a basis suitable for the expansion of the tangential magnetic field $H_y(x)$ (which is a continuous function of the coordinate x even at the borders between the layers) similarly to Eq. (25) since, contrary to the continuous TM-eigenmode x profiles $\varphi_n(x, k_x)$, the functions $\varphi_n(x, k_x)/\sqrt{\varepsilon(x)}$ are not continuous at the discontinuities of permittivity at the layers' borders, i.e., at $x = d_1$ and $x = d_1 + d_2$.

Note also that the x profile of TE and TM eigenmodes in Eqs. (9) and (26) is the same as that of the Kronig-Penney solution [88] for the wave function $\psi_n(x)$ of a particle of mass m moving in a 1D periodic potential $\tilde{V}(x)$:

$$\hat{H}\psi_n = \tilde{E}_n\psi_n, \quad \hat{H} = -(\hbar^2/2m)(d^2/dx^2) + \tilde{V}(x), \quad (39)$$

where \hat{H} is the Hamiltonian and \tilde{E}_n is the eigenenergy.

A discussion of an analogy between these solutions can be found in [3]. In the pioneering papers on diffraction gratings [10–13], the eigenmode solution was presented without a reference to the Kronig-Penney one. This analogy agrees with a well-known interpretation of the rays in geometrical optics as the paths of mechanical particles of mass m moving in an effective potential $\tilde{V}(x)$ determined by the permittivity $\varepsilon(x)$ of a dielectric medium. Namely, a region with a larger permittivity corresponds to a region with a lower potential energy, as is shown in Fig. 2 for the case of TE waves. So, the optical rays tend to turn toward and to concentrate into an optically thick region, like the mechanical particles which tend to move toward a minimum of the potential energy.

This analogy can be stated in terms of the dimensionless counterparts of the energy, potential, and coordinate:

$$E_n = \frac{\varepsilon_2}{\varepsilon_1} - \frac{c^2 k_{zn}^2}{\omega^2 \varepsilon_1}, \quad V(x) = \frac{\varepsilon_2 - \varepsilon(x)}{\varepsilon_1}, \quad X = \frac{x\omega\sqrt{\varepsilon_1}}{c}. \quad (40)$$

Then, Eq. (6) for the TE eigenmodes coincides with its dimensionless Kronig-Penney counterpart of Eq. (39),

$$-\frac{d^2 f_n}{dX^2} + [V(x) - E_n]f_n(X) = 0. \quad (41)$$

Equation (23) for the TM eigenmodes has a similar form,

$$-\frac{d^2 \varphi_n}{d\bar{X}^2} + V_M \varphi_n = 0, \quad V_M = \frac{V(x) - E_n}{(\varepsilon(x))^2}, \quad \bar{X} = \frac{\bar{x}\omega\sqrt{\varepsilon_1}}{c}, \quad (42)$$

to the Schrödinger equation for the eigen-wave-function of a particle moving in a modified 1D potential V_M along the scaled axis $\bar{x}(x)$, Eq. (22), and possessing zero energy.

In this interpretation, the latter (zero) energy of the particle is the eigenvalue of the Hamiltonian $\hat{H}_M = -d^2/d\bar{X}^2 + V_M$ that equals zero only at the particular discrete values of the parameter E_n defined in Eq. (40) and considered as a parameter of the modified potential V_M . In other words, instead of searching for the eigenvalues of the Hamiltonian, we are searching for the discrete set of the modified potential profiles $V_M(x)$ (determined by the discrete values of the parameter E_n) such that they have the discrete energy level with exactly zero energy. The same interpretation is valid for the usual Kronig-Penney eigenvalue problem in Eq. (6) for the TE eigenmodes if one considers $V_E = V(x) - E_n$ as a potential and looks for

the discrete values of the potential's parameter E_n that ensure an appearance of the zero-energy level among the eigenvalues of the Hamiltonian $\hat{H}_E = -d^2/dX^2 + V_E$.

Thus, we come to the *problem of eigenpotentials* with the zero-energy level for the given Hamiltonian, which is \hat{H}_E or \hat{H}_M for the TE or TM eigenmodes, respectively.

This can be viewed as a generalization of the standard eigenvalue problem. In the lossless case of the real-valued permittivities $\varepsilon_q > 0$, the eigenpotential problem is well defined since in this case both Hamiltonians \hat{H}_E and \hat{H}_M are self-adjoint, or Hermitian, operators and the real-valued eigenparameters E_n provide the solution to the problem. In the case of the absorptive permittivities, the complex eigenparameters E_n are needed (Appendix A).

III. DIFFRACTION OF FOURIER HARMONICS PROPAGATING THROUGH THE GRATING

The TE and TM eigenmodes constitute the natural basis for the solution to the problem of the wave propagation in a lamellar grating since they propagate without any diffraction, that is, without changing their x profiles. Let us use this remarkable fact to calculate the diffraction of Fourier harmonics of the TE- or TM-polarized field,

$$\begin{aligned} E_y(x, z) &= \sum_{p=-\infty}^{\infty} g_p(z) e^{i(k_x + pk_g)x}, \\ H_y(x, z) &= \sum_{p=-\infty}^{\infty} g_p(z) e^{i(k_x + pk_g)x}, \end{aligned} \quad (43)$$

due to its propagation through the grating in the $+z$ direction. The amplitudes g_p of the Fourier harmonics in the plane $z = 0$ could be set either by the current sheet embedded into the grating, such as in Eq. (1), or by a wave incident from an outside region $z < 0$ (including the case when $z = 0$ is the grating boundary as discussed in the next section, Sec. IV). We must now determine the *coefficients of diffraction conversion* $t_p^{p'}$ (z) from the p' harmonic of the field, say, at the boundary layer $z = 0$, into any p harmonic at a deeper layer $z > 0$, that is, the transfer matrix $P(z) = \{t_p^{p'}\}$, which gives the column vector of amplitudes $\mathbf{g}(z) = \{g_p(z)\}$ at $z > 0$ via the boundary value of this vector $\mathbf{g}(z = 0) = \{g_p(0)\}$:

$$g_p(z) = \sum_{p'=-\infty}^{\infty} t_p^{p'}(z) g_{p'}(0), \quad \text{or} \quad \mathbf{g}(z) = P(z) \mathbf{g}(0). \quad (44)$$

The problem can be easily solved in three steps. First, we transform the x -profile expansion of the field in Eq. (43) at $z = 0$ from the Fourier basis to the biorthogonal basis of TE or TM eigenmodes as per Eqs. (A7) and (A9),

$$\begin{aligned} e^{i(k_x + p'k_g)x} &= \sum_{n=1}^{\infty} c_n^{p'} f_n(x) \quad \text{or} \quad \sum_{n=1}^{\infty} c_n^{p'} \varphi_n(x), \\ c_n^{p'} &= \frac{\langle f_n^\dagger, e^{i(k_x + p'k_g)x} \rangle}{\langle f_n^\dagger, f_n \rangle} \quad (\text{TE wave}), \\ c_n^{p'} &= \frac{\langle \varphi_n^\dagger, e^{i(k_x + p'k_g)x} \rangle_M}{\langle \varphi_n^\dagger, \varphi_n \rangle_M} \quad (\text{TM wave}). \end{aligned} \quad (45)$$

Next, we find the field at any layer $z > 0$ located deeper inside the grating as a simple superposition of the nondiffracting eigenmodes which propagate in accord with the pure exponential factors $e^{ik_{zn}z}$, as in Eq. (8) or (25):

$$\begin{aligned} E_y(x, z) &= \sum_{p'=-\infty}^{\infty} g_{p'}(z = 0) \sum_{n=1}^{\infty} c_n^{p'} e^{ik_{zn}z} f_n(x), \\ H_y(x, z) &= \sum_{p'=-\infty}^{\infty} g_{p'}(z = 0) \sum_{n=1}^{\infty} c_n^{p'} e^{ik_{zn}z} \varphi_n(x). \end{aligned} \quad (46)$$

Finally, we return back to the expansion via Fourier basis:

$$\begin{aligned} f_n(x) \quad \text{or} \quad \varphi_n(x) &= \sum_{p=-\infty}^{\infty} \tilde{c}_p^n e^{i(k_x + pk_g)x}, \\ \tilde{c}_p^n &= \langle e^{i(k_x + pk_g)x}, f_n \rangle \quad (\text{TE wave}), \\ \tilde{c}_p^n &= \langle e^{i(k_x + pk_g)x}, \varphi_n \rangle \quad (\text{TM wave}). \end{aligned} \quad (47)$$

Note that Eq. (47) involves the standard, not modified (33), inner product (16) even for the case of the TM wave.

Thus, the harmonic amplitudes (44) of the propagating field (43) are given by the following transfer matrix of the diffraction-conversion coefficients:

$$\begin{aligned} t_p^{p'}(z) &= \sum_{n=1}^{n^*} \tilde{c}_p^n e^{ik_{zn}z} c_n^{p'}, \\ P(z) &= C^{-1} \Lambda(z) C, \quad \Lambda = \text{diag}\{e^{ik_{zn}z}\}. \end{aligned} \quad (48)$$

Here we use the fact that the matrix $\tilde{C} = \{\tilde{c}_p^n\}$ is the inverse of the matrix $C = \{c_n^{p'}\}$, $\tilde{C} = C^{-1}$, due to completeness and biorthogonality of the eigenmode basis, Eq. (A7). The diagonal matrix Λ has the entries $e^{ik_{zn}z} \delta_{n',n}$.

We find the explicit formulas for the entries of the basis-transformation matrices C and \tilde{C} as follows:

$$\begin{aligned} c_n^{p'} &= \sum_{q=1}^2 \frac{ie^{\frac{i}{2}k_x \lambda_g}}{\lambda_g N_s \varepsilon_q^s} \sum_{j=+,-} \left[\frac{b_q^{j\dagger} (e^{-ik_q^s d_q} - e^{-ik_x^{(p)} d_q})}{(k_x^{(p)} - jk_q^{\dagger}) e^{i(1-q)(k_q^{\dagger} d_q - k_x^{(p)} d_1)} \right]^*, \\ \tilde{c}_p^n &= \frac{ie^{-\frac{i}{2}k_x \lambda_g}}{\lambda_g} \sum_{q=1}^2 \sum_{j=+,-} \frac{b_q^j (e^{-ik_q d_q} - e^{-ik_x^{(p)} d_q})}{(jk_q - k_x^{(p)}) e^{i(1-q)(k_q d_q - k_x^{(p)} d_1)}}, \end{aligned} \quad (49)$$

where $k_x^{(p)} = k_x + pk_g$, $N_0 = \langle f_n^\dagger, f_n \rangle$, $N_1 = \langle \varphi_n^\dagger, \varphi_n \rangle_M$, and $s = 0$ or 1 for TE or TM eigenmodes, respectively.

The explicit analytic result in Eqs. (48) and (49) turns out to be amazingly efficient for the calculation of the diffraction conversion in the lamellar grating. The point is that only a small number n^* of the lower-band eigenmodes have a relatively small imaginary part of the propagation constant k_{zn} . The contribution from all other, higher-band eigenmodes vanishes exponentially with increasing propagation distance z . Hence, in order to find the coefficient of diffraction conversion $t_p^{p'}(z)$ for a sufficiently thick grating (that constitutes the case being the most difficult for an analysis), it is enough to keep in the sum (48) only very few lower-band eigenmodes $n \in [1, n^*]$. The opposite case of a very thin grating can be easily calculated by means of the perturbation theory within the usual Fourier-Rayleigh expansion method [4–6].

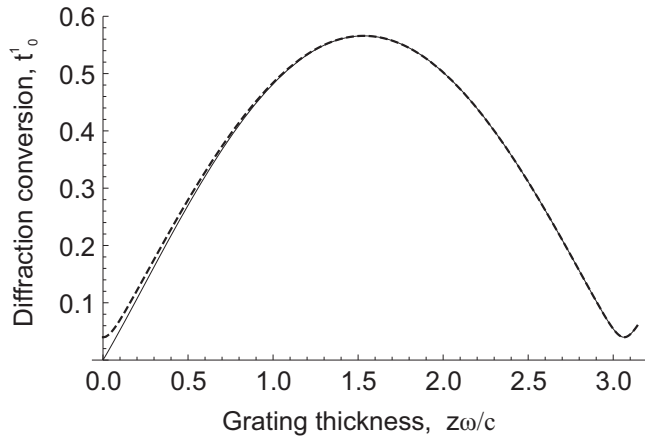


FIG. 5. The TM-wave coefficient of diffraction conversion $t_0^{p'}$ from the $p' = 1$ spatial Fourier harmonic into the $p = 0$ harmonic as a function of the scaled distance $z\omega/c$ that the TM wave propagates inside the grating. The dashed and solid curves are the approximate ($n^* = 3$) and exact ($n^* = \infty$) plots of the explicit analytic Eqs. (48) and (49) for the grating whose eigenmode spectrum is shown in Fig. 4. The Bloch x wave number of the TM wave is assumed to be zero, $k_x = 0$.

As an illustration, we plot the explicit result in Eqs. (48) and (49) for the diffraction conversion coefficient, $t_0^1(z)$, from the $p' = 1$ harmonic into the $p = 0$ harmonic, both with the zero Bloch wave number $k_x = 0$, as a function of the dimensionless grating thickness, $z\omega/c$, for the same grating parameters as the ones adopted in Figs. 3 and 4. It is shown in Fig. 5 for the case of TM waves. The graph of $t_0^1(z)$ for the case of TE waves looks similar. As expected, since only the first three lower-band eigenmodes do not attenuate strongly when penetrating deeper inside this grating, it is enough keeping just $n^* = 3$ eigenmodes in the sum (48) to achieve an accuracy on the order of 1% for all grating thicknesses except very thin ones, $z\omega/c \ll 1$. The other reason for this convenient fact is that the expansion (45) of the incoming $p' = 1$ harmonic over the eigenmodes contains the appreciable (with on the order of unity amplitudes $|c_n^1| \sim 1$) contributions only from the first three lower-band eigenmodes, while the amplitudes of all other eigenmodes are very small, $|c_n^1| \ll 1$. The latter is the reason why the difference between the approximate ($n^* = 3$, the dashed curve) and exact ($n^* = \infty$, the solid curve) values of the diffraction conversion coefficient is much less than unity even for a very thin grating thickness $z\omega/c \ll 1$.

Thus, the amplitude of each Fourier harmonic $g_p(z)$, Eq. (44), inside the grating at the large propagation distances is given by a simple explicit formula (48) as a quasiperiodic interference of just three ($n^* = n_0 = 3$) lower-band eigenmodes. This means that an amplitude modulation of the Fourier harmonic $g_0(z)$, shown in Fig. 5 just within a small distance interval $z\omega/c \in [0, \pi]$, will continue in a quasiperiodic pattern at larger distances. Such a phenomenon of interfering eigenmodes is common for all gratings except the ones which support just one ($n_0 = 1$) or no ($n_0 = 0$) nonevanescing eigenmodes. In both latter cases, the amplitude of each spatial Fourier harmonic does not experience the interference oscillations but tends to a finite or zero constant when the

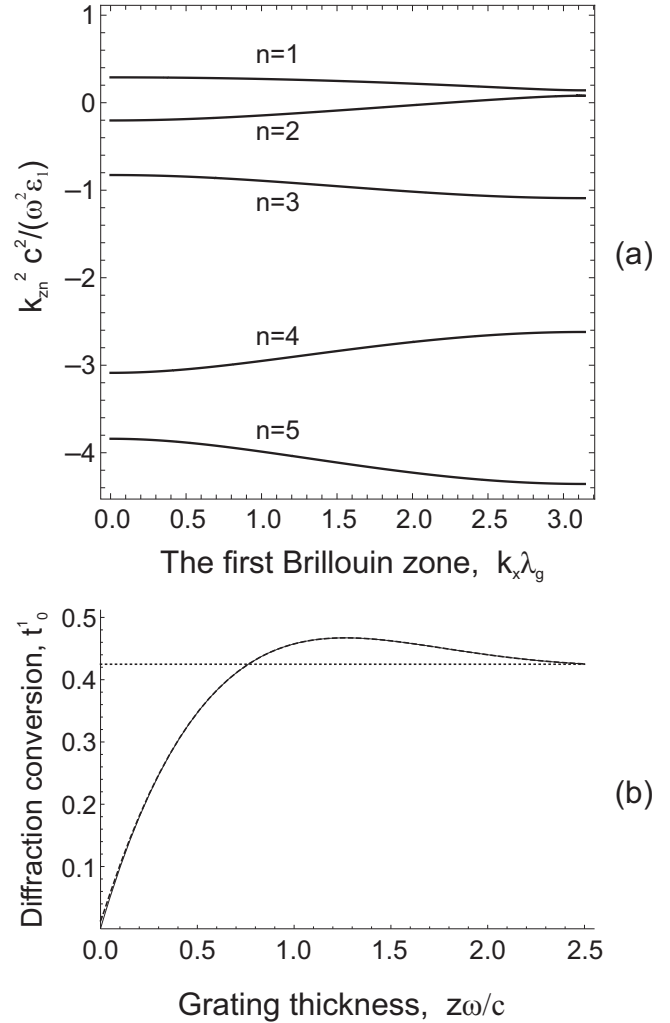


FIG. 6. (a) TM-eigenmode propagation constant squared, $k_{zn}^2 c^2 / (\omega^2 \epsilon_1)$, as a function of the scaled x wave number, $k_x \lambda_g$, within the positive half of the first Brillouin zone. (b) TM-wave coefficient of diffraction conversion $t_0^{p'}$ from $p' = 1$ spatial Fourier harmonic into $p = 0$ harmonic as a function of the scaled distance $z\omega/c$ that the TM wave propagates inside the grating. The dotted, dashed, and solid curves are the approximate ($n^* = 1$ and $n^* = 3$) and exact ($n^* = \infty$) plots of the explicit analytic Eqs. (48) and (49). The grating is the same as in Figs. 4 and 5, except that its period is shorter, $\lambda_g = 2\pi c \sqrt{\epsilon_1} / \omega$. The Bloch wave number is set to zero, $k_x = 0$.

amplitudes of all evanescent eigenmodes exponentially decay with increasing propagation distance.

This propagation pattern is illustrated in Fig. 6 plotted for the grating with the period two times shorter than that of Fig. 5, with all other parameters being kept exactly the same. From the analogy with the Kronig-Penney quantum mechanical problem (see Sec. II) it follows that decreasing the grating period makes the eigenmode propagation constants squared, k_{zn}^2 , more negative since squeezing of the potential wells pulls the energy levels E_n , Eq. (40), out of the wells toward the higher energies. When a dimensionless energy level crosses the critical value ϵ_2/ϵ_1 , so that $E_n > \epsilon_2/\epsilon_1$, the corresponding propagating eigenmode becomes evanescent.

As a result, the grating in Fig. 6 supports only one propagating eigenmode, $n_0 = 1$, and the amplitudes of spatial Fourier harmonics of the field propagating inside the grating tend to the constant values proportional to the weights of the harmonics in the Fourier expansion (47) of this eigenmode. Remarkably, in this example, the approximation in Eqs. (48) and (49) only takes into account three lower-band eigenmodes ($n^* = 3$, the dashed curve), yet is so good that it barely may be distinguished from the exact numerical solution ($n^* = \infty$, the solid curve) in Fig. 6.

A comparison of the diffraction for the cases of the TE and TM waves shows that qualitatively it behaves similarly for both cases; however quantitatively the coefficients of diffraction conversion of the TE and TM waves could be significantly different. For the grating parameters adopted in Fig. 6, they differ by about two times, with the TM-wave diffraction conversion being stronger.

Note that there would be an additional attenuation of the amplitudes of eigenmodes resulting in a decrease of the oscillations and amplitudes of Fourier harmonics in the course of propagation through the grating if a grating material is absorptive; this case is discussed in Appendix A.

IV. REFLECTION AND TRANSMISSION OF EIGENMODES AT A GRATING BOUNDARY

Let us consider the same, unbounded in the xy plane, lamellar grating as the one discussed in the previous sections (see Fig. 2) but occupying now only a finite layer $z \in [0, L]$ of a width L along the vertical axis z and surrounded by the homogeneous substrate ($z < 0$) and cover ($z > L$) media with the permittivities ε_1 and ε_2 , respectively. Suppose the n th TE eigenmode with a unity amplitude, $E_{yn} = f_n(x)e^{ik_{zn}(z-L)}$, propagates in the $+z$ direction inside the grating and is incident on a boundary surface $z = L$ between the grating and the cover medium. It produces a superposition of the TE eigenmodes propagating in the $-z$ direction inside the grating with the amplitudes equal to the coefficients of reflection $R_{n'}^n$,

$$E_y^{(r)} = \sum_{n'=1}^{\infty} R_{n'}^n f_{n'}(x) e^{-ik_{zn'}(z-L)}, \quad (50)$$

as well as a superposition of the plane waves, or the spatial Fourier harmonics, emitted from the grating into the cover medium with the amplitudes equal to the coefficients of transmission T_p^n ,

$$E_y^{(t)} = \sum_{p=-\infty}^{\infty} T_p^n e^{i(k_x + pk_g)x + ik_{zp}^{(2)}(z-L)}, \quad (51)$$

where $k_{zp}^{(2)} = [\varepsilon_2 \omega^2 / c^2 - (k_x + pk_g)^2]^{1/2}$. As is well known, the tangential components of the electric and magnetic fields are continuous at the boundary. That means $E_{yn}(z = L - 0) + E_y^{(r)}(z = L - 0) = E_y^{(t)}(z = L + 0)$ and a similar condition for the corresponding magnetic H_x fields given by Eq. (5). Using the expansion (45) of the Fourier harmonic $e^{i(k_x + pk_g)x}$ and equating coefficients in front of each eigenmode $f_{n'}$, we rewrite these two conditions as the system of two matrix equations

$$CT = R + I, \quad C\hat{k}_z^{(2)}T = \hat{k}_z(I - R) \quad (\text{TE wave}). \quad (52)$$

Here the matrix $C = \{c_n^n\}$ is defined in Eqs. (45) and (49), $I = \{\delta_{n'}^{n'}\}$ is the unity matrix written via the Kronecker delta $\delta_{n'}^{n'}$; $\hat{k}_z = \text{diag}\{k_{zn}\}$ and $\hat{k}_z^{(2)} = \text{diag}\{k_{zp}^{(2)}\}$ are the diagonal matrices composed of the z wave vectors of the eigenmodes in the grating and the plane waves in the cover medium with permittivity ε_2 , respectively; $T = \{T_p^n\}$ and $R = \{R_{n'}^n\}$ are the matrices of the transmittance and reflectance coefficients, respectively. The indices n and n' refer to the eigenmodes $\{f_n\}$ and run over the set of positive integers $\{1, 2, \dots\}$. The indices p and p' refer to the spatial Fourier harmonics $\{e^{i(k_x + pk_g)x}\}$ and run over the set of all integers $\{\dots, -2, -1, 0, 1, 2, \dots\}$. The lower and upper indices enumerate the rows and the columns of a matrix, respectively. The product of two matrices is defined as usual, e.g., $(CT)_{n'}^n = \sum_p c_n^p T_p^n$.

Solving the matrix equations (52), we find the matrices of the transmittance and reflectance coefficients for the case of the *outcoming from the grating TE eigenmode*:

$$T = 2(\hat{k}_z C + C\hat{k}_z^{(2)})^{-1} \hat{k}_z \quad (\text{TE wave}),$$

$$R = (\hat{k}_z + C\hat{k}_z^{(2)}C^{-1})^{-1} (\hat{k}_z - C\hat{k}_z^{(2)}C^{-1}). \quad (53)$$

Note that, as per Eqs. (50) and (51), the transmittance and reflectance coefficients for the TE eigenmode incident onto the cover-grating border from inside the grating are defined via the corresponding amplitudes of the transmitted Fourier harmonics and reflected TE eigenmodes in the tangential component of the electric field E_y .

Remarkably, the derived formula (53) has exactly the same form as that for the transmittance and reflectance coefficients of a plane TE wave incident from a medium with permittivity ε_1 into a medium with permittivity ε_2 ,

$$\frac{E_t}{E_i} = \frac{2k_z^{(1)}}{k_z^{(1)} + k_z^{(2)}}, \quad \frac{E_r}{E_i} = \frac{k_z^{(1)} - k_z^{(2)}}{k_z^{(1)} + k_z^{(2)}}, \quad (54)$$

where E_i, E_t, E_r are the amplitudes of the incident, transmitted, and reflected plane waves and $k_z^{(1)}$ and $k_z^{(2)}$ are their wave vector components, orthogonal to the boundary, in the media ε_1 and ε_2 , respectively [89]. Obviously, the generalization of these formulas to the case of transmission and reflection of eigenmodes in Eq. (53) is achieved via replacing $k_z^{(1)} \rightarrow \hat{k}_z$ and multiplying the z wave vectors by the appropriate matrices of transformation between the eigenmode and Fourier-harmonic bases.

Repeating this analysis for the case of the TM eigenmode, $H_{yn} = \varphi_n(x)e^{ik_{zn}(z-L)}$, incident onto the cover-grating border from inside the grating and defining now the reflectance, $\tilde{R}_{n'}^n$, and transmittance, \tilde{T}_p^n , coefficients via the corresponding amplitudes of the reflected TM eigenmodes and transmitted Fourier harmonics of the tangential component of magnetic field H_y , that is,

$$H_y^{(r)} = \sum_{n'=1}^{\infty} \tilde{R}_{n'}^n \varphi_{n'}(x) e^{-ik_{zn'}(z-L)}, \quad (55)$$

$$H_y^{(t)} = \sum_{p=-\infty}^{\infty} \tilde{T}_p^n e^{i(k_x + pk_g)x + ik_{zp}^{(2)}(z-L)}, \quad (56)$$

we again obtain two matrix equations as the conditions of the continuity of the tangential components of the magnetic and

electric fields at the grating-cover boundary:

$$C\tilde{T} = \tilde{R} + I, \quad MC\hat{k}_z^{(2)}\tilde{T}/\varepsilon_2 = \hat{k}_z(I - \tilde{R}) \quad (\text{TM wave}). \quad (57)$$

Here we take into account a nonorthogonality of the TM-eigenmode basis $\{\varphi_n(x)\}$ with respect to the standard inner product (16), despite its biorthogonality with respect to the modified inner product (33) (see Sec. IV), by means of the matrix $M = \{M_n^{n'}\}$ with the entries

$$M_n^{n'} = \frac{\langle \varphi_n^\dagger, \varphi_{n'} \rangle}{\langle \varphi_n^\dagger, \varphi_n \rangle_M}. \quad (58)$$

The result for the transmission and reflection matrices for the case of the *outcoming from the grating TM eigenmode* is

$$\begin{aligned} \tilde{T} &= 2(\hat{k}_z C + MC\hat{k}_z^{(2)}/\varepsilon_2)^{-1} \hat{k}_z \quad (\text{TM wave}), \\ \tilde{R} &= (\hat{k}_z + MC\hat{k}_z^{(2)}C^{-1}/\varepsilon_2)^{-1} (\hat{k}_z - MC\hat{k}_z^{(2)}C^{-1}/\varepsilon_2). \end{aligned} \quad (59)$$

Note that, as per Eqs. (55) and (56), the transmittance and reflectance coefficients for the TM eigenmode incident onto the cover-grating border from inside the grating are defined via the corresponding amplitudes of the transmitted Fourier harmonics and reflected TM eigenmodes in the tangential component of the magnetic field H_y .

The result (59) is a direct generalization of the classical optics formulas [89] for the transmittance and reflectance coefficients of a plane TM wave incident from a medium with permittivity ε_1 into a medium with permittivity ε_2 ,

$$\frac{H_t}{H_i} = \frac{2k_z^{(1)}/\varepsilon_1}{k_z^{(1)}/\varepsilon_1 + k_z^{(2)}/\varepsilon_2}, \quad \frac{H_r}{H_i} = \frac{k_z^{(1)}/\varepsilon_1 - k_z^{(2)}/\varepsilon_2}{k_z^{(1)}/\varepsilon_1 + k_z^{(2)}/\varepsilon_2}, \quad (60)$$

where H_i, H_t, H_r are the amplitudes of the magnetic field in the incident, transmitted, and reflected plane waves.

The analysis of the TE or TM plane wave (i.e., the p th spatial Fourier harmonic) with a unity amplitude,

$$E_{yp} \text{ or } H_{yp} = e^{i(k_x + pk_g)x + ik_{zp}^{(1)}z}, \quad p = 0, \pm 1, \pm 2, \dots, \quad (61)$$

incident on a bottom border $z = 0$ of the grating from the substrate with the z wave vector $k_{zp}^{(1)} = [\varepsilon_1 \omega^2 / c^2 - (k_x + pk_g)^2]^{1/2}$ is very similar to the above analysis. The reflectance, R_p^p or \tilde{R}_p^p , and transmittance, T_n^p or \tilde{T}_n^p , coefficients are defined by the corresponding amplitudes of the reflected plane waves and transmitted eigenmodes:

$$E_y^{(r)} = \sum_{p'=-\infty}^{\infty} R_{p'}^p e^{i(k_x + p'k_g)x - ik_{zp'}^{(1)}z} \quad (\text{TE wave}), \quad (62)$$

$$H_y^{(r)} = \sum_{p'=-\infty}^{\infty} \tilde{R}_{p'}^p e^{i(k_x + p'k_g)x - ik_{zp'}^{(1)}z} \quad (\text{TM wave}), \quad (63)$$

$$E_y^{(t)} = \sum_{n=1}^{\infty} T_n^p f_n(x) e^{ik_{zn}z} \quad (\text{TE wave}), \quad (64)$$

$$H_y^{(t)} = \sum_{n=1}^{\infty} \tilde{T}_n^p \varphi_n(x) e^{ik_{zn}z} \quad (\text{TM wave}). \quad (65)$$

As a result, we find the transmittance and reflectance coefficients for the case of the *incoming into the grating TE or TM*

plane wave in the explicit matrix form as follows:

$$\begin{aligned} T &= 2[\hat{k}_z^{(1)}C^{-1} + C^{-1}\hat{k}_z]^{-1}\hat{k}_z^{(1)}, \\ R &= [\hat{k}_z^{(1)} + C^{-1}\hat{k}_zC]^{-1}[\hat{k}_z^{(1)} - C^{-1}\hat{k}_zC] \end{aligned} \quad (66)$$

for the TE wave, or

$$\begin{aligned} \tilde{T} &= 2[\hat{k}_z^{(1)}C^{-1} + \varepsilon_1(MC)^{-1}\hat{k}_z]^{-1}\hat{k}_z^{(1)}, \\ \tilde{R} &= [\hat{k}_z^{(1)} + \varepsilon_1(MC)^{-1}\hat{k}_zC]^{-1}[\hat{k}_z^{(1)} - \varepsilon_1(MC)^{-1}\hat{k}_zC] \end{aligned} \quad (67)$$

for the TM wave. Again, they are the direct generalizations of the classical optics formulas [89] for the transmission and reflection of the wave incident from a medium with permittivity ε_1 into a medium with permittivity ε_2 . However, now the wave vector $k_z^{(2)}$, not $k_z^{(1)}$, corresponds to the eigenmode wave vectors \hat{k}_z inside the grating and, hence, should be replaced by \hat{k}_z in Eqs. (54) and (60).

V. GEOMETRICAL OPTICS: PROPAGATION OF THE EIGENMODES IN A WEAKLY INHOMOGENEOUS GRATING

In this and the following sections we consider gratings similar to Fig. 1 which have a lamellar-type structure shown in Fig. 2 for each xy plane at any particular z , but are inhomogeneous along the z axis. This means the parameters of the grating $\varepsilon_{1,2}(z)$ and $d_{1,2}(z)$ are the functions of the coordinate z ; i.e., the permittivities of the grating layers vary along the z axis, for instance, in virtue of a designed variation of the material composition during the grating growth process, and/or the groove profile is not strictly vertical but rather trapezoidal or curved. Nevertheless, we still can use the complete basis of the TE- or TM-eigenmode x profiles $\{f_n(x, z)\}$ or $\{\varphi_n(x, z)\}$ introduced in Sec. II and Appendix A, but now it is different at different levels z . So, we can expand the electromagnetic field of the TE or TM wave inside the grating over the eigenmode x profiles similar to Eqs. (8) or (25):

$$E_y = \sum_{n=1}^{\infty} E_{yn}(z) f_n(x, z) \quad \text{or} \quad H_y = \sum_{n=1}^{\infty} H_{yn}(z) \varphi_n(x, z). \quad (68)$$

However, the amplitudes of the eigenmode x profiles, $E_{yn}(z)$ or $H_{yn}(z)$, now vary in the course of the wave propagation, or diffraction, in a way that is different from a simple exponential function $e^{\pm ik_{zn}z}$. They should be calculated taking into account the z inhomogeneity of the grating. We address this problem in the case of a weakly inhomogeneous grating within the geometrical optics approximation in this section and within the theory of the nonadiabatic mode coupling in the following sections.

Let us assume that the scale Λ of variation of the grating parameters along the z axis is much larger than the wavelength λ_{zn} corresponding to the wave vector k_{zn} :

$$|d\lambda_{zn}/dz| = |\lambda_{zn}/\Lambda| \equiv \alpha \ll 1, \quad \lambda_{zn} = 2\pi/k_{zn}. \quad (69)$$

Then, we can solve Maxwell's Eq. (4) or (20) for the electromagnetic field of the TE or TM wave by means of its rigorous asymptotic expansion in a power series over the small parameter of weak inhomogeneity $\alpha \ll 1$.

The geometrical optics (WKB) approximation assumes that each eigenmode propagates independently from the others according to Eq. (7) or (24). Keeping only the zeroth- and first-order terms ($\sim \alpha^0, \alpha$) in these equations, one finds the solution for amplitudes $E_{yn}(z)$ or $H_{yn}(z)$ as a superposition of counterpropagating eigenmodes:

$$E_{yn} \text{ or } H_{yn} = \frac{\Phi_n(A_n^+ e^{i\phi_n} + A_n^- e^{-i\phi_n})}{k_{zn}^{1/2}}, \quad \phi_n = \int_0^z k_{zn} dz. \quad (70)$$

Here the geometrical optics TE- or TM-amplitude factor $\Phi_n(z)/k_{zn}^{1/2}$, where

$$\Phi_n(z) = \exp\left[-\int_0^z \psi_{nm} dz\right], \quad \psi_{nm}(z) = \frac{\langle f_n^\dagger, \frac{\partial f_m}{\partial z} \rangle}{\langle f_n^\dagger, f_n \rangle} \quad (\text{TE})$$

or

$$\psi_{nm}(z) = \frac{\langle \varphi_n^\dagger, \frac{\partial \varphi_m}{\partial z} - \frac{\varphi_m}{2\varepsilon} \frac{\partial \varepsilon}{\partial z} \rangle_M}{\langle \varphi_n^\dagger, \varphi_n \rangle_M} \quad (\text{TM}), \quad (71)$$

adiabatically follows a variation of wave vector $k_{zn}(z)$ and incorporates (via Φ_n) the effects of restructuring of the eigenmode and permittivity x profiles that go beyond the explicit dependence of the eigenmode functions f_n or φ_n on the z coordinate in Eq. (68). Thus, the amplitudes of the eigenmodes are not just the boundary-value constants A_n^\pm , but rather acquire an additional rescaling in the course of propagation through the inhomogeneous grating along the z axis due to the geometrical optics factor.

Importantly, the condition (69) of a weak inhomogeneity is not enough for a validity of this solution. It is also necessary that an inhomogeneity scale should be larger than the wavelength of beating between the eigenmodes:

$$\Lambda \gg 2\pi/|k_{zn} - k_{zn'}|, \quad n \neq n'. \quad (72)$$

Otherwise, the eigenmodes lose their independence in a region of the inhomogeneity-scale size if their parameters significantly change over this region but the accumulated phase difference between the eigenmodes $|k_{zn} - k_{zn'}|\Lambda$ remains less or on the order of 2π . In such regions those (close to a degeneracy) eigenmodes may experience a significant additional variation of their amplitudes that could be viewed as if the eigenmodes exchange by their amplitudes due to a mutual mode transformation [24]. We discuss this remarkable effect in the next section.

VI. NONADIABATIC MODE COUPLING IN THE NONLAMELLAR GRATINGS

The effect stated in the previous paragraph is known as the nonadiabatic mode coupling [24] and is generic for all nonlamellar gratings which could be viewed as the lamellar gratings inhomogeneous along the z axis. It implies that the amplitudes of the eigenmodes do not follow anymore the adiabatic WKB, or geometrical optics, law when propagate along the z axis but vary differently in virtue of the so-called linear mode coupling that appears due to the inhomogeneity along the propagation path. In particular, the reflection and transmission of the eigenmodes at the sharp grating boundaries considered in Sec. IV are the limiting case of the nonadiabatic

mode coupling involving many, both counterpropagating (reflection) and copropagating (transmission), eigenmodes.

First, we elaborate on the other two simple cases of the nonadiabatic mode coupling in inhomogeneous gratings:

(i) the mutual transformation (reflection) of the two counterpropagating eigenmodes of the same band type n with z wave vectors close to zero, $\pm k_{zn} \approx 0$, as is the case, for example, for the eigenmode $n = 3$ at $k_x \approx 2/\lambda_g$ in Figs. 3 and 4 or $n = 2$ at $k_x \approx \pi/\lambda_g$ in Fig. 6, as well as

(ii) the mutual transformation of the two copropagating eigenmodes of the different but neighboring band types n and $n' = n + 1$ with z wave vectors close to a Bragg band gap at x wave number $k_x \approx 0$ or π/λ_g , as is the case, e.g., at $k_x \approx 0$ for the eigenmodes $n = 4, n' = 5$ in Figs. 3 and 15 (at $d_2 \sim 1$) or $n = 2, n' = 3$ in Fig. 4.

The first case is common to most of the gratings and occurs often in practice. However, it is special since the very condition of a weak inhomogeneity (69) is violated for the two degenerate (being reflected into each other and possessing exactly the same x profiles) eigenmodes which are close (in the parameters' space) to the critical border $k_{zn} = 0$ between propagating and evanescent modes and could be either of them. The second case is simpler since the copropagating eigenmodes can efficiently transform into each other even when the condition of a weak inhomogeneity (69) is fulfilled.

The equation describing the linear coupling (reflection) of the two counterpropagating TE or TM eigenmodes with the same x profile, f_n or φ_n , can be obtained by plugging in their combined field written in the form

$$E_y = F_n(z)\Phi_n(z)f_n(x, z) \text{ or } H_y = F_n(z)\Phi_n(z)\varphi_n(x, z) \quad (73)$$

into the Maxwell equation (4) or (20) and projecting them onto the eigenmode x profile by calculating the corresponding scalar product $\langle f_n^\dagger, \dots \rangle$ or $\langle \varphi_n^\dagger, \dots \rangle_M$. The result is the second-order ordinary differential equation for the field amplitude $F(z)$ as a function of the coordinate z in the form of the canonical 1D Helmholtz equation

$$d^2 F_n / dz^2 + \tilde{k}_{zn}^2 F_n = 0 \quad (74)$$

with the renormalized propagation constant squared

$$\tilde{k}_{zn}^2 = k_{zn}^2 - \psi_{nn}^2 - \frac{d\psi_{nn}}{dz} + \eta_{nn}, \quad \eta_{nn} = \frac{\langle f_n^\dagger, \frac{\partial^2 f_n}{\partial z^2} \rangle}{\langle f_n^\dagger, f_n \rangle} \quad (\text{TE}),$$

$$\text{or } \eta_{nn} = \frac{\langle \varphi_n^\dagger, \varepsilon \frac{\partial}{\partial z} \left(\frac{1}{\varepsilon} \frac{\partial \varphi_n}{\partial z} \right) \rangle_M}{\langle \varphi_n^\dagger, \varphi_n \rangle_M} \quad (\text{TM}). \quad (75)$$

Introducing the amplitudes F_n^+ and F_n^- of the counterpropagating eigenmodes related to their total field F_n and its derivative dF_n/dz as follows,

$$F_n = \frac{F_n^+}{\tilde{k}_{zn}^{1/2}} e^{i\tilde{\phi}_n} + \frac{F_n^-}{\tilde{k}_{zn}^{1/2}} e^{-i\tilde{\phi}_n},$$

$$\frac{dF_n}{dz} = i\tilde{k}_{zn}^{1/2} F_n^+ e^{i\tilde{\phi}_n} - i\tilde{k}_{zn}^{1/2} F_n^- e^{-i\tilde{\phi}_n}, \quad \tilde{\phi}_n = \int_0^z \tilde{k}_{zn} dz, \quad (76)$$

we rewrite the equation of mode transformation for reflection (74) explicitly as the system of two coupled first-order

differential equations for eigenmode amplitudes:

$$\begin{aligned}\frac{dF_n^+}{dz} &= \frac{F_n^-}{2\tilde{k}_{zn}} \left(\frac{d\tilde{k}_{zn}}{dz} \right) e^{-2i\tilde{\phi}_n}, \\ \frac{dF_n^-}{dz} &= \frac{F_n^+}{2\tilde{k}_{zn}} \left(\frac{d\tilde{k}_{zn}}{dz} \right) e^{2i\tilde{\phi}_n}.\end{aligned}\quad (77)$$

The diagonal, self-coupling terms are absent because they have been taken care of by the explicit factor $\Phi_n(z)$ [see Eqs. (71) and (73)]. Equation (77) yields the first-order nonlinear differential equation, the well-known Riccati equation,

$$\frac{dR_n}{dz} + 2i\tilde{k}_{zn}R_n = \frac{1 - R_n^2}{2\tilde{k}_{zn}} \frac{d\tilde{k}_{zn}}{dz}, \quad (78)$$

directly for the coefficient of reflection

$$R_n(z) = \frac{F_n^-(z)}{F_n^+(z)} e^{-2i\tilde{\phi}_n(z)} \quad (79)$$

at any cross section z inside the inhomogeneous grating.

The equation describing the linear coupling of the two copropagating TE or TM eigenmodes with the different x profiles, f_n and f_m or φ_n and φ_m , can be obtained by plugging in their combined field written in the form

$$\begin{aligned}E_y &= \frac{F_n^+ \Phi_n e^{i\phi_n}}{k_{zn}^{1/2}} f_n(x, z) + \frac{F_m^+ \Phi_m e^{i\phi_m}}{k_{zm}^{1/2}} f_m(x, z) \quad (\text{TE}), \\ H_y &= \frac{F_n^+ \Phi_n e^{i\phi_n}}{k_{zn}^{1/2}} \varphi_n(x, z) + \frac{F_m^+ \Phi_m e^{i\phi_m}}{k_{zm}^{1/2}} \varphi_m(x, z) \quad (\text{TM})\end{aligned}\quad (80)$$

into the Maxwell equation (4) or (20) and projecting them onto the eigenmode x profiles by calculating the corresponding scalar products $\langle f_n^\dagger, \dots \rangle$ and $\langle f_m^\dagger, \dots \rangle$ or $\langle \varphi_n^\dagger, \dots \rangle_M$ and $\langle \varphi_m^\dagger, \dots \rangle_M$. For the coupling of the copropagating eigenmodes, contrary to the case of reflection considered above, we assume that the condition of a weak inhomogeneity in Eq. (69) is fulfilled and, hence, keep only the first-order terms $\sim \alpha$, neglecting the second-order derivatives. As a result, we obtain the linear system of the two coupled first-order differential equations for the eigenmode amplitudes F_n^+ and F_m^+ as follows:

$$dF_n^+/dz = g_n^m(z)F_m^+, \quad dF_m^+/dz = g_m^n(z)F_n^+. \quad (81)$$

This is similar to that in Eq. (77) for the two-mode coupling due to reflection. However, the coupling coefficients for the copropagating TE or TM eigenmodes,

$$\begin{aligned}g_n^m(z) &= -\frac{\langle f_n^\dagger, \frac{\partial f_m}{\partial z} \rangle}{\langle f_n^\dagger, f_n \rangle} \frac{\Phi_m k_{zm}^{1/2}}{\Phi_n k_{zn}^{1/2}} e^{i(\phi_m - \phi_n)} \quad (\text{TE wave}), \\ g_n^m(z) &= -\frac{\langle \varphi_n^\dagger, \frac{\partial \varphi_m}{\partial z} \rangle_M}{\langle \varphi_n^\dagger, \varphi_n \rangle_M} \frac{\Phi_m k_{zm}^{1/2}}{\Phi_n k_{zn}^{1/2}} e^{i(\phi_m - \phi_n)} \quad (\text{TM wave}),\end{aligned}\quad (82)$$

originate now from the overlapping of one eigenmode x profile, say, $f_n(x, z)$ or $\varphi_n(x, z)$, in the cross section at coordinate z with the other eigenmode x profile, say, $f_m(x, z + dz)$ or $\varphi_m(x, z + dz)$, in the neighboring cross section at coordinate $z + dz$. Such a slipping of one eigenmode into the other eigenmode in the course of propagation along the z axis

is the ultimate reason for the mode transformation in the inhomogeneous gratings.

The ratio of the field amplitudes of the two eigenmodes obeys the Riccati equation which is similar to Eq. (78) but also includes the coupling coefficients (82). So, the method of the Riccati equation is fully applicable for the analysis, both qualitative and quantitative, of the mode coupling of the two copropagating eigenmodes as well.

The derived Eqs. (74), (77), (78), and (81) are all of the standard types. The methods for their analysis are well developed and known from the theory of wave propagation and waveguides, quantum mechanics, etc. (see, e.g., [2,6,24,90,91] and references therein). The behavior of their solutions is well understood, both mathematically and physically, by means of various analytic and numerical methods, including a qualitative theory of linear mode coupling [24] and a rich set of standard, exactly solvable (in terms of known special functions) models like a transition layer with a linear profile of permittivity solvable via the Airy functions. Besides, there is a large number of well-studied analogous problems on a nonadiabatic mode transformation in various nonstationary processes where the time plays a part of the coordinate z along a propagation path (see, e.g., [92–94]). A famous example is an analytic solution to the Landau-Zener model of a transition between the states of a two-level quantum mechanical system governed by a time-dependent Hamiltonian.

In the case of the inhomogeneous lamellar gratings considered above, these equations can be studied in even greater detail since there are explicit analytic formulas for the profiles and other parameters of the eigenmodes. Obviously, such an analysis can be generalized to the cases when not just two, but a larger number of eigenmodes $n = n_1, n_2, \dots, n_Q$ experience a mutual linear mode coupling or even when a few groups of such coupled eigenmodes exist in the grating. Then, their combined field

$$\begin{aligned}E_y(x, z) &= \sum_{n=n_1, \dots, n_Q} \frac{F_n^+ \Phi_n e^{i\phi_n}}{k_{zn}^{1/2}} f_n(x, z) \quad (\text{TE wave}), \\ H_y(x, z) &= \sum_{n=n_1, \dots, n_Q} \frac{F_n^+ \Phi_n e^{i\phi_n}}{k_{zn}^{1/2}} \varphi_n(x, z) \quad (\text{TM wave})\end{aligned}\quad (83)$$

can be described as a superposition of those Q eigenmode fields whose amplitudes F_n^+ obey the system of coupled linear first-order ordinary differential equations

$$\frac{dF_n^+}{dz} = \sum_{m=n_1, \dots, n_Q} g_n^m(z)F_m^+, \quad n = n_1, \dots, n_Q. \quad (84)$$

VII. SIMULTANEOUS TRANSFORMATION AND REFLECTION OF EIGENMODES IN THE INHOMOGENEOUS LAMELLAR GRATINGS

Finally, we derive the system of the first-order differential equations for the amplitudes F_n^\pm of the co- and counter-propagating eigenmodes that takes into account exactly the mutual transformation and reflection between the eigenmodes within a given subset of eigenmodes $n = n_1, n_2, \dots, n_Q$. In other words, we truncate the functional space of the field x profiles to a subspace spanned by all superpositions of the

selected eigenmode basis x profiles $\{f_n(x, z)|n = n_1, \dots, n_Q\}$ or $\{\varphi_n(x, z)|n = n_1, \dots, n_Q\}$ at any particular cross section z . Thus, the only approximation made in deriving these equations is an omission of the contributions due to diffraction into the eigenmodes which do not belong to the selected subset of eigenmodes.

In order to get these equations, we represent the combined electric or magnetic field of the Q eigenmodes as

$$E_y = \sum_{m=n_1, \dots, n_Q} F_m(z) \Phi_m(z) f_m(x, z) \quad (\text{TE wave})$$

$$\text{or } H_y = \sum_{m=n_1, \dots, n_Q} F_m(z) \Phi_m(z) \varphi_m(x, z) \quad (\text{TM wave}), \quad (85)$$

plug in it into the Maxwell equation (4) or (20), and project the latter onto each eigenmode x profile by calculating a scalar product $\langle f_n^\dagger, \dots \rangle$ or $\langle \varphi_n^\dagger, \dots \rangle_M$ that yields

$$\frac{d^2 F_n}{dz^2} + \tilde{k}_{zn}^2 F_n = \sum_{m \neq n} \frac{\Phi_m}{\Phi_n} \left[2\psi_{nm} \left(\psi_{mm} - \frac{d}{dz} \right) - \eta_{nm} \right] F_m. \quad (86)$$

We then rewrite Eq. (86) in terms of the amplitudes (76) of the counterpropagating eigenmodes F_n^\pm . As a result, we find the system of $2Q$ first-order differential equations describing the linear mode coupling of the eigenmodes:

$$\frac{dF_n^\pm}{dz} = \frac{F_n^\mp}{2\tilde{k}_{zn}} \left(\frac{d\tilde{k}_{zn}}{dz} \right) e^{\mp 2i\tilde{\phi}_n} \pm \sum_{m \neq n} \frac{\Phi_m \tilde{k}_{zm}^{1/2}}{\Phi_n \tilde{k}_{zn}^{1/2}} e^{\mp i\tilde{\phi}_n}$$

$$\times \left\{ F_m^+ e^{i\tilde{\phi}_m} \left[\psi_{nm} \left(-1 - \frac{i\psi_{mm}}{\tilde{k}_{zm}} \right) + \frac{i\eta_{nm}}{2\tilde{k}_{zm}} \right] \right.$$

$$\left. + F_m^- e^{-i\tilde{\phi}_m} \left[\psi_{nm} \left(1 - \frac{i\psi_{mm}}{\tilde{k}_{zm}} \right) + \frac{i\eta_{nm}}{2\tilde{k}_{zm}} \right] \right\},$$

$$n = n_1, \dots, n_Q, \quad (87)$$

where the upper sign should be used in the equation for $\frac{dF_n^+}{dz}$ and the lower sign for $\frac{dF_n^-}{dz}$ (Appendix B). These equations for the cases of TE and TM polarization have exactly the same form, although the parameters ψ_{nm} and η_{nm} defined in Eqs. (71) and (75) are different.

Note that an explicit inclusion of the geometrical optics factor $\Phi_m(z)$ in the ansatz for the wave field (85) and its particular cases in Eqs. (73), (80), and (83) is beneficial for the analysis of the linear mode coupling in a transition layer. This allows one to explicitly factor out the known [in the form of a definite integral of a given function $\psi_{mm}(z)$] geometrical optics scaling of the mode fields stated in Eqs. (70) and (71). As a result, the left-hand side of the second-order differential equation (86) for each mode field acquires a canonical (i.e., without a first-order derivative) form and, moreover, all self-mode-coupling terms disappear from the right-hand side of the system (87) of the first-order differential equations for the mode amplitudes $F_n^\pm(z)$. So, Eqs. (87) acquire a canonical, off-diagonal form and, most importantly, allow one to focus on and describe the net, pure effect of the transformation of the mode amplitudes $F_n^\pm(z)$, that is, their deviation from the constant geometrical optics amplitudes A_m^\pm of Eq. (70), due to

an inhomogeneity of the transition layer of the nonlamellar grating along the propagation path.

A fundamental problem for any mode coupling is finding the transmittance, T_n^m , and reflectance, R_n^m , coefficients of linear mode transformation that are equal to the amplitudes of the copropagating, $F_n^+(z=L)$, and counterpropagating, $F_n^-(z=0)$, eigenmodes, respectively, at the exits from a given transition layer $z \in [0, L]$ of a thickness L generated by an m eigenmode of the unit amplitude, $F_m^+(z=0) = 1$, entering the layer in the $+z$ direction at $z=0$. Let us consider a general case of mode transformation or reflection described by Eqs. (87) for $2Q$ eigenmode amplitudes $F_n^\pm(z)$ assuming that a grating is homogeneous outside the transition layer. Let us introduce a column $2Q$ -vector \mathbf{F} as a 2-block vector—a column Q -vector \mathbf{F}^+ on top of a column Q -vector \mathbf{F}^- , where each Q -vector $\mathbf{F}^\pm = (F_{n_1}^\pm, \dots, F_{n_Q}^\pm)^T$ contains the amplitudes of the co- or counterpropagating eigenmodes, respectively. (A superscript T denotes a transpose operation.) The values of the amplitude vector $\mathbf{F}(z)$ at the opposite borders of the transition layer, $z=0$ and $z=L$, are related by a 2×2 -block state-transition matrix, or matriciant, M of the system of $2Q$ linear differential equations (87),

$$\mathbf{F}(L) = M\mathbf{F}(0), \quad M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad \mathbf{F} = \begin{bmatrix} \mathbf{F}^+ \\ \mathbf{F}^- \end{bmatrix}. \quad (88)$$

The matriciant and its $Q \times Q$ -matrix blocks A, B, C, D can be calculated analytically or numerically. The matriciant can be represented as $M = W(L)W^{-1}(0)$ via a fundamental $2Q \times 2Q$ -matrix $W(z)$ whose columns are linearly independent solutions of the system and whose Wronski determinant is nonzero, $\det W \neq 0$. Taking into account a radiation boundary condition at the exit $z=L$ from the transition layer, i.e., $\mathbf{F}^-(L) = 0$, we find the transmittance and reflectance coefficients due to linear mode transformation in the transition layer as follows:

$$T_n^m = (A - BD^{-1}C)_n^m, \quad R_n^m = -(D^{-1}C)_n^m. \quad (89)$$

They give the amplitudes of the eigenmodes at the exits from the transition layer at $z=L$ in the $+z$ direction and at $z=0$ in the $-z$ direction via the amplitudes of the eigenmodes entering the transition layer at $z=0$:

$$F_n^+(L) = \sum_m T_n^m F_m^+(0), \quad F_n^-(0) = \sum_m R_n^m F_m^+(0). \quad (90)$$

Obviously, the result in Eq. (90) can be easily extended to all eigenmodes participating in the optical response of the grating, i.e., beyond the subset of the coupled eigenmodes $n, m = n_1, \dots, n_Q$, since the transmittance and reflectance coefficients for all of the decoupled modes are given by the geometrical optics: $T_n^m = \delta_n^m$, $R_n^m = 0$.

Thus, in order to find the coefficients of transmittance, T_n^m , and reflectance, R_n^m , by the entire transition layer for a given m it is necessary to calculate the outcome amplitudes $F_n^+(z=L)$ and $F_n^-(z=0)$ of all modes for the special boundary condition of just one, m th mode entering the layer at $z=0$ in the $+z$ direction: $F_m^+(z=0) = \delta_{nm}$ and $F_n^-(z=L) = 0 \quad \forall n$. Then, for an arbitrary incident field represented by a combination of all incident modes $\{F_m^+(z=0)\}$, the field exiting the transition layer can be found in terms of the outcome mode amplitudes,

$F_n^+(L)$ and $F_n^-(0)$, as a sum in Eq. (90) over all incident modes.

The approach formulated above allows one to further simplify calculations of the propagation and diffraction of waves through the transition layer of a grating (within an already simple picture of very few lower-band, propagating or weakly evanescent, eigenmodes contributing to the transmitted field) by accounting for the linear mode coupling, including mutual reflections, between only some of those eigenmodes and assigning the known geometrical optics amplitudes to all of the other eigenmodes.

The equations for the reflection of a single eigenmode (77) or for the mutual transformation between the copropagating eigenmodes (81) and (84) follow from the general result in Eq. (87) if the coupling with all other eigenmodes except the counterpropagating one [i.e., the second and third lines in Eq. (87)] or the coupling with all counterpropagating eigenmodes [i.e., the first term and the third line in Eq. (87)] and the second order [$\sim \alpha^2$; see Eq. (69)] corrections, $\propto \psi_{nm}\psi_{mm}$ and η_{nm} , to the coupling coefficients, respectively, are negligible for a particular transition layer in an inhomogeneous grating.

The details of an application of this mode-coupling approach to the engineering and design of various gratings, in particular, a grating outcoupler for the nonlinear-mixing semiconductor lasers, go beyond the scope of this paper and will be presented elsewhere. However, we illustrate it below by two generic examples.

VIII. TWO TYPES OF MODE COUPLING IN THE WEAKLY INHOMOGENEOUS GRATINGS

We analyzed various possible scenarios of linear mode coupling in the weakly inhomogeneous lamellar dielectric gratings and came to the following practical observation. For the propagating or weakly evanescent eigenmodes, a significant eigenmode coupling usually occurs via three particular generic scenarios: a self-reflection of one eigenmode, a transformation between two copropagating eigenmodes [i.e., the cases (i) and (ii) introduced in Sec. VI above], and a mixed scenario when two copropagating modes simultaneously transform into each other and reflect into two counterpropagating modes of the same two eigenmode band types. Here we elaborate on the self-reflection and mixed scenarios by a detailed discussion of an example of the trapezoidal grating shown in Fig. 1. The results presented below constitute an application of the theory of linear mode coupling developed above and are based on a direct solution of Eqs. (87).

In order to focus on a net, pure effect of linear mode coupling in a transition layer $z \in [0, L]$ with a thickness L , we do not consider any reflection and mode coupling at the boundaries of the transition layer, in this section, say, assuming that the grating is smoothly extending beyond the transition layer so that the geometrical optics applies and the reflection and mode coupling at the layer's borders $z = 0$ and $z = L$ are absent. Below, we compute the transmittance and reflectance coefficients (89) of the transition layer. In order to accomplish this goal we need to solve Eqs. (87) for the special boundary conditions stated next to Eq. (90) and explained in Sec. VII.

Of course, reflection and mode coupling at the borders of an actual inhomogeneous grating of thickness L can be easily taken into account, in addition to the linear mode coupling inside the volume of the nonlamellar grating, by means of the explicit formulas derived in Sec. IV. Then, one can combine these contributions and find an actual optical response of the entire nonlamellar grating to an incident plane wave as is outlined in Sec. IX.

A. Partial reflection of a single eigenmode converting from propagating to evanescent regime in a transition layer of total internal reflection

The first scenario is a partial reflection of an eigenmode into the counterpropagating eigenmode of the same band type in a transition layer where the eigenmode z wave vector squared becomes zero and changes its sign while the z wave vectors of all other eigenmodes are far from zero—in other words, the case of a transition layer where only one eigenmode experiences a total internal reflection and changes its character from propagating to evanescent or vice versa.

For a typical example, we consider a case of the TE eigenmode of the band type $n = 3$ and x wave number $k_x = \pi/(2\lambda_g)$ (see Fig. 3) propagating through the trapezoidal transition layer of a grating which consists of two, alternating in the x direction, dielectric-medium layers with permittivities $\varepsilon_1 = 12$, $\varepsilon_2 = 1$ relevant to the case of a GaAs heterostructure and has a constant grating period $\lambda_g = 2\lambda_1$ with a groove width $d_2(z)$ linearly increasing from $d_2(z = 0) = 4.8\lambda_1/(2\pi)$ at the transition layer entrance to $d_2(z = L) = 8.4\lambda_1/(2\pi)$ at the exit from the transition layer (see Fig. 1). Here λ_1 is the wavelength of a plane wave in the medium with the permittivity ε_1 . It is shorter than the wavelength λ_2 of a plane wave in the second medium, $\lambda_2 = \sqrt{\varepsilon_2/\varepsilon_1}\lambda_1 \approx 3.5\lambda_1$, and corresponds to the wave number $k_1^{(0)} \equiv 2\pi/\lambda_1 = \omega/(c\sqrt{\varepsilon_1})$. Let us consider a transition layer of a relatively large thickness, say, $L = 16\lambda_1/(2\pi)$, corresponding to a weak inhomogeneity for all eigenmodes except the eigenmode $n = 3$, so that the selected eigenmode is decoupled from the eigenmodes of all other band types $n \neq 3$. This decoupling is justified by a fulfillment of the condition (72) of a large detuning between the eigenmodes. For instance, the eigenmode $n = 2$ acquires a relative change of its z wave vector $\Delta k_{z2}/k_{z2} \sim 1/2$ at a transition layer thickness L that amounts to an inhomogeneity scale $\Lambda \sim 2L$. Its detuning from the eigenmode $n = 3$ is $|k_{z2} - k_{z3}| \sim 0.7k_1^{(0)}$. Hence, their beating phase shift, $|k_{z2} - k_{z3}|\Lambda \sim 20$, accumulated over the inhomogeneity scale Λ is larger than 2π .

In the subsequent numerical simulations and graphs, we will assume that the unit of length is equal to $\lambda_1/(2\pi)$. This implies the following values of the grating parameters: $k_1^{(0)} = 1$, $\lambda_1 = 2\pi$, $\lambda_g = 4\pi$, $L = 16$, $d_2(z) = 4.8 + 3.6z/L$.

For the first five lower-band eigenmodes existing in the trapezoidal grating specified above, the dependence of the eigenmode z wave vectors squared on the groove thickness d_2 varying along the inhomogeneous grating is shown in Fig. 7, which presents the solutions to the dispersion equation (14). Of course, one has to look only at the range of groove thicknesses, $4.8 < d_2 < 8.4$, which are actually present in the grating. Here we consider the TE eigenmode $n = 3$ with

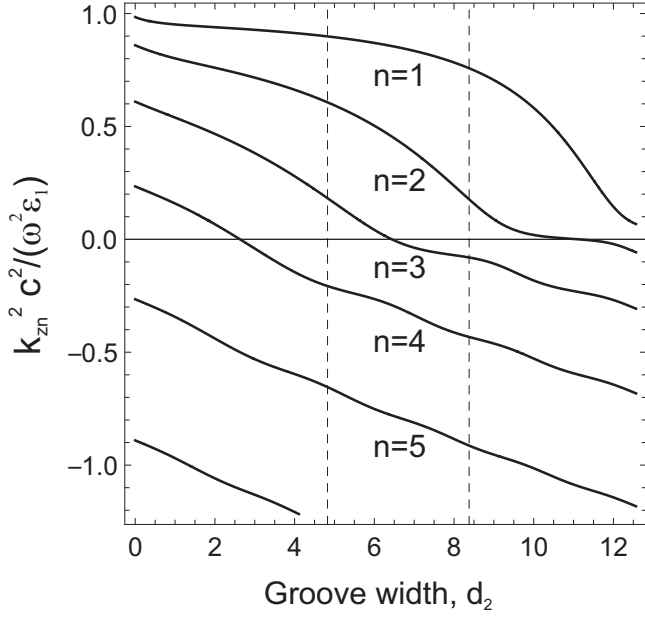


FIG. 7. Dependence of the z wave vectors squared, $\frac{k_{zn}^2 c^2}{\omega^2 \epsilon_1}$, of the TE eigenmodes on the groove width d_2 increasing from $d_2(z=0) = 4.8$ to $d_2(z=L) = 8.4$ along the z axis: the case of the trapezoidal grating of Sec. VIII A (Fig. 1), $k_x = \frac{\pi}{2\lambda_g}$. The unit of length is $\lambda_1/(2\pi)$.

$k_x = \pi/(2\lambda_g)$ whose propagation constant converts from the real values $k_{z3} = \text{Re}(k_{z3})$ at the groove widths $d_2(z) < d_2(z_0)$ to the pure imaginary values $k_{z3} = i\text{Im}(k_{z3})$ at the groove widths $d_2(z) > d_2(z_0)$ as is shown in Fig. 8. The critical groove width $d_2(z_0) = 6.44$ is achieved in the cross section of a total internal reflection located at a level $z_0 = 7.288$ where $k_{z3}(z_0) = 0$. This conversion means a transition of the eigenmode from the propagating regime to the evanescent regime.

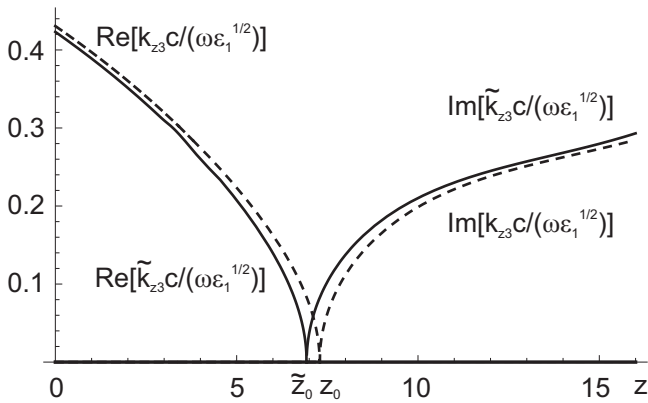


FIG. 8. The bare, $ck_{z3}/(\omega\sqrt{\epsilon_1})$ (dashed curve), and renormalized, $c\tilde{k}_{z3}/(\omega\sqrt{\epsilon_1})$ [Eq. (75); solid curve], propagation constants of the TE eigenmode $n=3$ represented by their real and imaginary parts below ($z < z_0$ or $z < \tilde{z}_0$) and above ($z > z_0$ or $z > \tilde{z}_0$) the related bare [$k_{z3}(z_0) = 0$] or renormalized [$\tilde{k}_{z3}(\tilde{z}_0) = 0$] layers of total internal reflection, respectively, as functions of the coordinate z along the z axis for the case of the trapezoidal grating of Sec. VIII A (Fig. 1). The unit of length is $\lambda_1/(2\pi)$.

Note that due to a grating inhomogeneity and Eq. (75), the propagation constant \tilde{k}_{z3} entering Eqs. (74) and (77) for the amplitudes of the co- and counterpropagating modes is renormalized, i.e., differs from the bare one k_{z3} . This amounts to a shift and deformation of the z profile of the propagation constant, namely, from a dashed-curve profile to a solid-curve profile in Fig. 8. In particular, the layer of a total internal reflection shifts to a new position $\tilde{z}_0 = 6.94$ set by condition $\tilde{k}_{z3}(\tilde{z}_0) = 0$.

It is worthwhile to represent this change of the propagation constant squared in Eq. (75) in a symmetric form

$$\tilde{k}_{zn}^2 - k_{zn}^2 = |\psi_{nm}(z)|^2 - \mu_{nm}(z) \quad (91)$$

that does not involve the second derivative $\partial^2 f_n / \partial z^2$ of the eigenmode profile $f_n(x, z)$, but includes only its first derivative via its scalar product ψ_{nm} with the eigenmode profile in Eq. (71) as well as its squared and scaled norm

$$\mu_{nm}(z) = \left\langle \frac{\partial f_n^\dagger}{\partial z}, \frac{\partial f_n}{\partial z} \right\rangle / \langle f_n^\dagger, f_n \rangle. \quad (92)$$

Similarly to the eigenmode norm in Eq. (18), the latter can be found analytically via the following explicit formula:

$$\begin{aligned} \left\langle \frac{\partial f_n^\dagger}{\partial z}, \frac{\partial f_n}{\partial z} \right\rangle &= \frac{1}{\lambda_g} \sum_{q=1,2} \sum_{s=\pm 1} \sum_{j=\pm 1} \{ d_q c_q^j c_q^{s*} v_1(u) \\ &\quad + i(-1)^q d_q^2 [j c_q^{s*} k'_q b_q^{(j)} - s c_q^j k'_q b_q^{(s)*}] v_2(u) \\ &\quad + d_q^3 s j |k'_q|^2 b_q^{(j)} b_q^{(s)*} v_3(u) \}, \\ c_q^j &= \frac{db_q^{(j)}}{dz} - i j k_q b_q^{(j)} \frac{dd_1}{dz}, \end{aligned} \quad (93)$$

where $u = i(jk_q - sk_q^*)(-1)^q d_q$, $k'_q = dk_q/dz$, $b_q^{(j)} = b_q^\pm$ for $j = \pm 1$. It is expressed via the elementary functions

$$\begin{aligned} v_1(u) &= \frac{e^u - 1}{u}, \quad v_2(u) = \frac{e^u(u-1) + 1}{u^2}, \\ v_3(u) &= \frac{e^u(1-u+u^2/2) - 1}{u^3}, \end{aligned} \quad (94)$$

which behave analytically and tend to unity, one-half, and one-third, respectively, when their argument tends to zero, $u \rightarrow 0$, that is, $v_1(0) = 1$, $v_2(0) = 1/2$, $v_3(0) = 1/3$.

It is easy to prove that the change (91) of the propagation constant squared due to a grating inhomogeneity is invariant under rescaling of the eigenmode profile $f_n \rightarrow f_n/c(z)$ by any function $c(z)$ which depends on z , but not on x . There is a canonical scaling that gives

$$\tilde{k}_{zn}^2 - k_{zn}^2 = \left\langle \bar{f}_n^\dagger, \frac{\partial \bar{f}_n}{\partial z} \right\rangle^2 - \left\langle \frac{\partial \bar{f}_n^\dagger}{\partial z}, \frac{\partial \bar{f}_n}{\partial z} \right\rangle, \quad \bar{f}_n = \frac{f_n}{\|f_n\|}. \quad (95)$$

Due to the latter fact and advantages related to numerical simulations, we employ the eigenmode profile \bar{f}_n with the canonical normalization assuring the unity norm, $\|\bar{f}_n\| = 1$, for all subsequent calculations and graphs.

The profile of the TE($n=3$) eigenfunction, namely, its absolute value $|\bar{f}_3|$ as a function of the coordinates $x \in [0, \lambda_g]$ and $z \in [0, L]$ in the trapezoidal grating, is shown in Fig. 9. It is continuous everywhere, including the boundary between

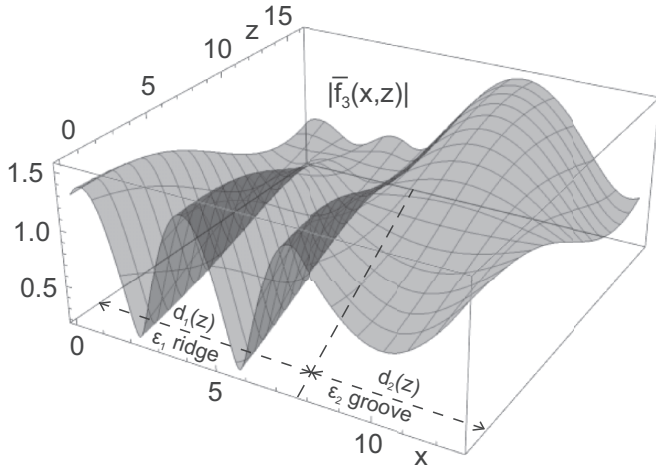


FIG. 9. A 3D plot of the TE($n = 3$)-eigenmode profile, $|\bar{f}_3(x, z)|$, for the case of the trapezoidal grating of Sec. VIII A (Fig. 1). The unit of length is $\lambda_1/(2\pi)$.

the ridge of optically thick material, $\epsilon_1 = 12$, and the groove of optically thin material, $\epsilon_2 = 1$. This boundary crosses the x - z plane along the straight line $x = d_1(z) = \lambda_g - d_2(z)$ going from the point $x = 7.77$ at $z = 0$ to the point $x = 4.17$ at $z = 16$ following the shrinking width $d_1(z)$ of the ridge; cf. Fig. 1.

The phenomenon of linear mode coupling takes place around the critical layer $\tilde{z}_0 = 6.94$ of a total internal reflection where $\tilde{k}_{z3}(\tilde{z}_0) = 0$ (see Fig. 8) and manifests itself in a significant variation of the co- and counterpropagating TE($n = 3$)-eigenmode amplitudes $F_3^+(z)$ and $F_3^-(z)$. A violation of the geometrical optics approximation of constant amplitudes occurs due to an inhomogeneity of the grating along the z axis and is strongly pronounced near the critical layer as is shown in Fig. 10.

Both amplitudes grow to large values in a close vicinity of the critical layer. However, their contributions to the

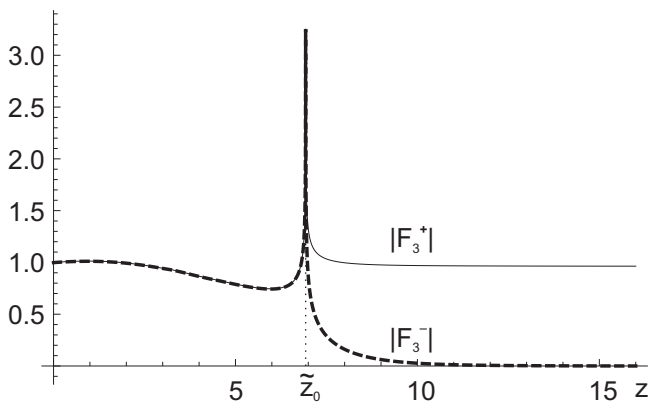


FIG. 10. Spatial variation of the co- and counterpropagating TE($n = 3$)-eigenmode amplitudes $|F_3^+(z)|$ (solid curve) and $|F_3^-(z)|$ (dashed curve) due to linear mode coupling in the course of propagation through a layer of total internal reflection in the trapezoidal grating of Sec. VIII A (Figs. 1, 7–9). An incoming-mode amplitude equals unity, $F_3^+(z = 0) = 1$. The unit of length is $\lambda_1/(2\pi)$.

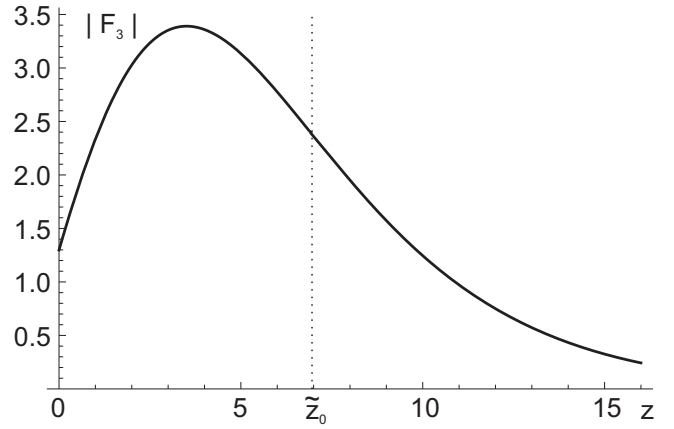


FIG. 11. A profile of the amplitude of the electric field $|F_3(z)|$ that has emerged due to TE($n = 3$)-eigenmode reflection in the inhomogeneous (trapezoidal) grating of Sec. VIII A. It corresponds, via Eqs. (73) and (76), to the linear mode coupling shown in Fig. 10. The unit of length is $\lambda_1/(2\pi)$.

amplitude of the combined field in Eq. (76) mostly cancel, which results in a smooth profile of the total field shown in Fig. 11. Actually, the co- and counterpropagating mode amplitudes remain finite everywhere due to a presence of a finite, relatively small absorption, $\text{Im}(\epsilon_{1,2}) \neq 0$, which is taken into account via formulas derived in Appendix A. Immediately behind the critical layer, the amplitude of the counterpropagating, reflected mode drops to zero in a steep manner while the copropagating, transmitted mode amplitude returns to the unity value. This pattern corresponds to an exponential attenuation of the combined field of the two, co- and counterpropagating modes at $z > \tilde{z}_0$ due to a presence of the geometrical optics factor $\exp[-\int_0^z \text{Im}(\tilde{k}_{z3}) dz]$. The latter suppresses the field transmitted by the eigenmode which acquires an evanescent wave number $\text{Im}(\tilde{k}_{z3}) > 0$ at $z > \tilde{z}_0$. At the entrance to the trapezoidal grating layer, $z = 0$, the amplitude of the reflected, counterpropagating mode is almost equal to the unity amplitude of the incident mode since we consider a relatively thick transition layer capable of a nearly total reflection.

The resulting profile of the amplitude of the electric field [given by Eqs. (73) and (76)] that emerges in the inhomogeneous grating is shown in Fig. 11. It demonstrates a typical, nearly threefold amplification of the electric field relatively close to, but notably in front of, the critical layer of a total internal reflection. Obviously, the effect of linear mode coupling in the inhomogeneous gratings is significantly complicated by a nontrivial dependence of the eigenmode profile f_n and z wave vector k_{zn} on the varying parameters of the inhomogeneous grating.

At the same time, the linear mode coupling described above for the case of a single-eigenmode reflection in the inhomogeneous grating is similar to the one taking place in the classical one-dimensional problem of a plane wave propagating and reflecting along the z axis within a dielectric medium which is homogeneous and unbounded in the x - y plane, but inhomogeneous along the z axis, and which permittivity is linearly decreasing from positive to negative values,

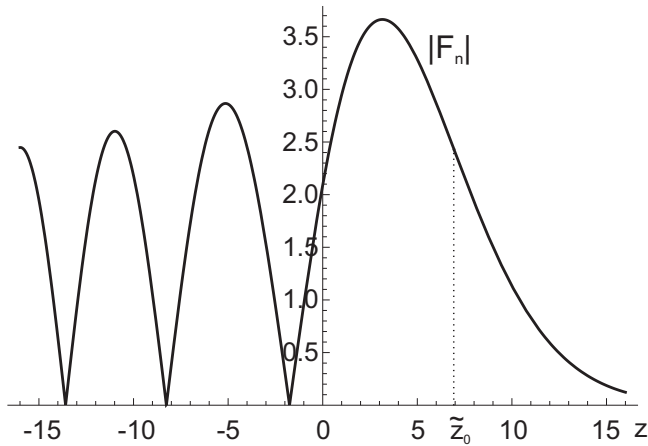


FIG. 12. A profile of the amplitude of the electric field $|F_n(z)|$ that has emerged due to plane-wave reflection in the unbounded dielectric medium whose permittivity is inhomogeneous only along the z axis, $\varepsilon(z) = \varepsilon'(\tilde{z}_0)(\tilde{z}_0 - z)$. The slope of the plane-wave wave number squared, $\varepsilon(z)\omega^2/c^2$, is the same as the slope of the eigenmode wave number squared, \tilde{k}_{z3}^2 , in Figs. 10 and 11 for the trapezoidal grating of Sec. VIII A at $z = \tilde{z}_0$. An incoming-mode amplitude equals unity, $F_n^+(z = -16) = 1$. The unit of length is $\lambda_1/(2\pi)$.

$\varepsilon(z) = \varepsilon'(\tilde{z}_0)(\tilde{z}_0 - z)$, crossing zero value at the critical layer \tilde{z}_0 of a total internal reflection. The latter, classical problem can be solved by means of the same equations of linear mode coupling (73)–(77) if one sets $f_n(x, z) \equiv 1$, $\Phi_n(z) \equiv 1$, $k_1 = k_2 = 0$, $\varepsilon_1 = \varepsilon_2 = \varepsilon(z)$, and $\tilde{k}_{zn}^2 = \varepsilon\omega^2/c^2$, i.e., considers a plane wave in a homogeneous medium instead of a nontrivial eigenmode in a lamellar grating. The corresponding solution for the absolute value of the total electric field is shown in Fig. 12. We skip the graph for the amplitudes of the co- and counterpropagating plane waves since it is almost the same as the one shown in Fig. 10. The similarity of these plane-wave graphs to the graphs in Figs. 10 and 11 for the eigenmode reflection in an inhomogeneous grating is remarkable.

In particular, this means that the spatial profile of the co- and counterpropagating mode amplitudes in the vicinity of the critical layer of a total internal reflection in an inhomogeneous grating can be described analytically in a universal form via the Airy functions. The Airy functions constitute a well-known solution to the Helmholtz equation (7) for the plane-wave reflection in an unbounded medium. Note also that in Fig. 12 we extended the considered layer of the dielectric medium to a wider interval of z coordinate, $z \in [-16, 16]$, as compared to the trapezoidal grating's interval $z \in [0, 16]$ depicted in Figs. 10, 11, in order to demonstrate a formation of a standing wave in front of the total internal reflection layer $\tilde{z}_0 = 6.94$ due to an interference of the incident and reflected waves whose amplitudes are almost equal at $z < 0$. Yet, we employed the same boundary conditions at the layer's borders, namely, $F_n^+(z = -16) = 1$, $F_n^-(z = 16) = 0$.

The effect of the linear mode coupling in the grating transition layer of a finite thickness L (see Fig. 1) can also be presented in terms of (a) the amplitude $F_3^-(z = 0)$ of the reflected, counterpropagating mode exiting the grating at the bottom border $z = 0$ in the $-z$ direction which is related to the internal reflectance coefficient $R(L) = F_3^-$

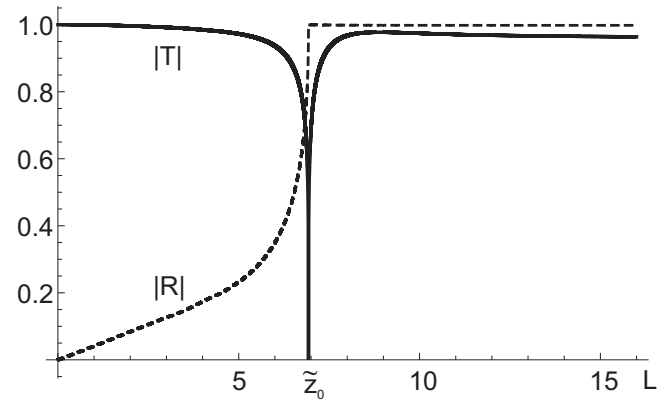


FIG. 13. Absolute values of the internal transmittance and reflectance coefficients, $|T|$ (solid curve) and $|R|$ (dashed curve), which give the amplitudes of the co- and counterpropagating modes, $F_3^+(z = L)$ and $F_3^-(z = 0)$, respectively, as functions of thickness L of the trapezoidal grating transition layer of Sec. VIII A (Fig. 1) with a fixed gradient of the groove width $d_2(z) = 4.8 + 3.6z/16$. The unit of length is $\lambda_1/(2\pi)$.

($z = 0$)/ $F_3^+(z = 0)$ and (b) the amplitude $F_3^+(z = L)$ of the transmitted, copropagating mode exiting the grating at the top border $z = L$ in the $+z$ direction which is related to the internal transmittance coefficient $T(L) = F_3^+(z = L)/F_3^+(z = 0)$. The dependence of these reflectance and transmittance coefficients on the thickness L of the trapezoidal grating layer with a fixed gradient of the groove width $d_2(z) = 4.8 + 3.6z/16$, but with the groove width at the top of the grating transition layer, $d_2(L) = 4.8 + 3.6L/16$, depending on the layer thickness L , is shown in Fig. 13. With an increasing thickness L , the reflectance accumulated in the grating transition layer is growing and reaches the maximum value $|R| = 1$ when the layer of total internal reflection $z = \tilde{z}_0 = 6.94$ appears inside the grating transition layer. At the same time, the transmittance drops from the maximum value $|T| = 1$ to zero at the grating transition layer thickness close to the position of the total internal reflection layer, $L \approx \tilde{z}_0 = 6.94$, and then returns to its new maximum value when the layer of total internal reflection enters deeper inside the grating.

Note that the transmittance coefficient shown in Fig. 13 is defined in terms of the amplitude of the copropagating mode $F_3^+(L)$. Of course, the value of the actual field penetrating to the top of the grating transition layer, which is given by Eqs. (73) and (76) at $z = L$ and plotted in Fig. 14, does not drop to zero at $L = \tilde{z}_0$, but becomes exponentially attenuated at $L > \tilde{z}_0$ in virtue of the geometrical optics factor $\exp[-\int_{\tilde{z}_0}^L \text{Im}(\tilde{k}_{z3})dz]$ accumulated on a path $z \in [\tilde{z}_0, L]$ where the eigenmode is evanescent. The field at the top of the trapezoidal grating layer achieves a sharp, finite maximum when the layer thickness is such that the layer of total internal reflection is located near the top of the trapezoidal grating, $L \approx \tilde{z}_0$.

B. Mutual transformation of two eigenmodes traversing a region of their degeneracy and reflection

The second scenario takes place when two eigenmodes of neighboring orders n , $n + 1$ and of an x wave vector k_x

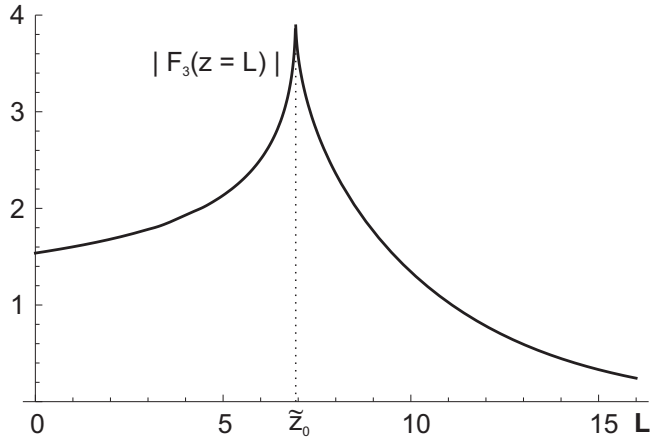


FIG. 14. The amplitude, $|F_3(z=L)|$, of the field penetrated to the top of the trapezoidal grating transition layer (Fig. 1) with a fixed gradient of the groove width $d_2(z) = 4.8 + 3.6z/16$ as a function of the layer thickness L . The amplitude of the mode which is entering the layer at its bottom in the $+z$ direction is unity, $F_3^+(z=0) = 1$. The unit of length is $\frac{\lambda_1}{2\pi}$.

near zero or $\pm\pi/\lambda_g$ have very close z wave vectors squared, separated only by a relatively narrow band gap, and enter a transition layer of total internal reflection. Thus, their z wave vectors squared change their signs almost simultaneously and the two eigenmodes experience a substantial mutual transformation into each other and reflection in the layer where both of them are converting from propagating to evanescent or vice versa. We are interested in the mode coupling at the transition layer, so we just compute the transmittance and reflectance coefficients (89) of the transition layer by solving Eqs. (87) for the appropriate boundary conditions explained in Sec. VII and skip a discussion of reflection and mode coupling at the layer's boundaries.

As a typical example, we consider two TE eigenmodes of the band types $n = 2, 3$ and x wave number $k_x = 0$ (see Fig. 15) propagating through the trapezoidal transition layer of a grating made out of GaAs similar to the one discussed in Sec. VIII A (Fig. 1). However, we change some parameters of the trapezoidal grating. Namely, now the constant grating period is $\lambda_g = 2.5\lambda_1$, the groove width $d_2(z)$ is linearly increasing from $d_2(z=0) = 9.3\lambda_1/(2\pi)$ to $d_2(z=L) = 14\lambda_1/(2\pi)$ from the transition layer entrance to its exit, and the thickness of the transition layer is two times shorter, $L = 8\lambda_1/(2\pi)$, which makes the effect of the linear mode coupling more pronounced but keeps the inhomogeneity relatively weak. In this case the two selected eigenmodes $n = 2, 3$ are strongly coupled between themselves, but are decoupled from the eigenmodes of all other band types $n \neq 2, 3$. This decoupling is justified by a fulfillment of the condition (72) of a large detuning between the eigenmodes. For instance, the eigenmode $n = 1$ acquires a relative change of its z wave vector $\Delta k_{z1}/k_{z1} \sim 1/3$ at a transition layer thickness L that amounts to an inhomogeneity scale $\Lambda \sim 3L$. Its detuning from the eigenmodes $n = 2, 3$ is $|k_{z1} - k_{z2,3}| \sim 0.4k_1^{(0)}$, so the beating phase shift $|k_{z1} - k_{z2,3}|\Lambda \sim 10$ accumulated over the inhomogeneity scale Λ is larger than 2π .

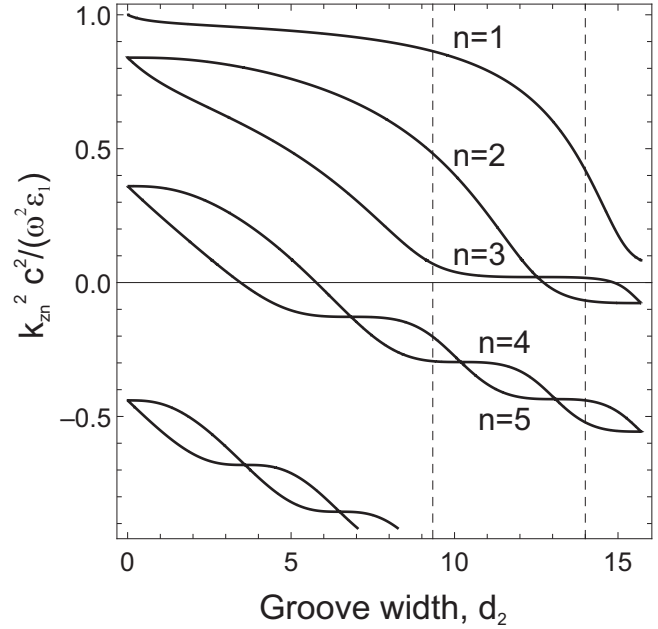


FIG. 15. Dependence of the z wave vectors squared, $\frac{k_{zn}^2 c^2}{\omega^2 \epsilon_1}$, of the TE eigenmodes on the groove width d_2 increasing from $d_2(z=0) = 9.3$ to $d_2(z=L) = 14$ along the z axis: the case of the trapezoidal grating of Sec. VIII B (Fig. 1), $k_x = 0$. The unit of length is $\lambda_1/(2\pi)$.

We again set $\lambda_1/(2\pi)$ to be the unit of length, which implies the following values of the grating parameters: $k_1^{(0)} = 1$, $\lambda_1 = 2\pi$, $\lambda_g = 5\pi$, $L = 8$, $d_2(z) = 9.3 + 4.7z/L$.

The dependence of the eigenmode z wave vectors squared on the groove thickness d_2 varying along the inhomogeneous grating is shown in Fig. 15 for the first five lower-band eigenmodes living in this trapezoidal grating. They correspond to the first five solutions to the dispersion equation (14). Of course, one has to look only at the range of groove thicknesses, $9.3 < d_2 < 14$, which are actually present in the grating. The eigenmode $n = 2$ has the propagation constant which converts from the real values $k_{z2} = \text{Re}(k_{z2})$ at the groove widths $d_2(z) < d_2(z_0)$ to the pure imaginary values $k_{z2} = i\text{Im}(k_{z2})$ at the groove widths $d_2(z) > d_2(z_0)$ as is shown in Fig. 16. The critical groove width $d_2(z_0) \approx 12.72$ is achieved in the cross section of the ($n = 2$)-mode total internal reflection located at a level $z_0 \approx 5.83$ where $k_{z2}(z_0) = 0$. This conversion means a transition of the eigenmode $n = 2$ from the propagating regime to the evanescent regime. The propagation constant of the eigenmode $n = 3$ is real-valued, $k_{z3} = \text{Re}(k_{z3})$, throughout the entire transition layer.

Due to the grating inhomogeneity and Eqs. (75) and (91), the propagation constants \tilde{k}_{zn} entering Eqs. (87) for the amplitudes of the coupled co- and counterpropagating modes are renormalized, i.e., differ from the bare ones k_{zn} . This amounts to a shift and deformation of the z profile of the propagation constants, namely, from the dashed-curve profiles to the solid-curve profiles in Fig. 16. In particular, the layer of total internal reflection for the $n = 2$ eigenmode shifts from $z_0 = 5.83$ to a new position $\tilde{z}_0 \approx 5.2$ set by the condition $\tilde{k}_{z2}(\tilde{z}_0) = 0$.

In the subsequent calculations, we again employ the eigenfunctions \tilde{f}_n with the canonical normalization, $||\tilde{f}_n|| = 1$, as is

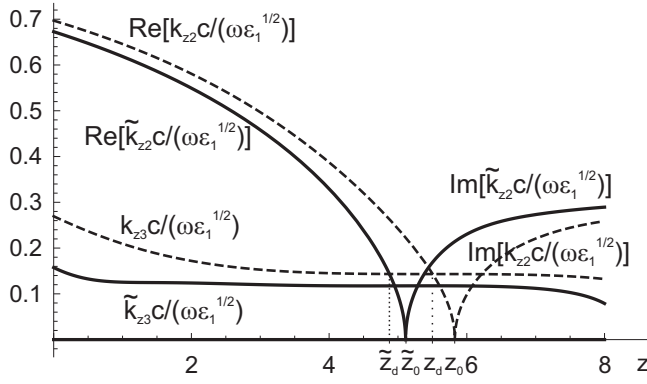


FIG. 16. The bare, $ck_{z2}/(\omega\sqrt{\epsilon_1})$ and $ck_{z3}/(\omega\sqrt{\epsilon_1})$ (dashed curves), and renormalized, $\tilde{c}k_{z2}/(\omega\sqrt{\epsilon_1})$ and $\tilde{c}k_{z3}/(\omega\sqrt{\epsilon_1})$ [Eq. (75); solid curves], propagation constants of the TE eigenmodes $n = 2, 3$ represented (i) by the real and imaginary parts of the eigenmode z wave number below ($z < z_0$ or $z < \tilde{z}_0$) and above ($z > z_0$ or $z > \tilde{z}_0$) the related bare [$k_{z2}(z_0) = 0$] or renormalized [$\tilde{k}_{z2}(\tilde{z}_0) = 0$] layers of total internal reflection, respectively, for the eigenmode $n = 2$, and (ii) by the real-valued z wave number for the eigenmode $n = 3$ as a function of the coordinate z along the z axis for the case of the trapezoidal grating of Sec. VIII B (Fig. 1). The unit of length is $\lambda_1/(2\pi)$.

defined in Eq. (95). Importantly, the spatial profiles of these eigenfunctions for the TE eigenmodes $n = 2$ and $n = 3$ are different from each other even at the critical level $z_d \approx 5.5$ where their propagation constants become degenerate, that is, $k_{z2}(z_d) = k_{z3}(z_d)$. The latter point corresponds to an intersection of the dispersion curves $n = 2$ and $n = 3$ in Fig. 15 at the groove width $d_2(z_d) \approx 12.53$. Qualitatively, the profiles of the eigenfunctions f_2 and f_3 look similar to the one shown in Fig. 9, and we skip their plots here.

The phenomenon of a cross-band transformation between the amplitudes of the modes copropagating in the $+z$ direction, F_2^+ and F_3^+ , and, simultaneously, between the amplitudes of the modes copropagating in the $-z$ direction, F_2^- and F_3^- , takes place in the region of degeneracy around the degeneracy level $\tilde{z}_d \approx 4.9$, where the eigenmode propagation constants, renormalized by an inhomogeneity, approach each other and coincide, $\tilde{k}_{z2}(\tilde{z}_d) = \tilde{k}_{z3}(\tilde{z}_d)$. (The renormalized degeneracy level $\tilde{z}_d \approx 4.9$ is shifted from its bare position $z_d \approx 5.5$; see Fig. 16.) The reason for this cross-band transformation is an increased wavelength of beating between the eigenmodes which, in the region of degeneracy, becomes larger than the inhomogeneity scale. This leads to a violation of the condition (72), required for the applicability of the geometrical optics approximation. In the present case, when both eigenmodes are close to their critical layers of total internal reflection, this phenomenon is superposed on the phenomenon of linear mode coupling between the co- and counterpropagating modes within each eigenmode band, i.e., between the amplitudes F_2^+ and F_2^- as well as between the amplitudes F_3^+ and F_3^- , which corresponds to the reflection discussed in Sec. VIII A. Thus, all four mode amplitudes, F_2^\pm and F_3^\pm , become coupled within the system of four differential equations (87) and transform into each other in the course of propagation.

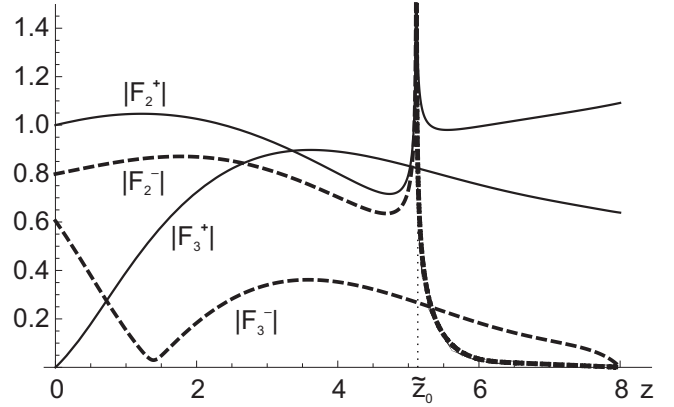


FIG. 17. Mutual transformation between all co- and counter-propagating mode amplitudes of the TE eigenmodes $n = 2, 3$ due to the linear mode coupling in the course of propagation through the trapezoidal grating layer of Sec. VIII B (Figs. 1 and 15): $|F_{2,3}^+(z)|$ (solid curves), $|F_{2,3}^-(z)|$ (dashed curves). Only one mode $n = 2$, and only from one side ($z = 0$), is entering the grating: $F_2^+(z = 0) = 1$, $F_3^+(z = 0) = 0$, $F_{2,3}^-(z = L) = 0$. The unit of length is $\lambda_1/(2\pi)$.

This phenomenon is illustrated in Figs. 17 and 18 for the case when only the eigenmode $n = 2$ enters the trapezoidal grating layer at $z = 0$, namely, for the boundary condition $F_2^+(z = 0) = 1$, $F_3^+(z = 0) = 0$.

Figure 17 shows the spatial profiles of the co- and counter-propagating mode amplitudes for both eigenmodes. It reveals two major effects. First, the incoming mode of the amplitude $F_2^+(z = 0) = 1$ generates the reflected, counterpropagating mode, F_2^- , in the vicinity of its total-internal-reflection layer $\tilde{z}_0 \approx 5.2$ in a pattern very similar to the one discussed in Sec. VIII A, Fig. 10. However, now the amplitude of the reflected mode, $F_2^-(z = 0) \approx 0.8$, is less than the incoming mode amplitude $F_2^+(z = 0) = 1$. This means that the reflection within the same eigenmode band, $n = 2$, becomes only partial; i.e., the related reflectance coefficient defined in Eqs. (89) and (90) is substantially less than unity, $|R_2^2| \approx$

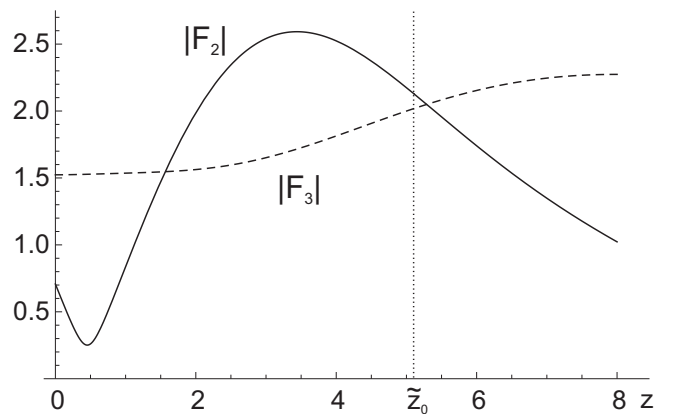


FIG. 18. The profiles of the total amplitude of the electric field, Eqs. (76) and (85), in each of the two TE eigenmodes $n = 2$ and $n = 3$, $|F_2(z)|$ (solid curve) and $|F_3(z)|$ (dashed curve), corresponding to the linear mode coupling in the inhomogeneous (trapezoidal) grating of Sec. VIII B shown in Fig. 17. The unit of length is $\lambda_1/(2\pi)$.

$0.8 < 1$, due to the cross-band mode transformation from the band $n = 2$ into the other band $n = 3$. Second, the incoming eigenmode of the band $n = 2$, $F_2^+(z = 0) = 1$, generates the copropagating mode of the other band $n = 3$ whose amplitude achieves a value on the order of unity, $|F_3^+| \sim 1$, already in the second quarter of the trapezoidal grating and remains such throughout the rest of the grating. This results in a large cross-band transmittance coefficient $|T_3^2| \sim 1$ [see Eqs. (89) and (90)]. Moreover, its counterpropagating counterpart, F_3^- , also accumulates a significant amount of reflection growing from the zero value, $F_3^-(z = L) = 0$, at the top of the trapezoidal grating layer $z = L$ toward a relatively large value $|F_3^-(z = 0)| \approx 0.6$ at the bottom. Thus, the cross-band reflectance is also well pronounced: $|R_3^2| \approx 0.6$.

Figure 18 shows the corresponding spatial profiles of the total amplitude of the electric field in each of the two eigenmodes, $|F_2(z)|$ and $|F_3(z)|$, defined in Eqs. (76) and (85). The field profile of the eigenmode $n = 2$ is reminiscent of the one emerging due to reflection from the layer of total internal reflection (see Fig. 11 and its discussion in Sec. VIII A). However, in the present case the field amplification in front of the layer of total internal reflection is significantly less pronounced (only about 2.5 instead of 3.5) since, as is stated above, the same-band reflection is only partial, not complete. Besides, an exponential attenuation of the eigenmode amplitude $|F_2(z)|$ is also less pronounced since the geometrical optics factor $\exp[-\int_0^z \text{Im}(\tilde{k}_{z2})dz]$ remains on the order of unity everywhere behind the total-internal-reflection layer at $\tilde{z}_0 < z < L$ due to relatively small values of the evanescent wave number $\text{Im}(\tilde{k}_{z2}) > 0$ (see Fig. 16). The field of the other eigenmode, $n = 3$, acquires an amplitude of the same order of magnitude, $|F_3| \sim 2$, due to a strong cross-band transformation. Its profile is almost constant, with only a slight increase toward the top exit from the trapezoidal grating, since a level of the total-internal-reflection layer for the eigenmode $n = 3$ corresponds to the groove width $d_2 \approx 15$ which is beyond the range of groove widths in the actual grating layer, $9.3 < d_2 < 14$ (see Fig. 15). Besides, the standing-wave structure (like in Fig. 12) of the eigenmode $n = 3$ cannot be seen on a scale of the trapezoidal grating layer since, according to Fig. 16, the eigenmode z wave number is so small, $\tilde{k}_{z3} \approx 0.1$, that the corresponding wavelength is much larger than the grating layer thickness, $2\pi/\tilde{k}_{z3} \approx 20\pi \gg L = 8$.

The result of the mutual mode transformation is significantly different in the case when only eigenmode $n = 3$ enters the trapezoidal grating layer at $z = 0$, i.e., for the boundary condition $F_3^+(z = 0) = 1$, $F_2^+(z = 0) = 0$. It is shown in Fig. 19 and should be compared against the one in Figs. 17 and 18 discussed above.

In this case, the amplitude of the incoming mode remains close to unity everywhere inside the grating layer, $|F_3^+| \sim 1$. It is only weakly coupled to the counterpropagating mode F_3^- since the eigenmode propagation constant \tilde{k}_3 shown in Fig. 16 varies only a little, and hence, slowly across the grating layer, so that the resulting reflection is small, $|F_3^-(z = 0)| \approx 0.42$. The cross-band transformation of F_3^+ into the copropagating mode of the other band, $n = 2$, does take place, but it is not very strong. The amplitude of this mode at the top exit from the grating layer only achieves a small value $|F_2^+(z = L)| \sim 0.2$; i.e., the cross-band transmittance

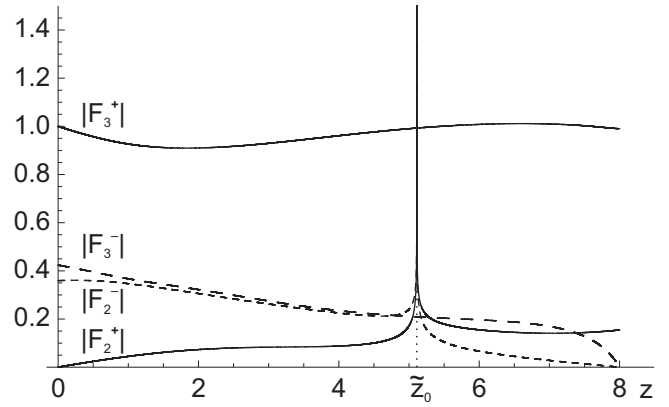


FIG. 19. Mutual transformation between all co- and counter-propagating mode amplitudes of the TE eigenmodes $n = 2, 3$ due to the linear mode coupling in the course of propagation through the trapezoidal grating layer of Sec. VIII B (Figs. 1 and 15): $|F_{2,3}^+(z)|$ (solid curves), $|F_{2,3}^-(z)|$ (dashed curves). Only one mode $n = 3$, and only from one side ($z = 0$), is entering the grating: $F_3^+(z = 0) = 1$, $F_2^+(z = 0) = 0$, $F_{2,3}^-(z = L) = 0$. The unit of length is $\lambda_1/(2\pi)$.

coefficient defined in Eqs. (89) and (90) is small, $|T_3^2| \sim 0.2$. In the vicinity of its total-internal-reflection layer $\tilde{z}_0 \approx 5.2$, this mode also generates the counterpropagating mode which reaches the amplitude $|F_2^-(z = 0)| \approx 0.36$ at the bottom exit from the trapezoidal grating layer. Thus, there is a significant cross-band reflectance, $|R_3^2| \approx 0.36$. The corresponding spatial profiles of the total amplitude of electric field in each of the two eigenmodes, $|F_2(z)|$ and $|F_3(z)|$, are similar to that shown in Fig. 18, and we skip their plots here. The main difference now is that the amplitude of the eigenmode $n = 2$ is an order of magnitude smaller than the amplitude of the eigenmode $n = 3$.

It is instructive to analyze the integral transmittance and reflectance of the grating transition layer as functions of the thickness L of the trapezoidal grating transition layer (Fig. 1) given by Eqs. (89) and (90). The corresponding intra- and cross-band internal transmittance and reflectance coefficients, $|T_n^2|$ and $|R_n^2|$ for $n = 2$ and 3 , respectively, are shown in Fig. 20 for the case of the trapezoidal grating transition layer of Sec. VIII B (Fig. 1) with a fixed gradient of the groove width $d_2(z) = 9.3 + 4.7z/8$. They describe the particular boundary conditions when only one mode $m = 2$, from only one side ($z = 0$), is entering the grating: $F_2^+(z = 0) = 1$, $F_3^+(z = 0) = 0$, $F_{2,3}^-(z = L) = 0$. The intraband transmittance $|T_2^2(L)|$ behaves similarly to the case of a single reflecting eigenmode discussed in Sec. VIII A, Fig. 13. However, the intraband reflectance is noticeably suppressed, compared to Fig. 13, and only grows to a value $|R_2^2| \approx 0.8$ even when the total-internal-reflection layer, $\tilde{z}_0 = 5.2$, of the eigenmode $n = 2$ appears inside the grating transition layer. This occurs due to a significant cross-band transformation of the eigenmode $n = 2$ into the other eigenmode $n = 3$. The latter phenomenon also results in an appearance of the cross-band transmittance, T_3^2 , and reflectance, R_3^2 , both of which grow with an increasing thickness of the grating layer, reach a maximum value when the total-internal-reflection layer of the eigenmode $n = 2$

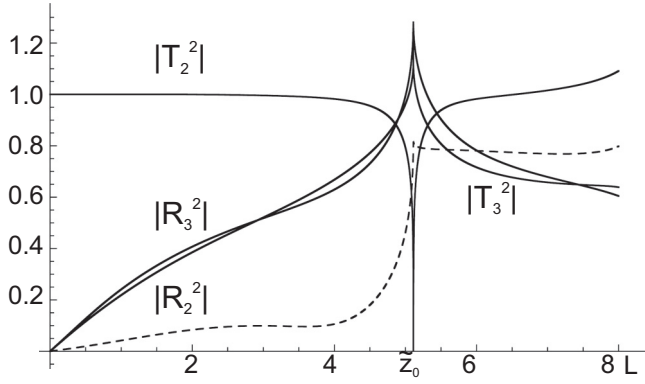


FIG. 20. The absolute values of the internal transmittance and reflectance coefficients, $|T_n^2|$ (solid curve) and $|R_n^2|$ (dashed curve), which are defined in Eqs. (89) and (90) and give the amplitudes of the co- and counterpropagating modes, $F_n^+(z=L)$ and $F_n^-(z=0)$, $n = 2, 3$, respectively, as the functions of thickness L of the trapezoidal grating transition layer of Sec. VIII B (Fig. 1) with a fixed gradient of the groove width $d_2(z) = 9.3 + 4.7z/8$. They describe the integral transmittance and reflectance of the grating transition layer if only one mode $m = 2$, and only from one side ($z = 0$), is entering the grating: $F_2^+(z=0) = 1$, $F_3^+(z=0) = 0$, $F_{2,3}^-(z=L) = 0$. The unit of length is $\lambda_1/(2\pi)$.

enters the grating transition layer, and remain substantial (~ 0.6) even at a larger grating-layer thickness, $L > \tilde{z}_0 = 5.2$.

IX. AN OVERALL OPTICAL RESPONSE AND MULTIPLE MODE COUPLING IN THE NONLAMELLAR GRATINGS

Remarkably, by taking all of the propagating low-band eigenmodes $n = 1, \dots, n_0$ and a sufficient number of the least-evanescent eigenmodes $n = n_0 + 1, \dots, n^*$ into consideration we can obtain a very accurate description of an overall optical response, that is, the transmission and reflection of an incident wave, for any inhomogeneous, nonlamellar grating, even if a grating inhomogeneity does not satisfy the conditions of weakness in Eqs. (69) and (72). It is sufficient to simply combine the effects of the mode coupling and reflection at the sharp boundaries and in the transition layer of the nonlamellar grating described in Sec. IV and Sec. VII, respectively.

Yet, the number Q of eigenmodes n_1, \dots, n_Q to be included in Eqs. (87) for an accurate description of the transition layer in practice remains quite small if the grating is relatively thick. For instance, we calculated accurately the mode transformation for the trapezoidal grating transition layer (Fig. 1) for the cases discussed in Sec. VIII by taking into account the first five eigenmodes in Eqs. (87). The results prove that the approximation of a single reflecting eigenmode or two coupled, close to degeneracy, eigenmodes employed in Sec. VIII correctly describe the main features of linear mode coupling in the inhomogeneous gratings. Of course, the mode coupling at the sharp boundaries of the grating could excite many other eigenmodes, outside of the subset $n = n_1, \dots, n_Q$. However, after penetrating inside the transition layer, they stay decoupled and just follow their geometrical optics approximation. So, their contribution to the grating's optical response can be easily taking into account via the standard coefficients

of transmittance and reflectance discussed in Sec. IV. Also, Eqs. (87) are convenient for numerical calculations since the accuracy of the approximation can be estimated by simply comparing the results for $Q = n^*$ and, say, $Q = n^* + 1$ or $n^* + 2$ eigenmodes.

Clearly, the approach of mode coupling is complementary to the standard perturbation theory within the usual Fourier-Rayleigh expansion method [4–6], which is great for the opposite, well-known case of a very thin grating.

It is straightforward to extend the mode-coupling approach illustrated in the present paper to the nonlamellar gratings characterized by different, e.g., parabolic, profiles of the groove width and other grating parameters along a transition layer. However, for the sake of space, we skip the discussion of these simulations.

Finally, we present the explicit general formulas for the overall optical response of a nonlamellar grating to an incident wave. Let us consider a generic setup similar to that of Sec. IV, that is, a finite grating layer $z \in [0, L]$ of a width L along the vertical axis z sandwiched between the homogeneous substrate ($z < 0$) and cover ($z > L$) media with the permittivities ε_1 and ε_2 , respectively. However, now we deal with a transition layer of an inhomogeneous grating, not a lamellar grating. Suppose a superposition of the plane-wave harmonics, such as in Eqs. (43) and (61), with the amplitudes $g_p(z) = g_p^+(0)e^{ik_{z,p}^{(1)}z}$ is incident onto this nonlamellar grating at its bottom boundary $z = 0$ in the $+z$ direction from the substrate. The top boundary $z = L$ is subject to the radiation boundary conditions. The optical response of the grating, that is, the amplitudes $g_p^+(z) = g_p^+e^{ik_{z,p}^{(2)}z}$ of the plane-wave harmonics emitted from the top grating boundary $z = L$ into the cover in the $+z$ direction and the amplitudes $g_p^-(z) = g_p^-e^{-ik_{z,p}^{(1)}z}$ of the plane-wave harmonics emitted from the bottom grating boundary $z = 0$ back into the substrate in the $-z$ direction, can be found as follows:

$$\mathbf{g}^+ = T^+(L)\mathbf{F}^+(L), \quad \mathbf{g}^- = R^+(0)\mathbf{g}^+(0) + T^-(0)\mathbf{F}^-(0). \quad (96)$$

Here $\mathbf{g}^+ = \{g_p^+\}^T$, $\mathbf{g}^- = \{g_p^-\}^T$, $\mathbf{g}^+(0) = \{g_p^+(0)\}^T$ are the column vectors of the amplitudes of the corresponding plane-wave harmonics. The column vectors $\mathbf{F}^+(L) = \{F_n^+(L)\}^T$ and $\mathbf{F}^-(0) = \{F_n^-(0)\}^T$ stand for the amplitudes of the eigenmodes incident from inside the grating onto its top boundary $z = L$ in the $+z$ direction and bottom boundary $z = 0$ in the $-z$ direction, respectively. The matrix $T^+(L)$ is the matrix T_p^n , Eq. (53), or \tilde{T}_p^n , Eq. (59), of transmittance of the eigenmodes into the plane-wave harmonics for the case of the TE or TM eigenmodes outcoming in the $+z$ direction from the grating at its top boundary $z = L$. It was calculated analytically in Sec. IV. The matrix $T^-(0)$ is a similar transmittance matrix for the case of the TE or TM eigenmodes outcoming in the $-z$ direction from the grating at its bottom boundary $z = 0$. The matrix $R^+(0)$ is the matrix R_p^{\prime} , Eq. (66), or \tilde{R}_p^{\prime} , Eq. (67), of reflectance of the plane-wave harmonics for the case of the TE or TM wave incoming in the $+z$ direction from the substrate into the grating at its bottom boundary $z = 0$. It was also calculated analytically in Sec. IV.

The amplitudes $\mathbf{F}^\pm(0)$ or $\mathbf{F}^\pm(L)$ of the eigenmodes propagating inside the nonlamellar grating in the $\pm z$ direction

in the vicinity of the bottom, $z = +0$, or top, $z = L - 0$, boundaries of the grating, respectively, can be found from the four boundary conditions for these amplitudes at $z = 0$ and $z = L$ written via the matrices of transmittance and reflectance by (i) the entire transition layer, T^\pm and R^\pm , and (ii) the boundaries, $T^+(0)$, $R^-(0)$, $R^+(L)$, as follows:

$$\begin{aligned} \mathbf{F}^+(0) &= R^-(0)\mathbf{F}^-(0) + T^+(0)\mathbf{g}^+(0), \\ \mathbf{F}^-(0) &= R^+\mathbf{F}^+(0) + T^-\mathbf{F}^-(L), \\ \mathbf{F}^+(L) &= R^-\mathbf{F}^-(L) + T^+\mathbf{F}^+(0), \\ \mathbf{F}^-(L) &= R^+(L)\mathbf{F}^+(L). \end{aligned} \quad (97)$$

Here the matrices T^+ and R^+ are the matrices T_n^m and R_n^m of the integral transmittance and reflectance of the eigenmodes propagating in the $+z$ direction through the entire transition layer as defined by Eqs. (90). They are calculated in a general form in Sec. VII, Eq. (89), and differ from the trivial unit, $T = 1$, and zero, $R = 0$, matrices due to the distributed mode coupling in the transition layer as is illustrated in Sec. VIII. The similar matrices T^- and R^- describe the integral transmittance and reflectance of the eigenmodes propagating in the $-z$ direction. The matrix $R^-(0)$ or $R^+(L)$ gives the reflectance coefficient R_n^m equal to a relative amplitude of the n eigenmode generated due to reflection from the bottom, $z = 0$, or top, $z = L$, grating boundary by the m eigenmode incident in the $-z$ or $+z$ direction onto the boundary, respectively. The matrix $T^+(0)$ gives the transmittance coefficient T_n^p equal to a relative amplitude of the n eigenmode propagating in the $+z$ direction which is generated at $z = +0$ inside the grating by the plane-wave p harmonic incident from the substrate in the $+z$ direction onto the bottom grating boundary $z = 0$. These matrices of transmittance and reflectance at the sharp boundaries of the grating $T^+(0)$, $R^-(0)$, $R^+(L)$ are similar to the matrices $T^+(L)$, $T^-(0)$, $R^+(0)$ entering Eq. (96) and known from Sec. IV. (They should not be confused with the integral transmittance and reflectance matrices T^\pm , R^\pm , which are not marked by the coordinates of boundaries $z = 0$ or $z = L$, but refer to the entire transition layer.)

The derivation of Eqs. (97) immediately follows from the meaning of their terms and the fact that the boundary conditions of continuity of the tangential components of the electric and magnetic fields are equivalent to the relation between the amplitudes of the incident, transmitted, and reflected modes via the transmittance and reflectance coefficients found in Sec. IV.

A straightforward solution of the linear algebraic equations (97) yields the explicit formulas for the amplitudes of all eigenmodes propagating in the $\pm z$ direction in an inner vicinity of the nonlamellar grating boundaries:

$$\begin{aligned} \mathbf{F}^+(0) &= [1 - R^-(0)(R^+ + T^-M)]^{-1}T^+(0)\mathbf{g}^+(0), \\ \mathbf{F}^-(0) &= (R^+ + T^-M)\mathbf{F}^+(0), \\ \mathbf{F}^+(L) &= (T^+ + R^-M)\mathbf{F}^+(0), \\ \mathbf{F}^-(L) &= M\mathbf{F}^+(0); \quad M = [1 - R^+(L)R^-]^{-1}R^+(L)T^+. \end{aligned} \quad (98)$$

By plugging $\mathbf{F}^-(0)$ and $\mathbf{F}^+(L)$ into Eqs. (96), we get an ultimate result for the overall optical response of the nonlamellar grating. It has a transparent form and requires only

algebraic operations, namely, the multiplication and inversion of the matrices composed of the transmittance and reflectance matrices of the mode coupling both at the sharp boundaries and in the volume of the transition layer. It does not even require any additional integration of the differential equations for the wave propagation in the inhomogeneous transition layer since it already has been fully incorporated into the integral transmittance and reflectance matrices of the entire transition layer T^\pm and R^\pm [see Eq. (90)] via the matrix of the mode-coupling equations (87) in Sec. VII.

Clearly, in a general case, the effect of the distributed mode coupling in the volume of the transition layer of the nonlamellar grating, which manifests itself via the nontrivial integral transmittance and reflectance matrices $T^\pm \neq 1$ and $R^\pm \neq 0$, is as important for the overall optical response of the grating as the usual mode coupling at the sharp boundaries of the grating. In the absence of the distributed mode coupling, when $T^\pm = 1$ and $R^\pm = 0$, the solution for the eigenmode amplitudes in Eq. (98) and, hence, for the overall optical response of the grating in Eq. (96), reduces to a simple form,

$$\begin{aligned} \mathbf{F}^+(0) &= \mathbf{F}^+(L) = [1 - R^-(0)R^+(L)]^{-1}T^+(0)\mathbf{g}^+(0), \\ \mathbf{F}^-(0) &= \mathbf{F}^-(L) = R^+(L)\mathbf{F}^+(0), \end{aligned} \quad (99)$$

that is valid for the homogeneous, lamellar grating.

X. CONCLUSIONS

We presented the explicit analytic solution and a clear physical picture for the structure, propagation, and diffraction of the TE and TM eigenmodes in a lamellar grating, including their geometrical optics and nonadiabatic transformation in the inhomogeneous gratings as well as their reflection and transmission at a grating boundary. In particular, we emphasize the following.

The results are given in a symmetric form that reveals a symmetry of eigenmode structure and includes explicit formulas for the norm of the eigenmode x profile and its derivative, matrices of transformation between the eigenmode and Fourier bases, and the diffraction conversion between the spatial Fourier harmonics of the field due to their propagation through the grating.

We generalized the classical formulas of the reflection and transmission of a plane wave incident onto a plane border between two media to the case of gratings. Namely, we found the reflectance and transmittance coefficients of the eigenmodes and spatial Fourier harmonics at a grating border in the canonical form involving the wave vectors orthogonal to the border, but now these wave vectors are the wave vectors of the eigenmodes and Fourier harmonics. A significant feature of these formulas is that the wave vectors should be multiplied by the appropriate matrices of transformation between the eigenmode and Fourier bases.

We derived the equations for the nonadiabatic mode transformation in the weakly inhomogeneous, nonlamellar gratings and provided the explicit formulas for all coefficients entering these equations. They set the stage for an application of the theory of linear mode coupling [24] to various systems involving optical gratings.

Finally, we derived the system of equations (87) for the amplitudes F_n^\pm of the co- and counterpropagating modes that takes into account exactly the simultaneous mutual transformation and reflection of the eigenmodes within a given subset of eigenmodes and does not assume a weakness of inhomogeneity. Such a truncation of the functional space of the field x profiles provides an efficient method for the analysis of the wave propagation and diffraction in the inhomogeneous, nonlamellar gratings and is convenient for numerical simulations.

A remarkable practical observation regarding the wave propagation and diffraction in all of the weakly inhomogeneous lamellar dielectric gratings we have been dealing with is that a main contribution to the eigenmode reflection and/or transformation, in the case of the propagating or weakly evanescent eigenmodes, comes from the regions surrounding the layers in which either one eigenmode or two neighboring (separated by a narrow band gap) eigenmodes experience a transition through a degeneracy cross section, that is, either (a) $k_{zn} \approx 0$ or (b) $k_{zn} \approx k_{z(n+1)}$ or even (c) $k_{zn} \approx k_{z(n+1)} \approx 0$. Usually there are only these three generic scenarios for substantial linear mode coupling. We illustrated the predicted effects of the eigenmode coupling for these scenarios by their detailed description in the case of a GaAs grating.

Finally, we presented an overall optical response of a generic nonlamellar grating in the transparent form of Eqs. (96) and (98). They explicitly incorporate a combination of the usual mode coupling at the sharp boundaries of the grating (described in Sec. IV) and a less evident mode coupling accumulated over a distributed transition layer of a nonlamellar grating (described in Secs. VI, VII, VIII, and IX). In a general case, both of these mode-coupling effects are important and their interplay can significantly contribute to the real optical response.

The provided formulas constitute a basis for the calculation of the resonant spectral characteristics and other parameters of the cavities formed by gratings. They reveal a simple physical picture and understanding of the generic situations, critical parameters, and main effects of the eigenmode coupling and geometrical optics responsible for the ultimate optical properties of gratings. Such a qualitative, analytic approach is indispensable for the design of new optical systems involving gratings since it could prompt a vision of an optimal set of major parameters of grating without the need to perform a massive, routine numerical search for those. In particular, a nontrivial structure of the dispersion curves of the lower-band eigenmodes in a nonlamellar, inhomogeneous grating, like the one with the intertwining pairs of the propagation constants found in Fig. 15 (cf. Fig. 7), and an understanding of the linear mode coupling effects described above could immediately point to a specific design and parameters of a grating which would be optimal for achieving a particular grating functionality.

The analysis presented in the paper could help to better understand the physical principles and design of various gratings, to solve the difficult problem of finding the analytic solutions for the 2D gratings and 3D photonic crystals, to study the bound states in the continuum in various periodical lamellar structures (widely discussed in the current literature [95–104]), and to develop a theory of mode coupling [24] for the gratings with varying parameters. For comparison,

recall a staircase approximation of various (e.g., sinusoidal or trapezoidal) grating profiles that has been widely used and discussed within the eigenmode technique of numerical simulations of diffraction in the gratings (see, for a review, chapters 7 and 10 in [4] and references therein). It requires explicit calculation of the nonadiabatic mode transformation at the sharp boundaries between a sufficient number of lamellar staircase steps, which depends on the groove profile, and can be done similarly to Sec. IV. This numerical technique has yielded perfect results in the case of TE waves but the results of its application to the case of TM waves have been somewhat controversial.

However, all of the latter problems, including the theory's application to a new, grating-outcoupled surface-emitting laser (GOSEL) design of the semiconductor diode lasers generating mid- or far-IR radiation by means of the intracavity nonlinear mixing [25–36], are beyond the scope of the present paper. We will address them elsewhere.

ACKNOWLEDGMENTS

The work was supported by the Program of Fundamental Studies of the Presidium of the Russian Academy of Science “Nanostructures: Physics, Chemistry, Biology, Basic Technology.” We thank the Robert A. Welch Foundation and King Abdulaziz City for Science and Technology (KACST) for their support.

APPENDIX A: BIORTHOGONALITY OF THE EIGENMODE BASIS FOR AN ABSORPTIVE PERMITTIVITY

The analysis of the eigenmodes presented above and, in particular, all explicit analytic formulas for the TE eigenmodes, Eqs. (4)–(18), and TM eigenmodes, Eqs. (20)–(35), in a lamellar grating are fully valid for both the lossless (real-valued) and absorptive (complex-valued) permittivities ε_q of the grating layers. In both cases, the set of the eigenmode x profiles, $\{f_n(x)|n = 1, 2, \dots\}$ for the TE eigenmodes or $\{\varphi_n(x)|n = 1, 2, \dots\}$ for the TM eigenmodes, forms a complete basis for the expansion of the field in the grating as in Eq. (8) or Eq. (25). The only difference is that in the lossless case this basis is orthogonal, but in the absorptive case it is biorthogonal.

The latter complication stems from the fact that the operators \hat{L} and \hat{L}_M in Eq. (6) or (23) for the TE or TM eigenmodes determined by the corresponding Helmholtz equations (4) or (20) lose their self-adjoint (Hermitian) property in the case of the complex-valued permittivities. It means that the operators \hat{L} and \hat{L}_M do not coincide anymore with their adjoint (Hermitian conjugate) operators \hat{L}^\dagger and \hat{L}_M^\dagger but differ simply by the complex conjugation of the permittivity $\varepsilon(x)$, namely,

$$\hat{L}^\dagger = \frac{d^2}{dx^2} + \frac{\varepsilon^* \omega^2}{c^2}, \quad \hat{L}_M^\dagger = \varepsilon^* \frac{d}{dx} \left(\frac{d}{\varepsilon^* dx} \right) + \frac{\varepsilon^* \omega^2}{c^2}. \quad (\text{A1})$$

Consequently, the eigenvalues of the adjoint operators are equal to the complex conjugate of the eigenmode propagation constants squared, $(k_{zn}^2)^*$, and the eigenfunctions of the adjoint operators, defined as

$$\hat{L}^\dagger f_n^\dagger = (k_{zn}^2)^* f_n^\dagger, \quad \hat{L}_M^\dagger \varphi_n^\dagger = (k_{zn}^2)^* \varphi_n^\dagger, \quad (\text{A2})$$

are given by the same explicit formulas as those for the eigenmodes in Eqs. (9) and (10) or (26) and (27) but with the additional substitute $\varepsilon_q \rightarrow \varepsilon_q^*$, $k_{zn}^2 \rightarrow (k_{zn}^2)^*$, that is,

$$f_n^\dagger(x) \text{ or } \varphi_n^\dagger(x) = b_q^{+\dagger} e^{ik_q^\dagger(x-d_1)} + b_q^{-\dagger} e^{-ik_q^\dagger(x-d_1)}. \quad (\text{A3})$$

Here $k_q^\dagger = [\omega^2 \varepsilon_q^*/c^2 - (k_{zn}^2)^*]^{1/2}$ and the amplitudes are

$$b_1^{\pm\dagger} = \pm \left[\left(-\frac{k_2^\dagger}{\varepsilon_2^{*s}} \pm \frac{k_1^\dagger}{\varepsilon_1^{*s}} \right) e^{ik_2^\dagger d_2 - ik_x \lambda_g/2} + \frac{2k_2^\dagger}{\varepsilon_2^{*s}} e^{\pm ik_1^\dagger d_1 + ik_x \lambda_g/2} - \left(\frac{k_2^\dagger}{\varepsilon_2^{*s}} \pm \frac{k_1^\dagger}{\varepsilon_1^{*s}} \right) e^{-ik_2^\dagger d_2 - ik_x \lambda_g/2} \right],$$

$$b_2^{\pm\dagger} = \pm \left[\left(\frac{k_1^\dagger}{\varepsilon_1^{*s}} \pm \frac{k_2^\dagger}{\varepsilon_2^{*s}} \right) e^{ik_1^\dagger d_1 + ik_x \lambda_g/2} - \frac{2k_1^\dagger}{\varepsilon_1^{*s}} e^{\mp ik_2^\dagger d_2 - ik_x \lambda_g/2} + \left(\frac{k_1^\dagger}{\varepsilon_1^{*s}} \mp \frac{k_2^\dagger}{\varepsilon_2^{*s}} \right) e^{-ik_1^\dagger d_1 + ik_x \lambda_g/2} \right] \quad (\text{A4})$$

with $s = 0$ or 1 for TE or TM eigenmodes, respectively.

This means that the explicit formulas for the complex conjugate of the adjoint eigenmode, $(f_n^\dagger)^*$ or $(\varphi_n^\dagger)^*$, are the same Eqs. (9) and (10) or Eqs. (26) and (27) that describe the original eigenmode f_n or φ_n but with the sign in front of each imaginary unity explicitly entering these formulas being changed to the opposite sign, $i \rightarrow -i$. Hence, in the lossless limit, when the permittivities are real-valued quantities, the adjoint eigenmodes coincide with the original eigenmodes, $f_n^\dagger = f_n$ and $\varphi_n^\dagger = \varphi_n$, and their inner products become the norm squared given by Eqs. (18) and (35), $\langle f_n^\dagger, f_n \rangle = \|f_n\|^2$ and $\langle \varphi_n^\dagger, \varphi_n \rangle_M = \|\varphi_n\|_M^2$.

In the theory of wave propagation, the adjoint eigenmodes are known as the transferring modes [24].

The defining property of an adjoint operator,

$$\langle \hat{L}^\dagger g, f \rangle = \langle g, \hat{L} f \rangle, \quad (\text{A5})$$

written for the TE or TM eigenfunctions as follows,

$$\langle \hat{L}^\dagger f_{n'}^\dagger, f_n \rangle = \langle f_{n'}^\dagger, \hat{L} f_n \rangle, \quad \langle \hat{L}_M^\dagger \varphi_{n'}^\dagger, \varphi_n \rangle = \langle \varphi_{n'}^\dagger, \hat{L}_M \varphi_n \rangle, \quad (\text{A6})$$

gives $(k_{zn'}^2 - k_{zn}^2) \langle f_{n'}^\dagger, f_n \rangle = 0$, $(k_{zn'}^2 - k_{zn}^2) \langle \varphi_{n'}^\dagger, \varphi_n \rangle_M = 0$, which proves the biorthogonality of the eigenmode basis: The eigenmodes of the TE operator \hat{L} (or TM operator \hat{L}_M) and its adjoint operator \hat{L}^\dagger (or TM operator \hat{L}_M^\dagger) corresponding to the different band indexes n and n' are orthogonal and may be made biorthonormal:

$$\frac{\langle f_{n'}^\dagger, f_n \rangle}{\langle f_n^\dagger, f_n \rangle} = \delta_{n',n}, \quad \frac{\langle \varphi_{n'}^\dagger, \varphi_n \rangle_M}{\langle \varphi_n^\dagger, \varphi_n \rangle_M} = \delta_{n',n}. \quad (\text{A7})$$

In terms of such a biorthogonal basis, one has

$$\hat{L} = \sum_n k_{zn}^2 \frac{|f_n\rangle \langle f_n^\dagger|}{\langle f_n^\dagger, f_n \rangle}, \quad \hat{L}_M = \sum_n k_{zn}^2 \frac{|\varphi_n\rangle_{MM} \langle \varphi_n^\dagger|}{\langle \varphi_n^\dagger, \varphi_n \rangle_M}. \quad (\text{A8})$$

Thus, the only major difference of the absorptive case from the lossless one is the necessity to use the biorthonormal basis and Eq. (A7) instead of the orthonormal basis and Eq. (19) or

(37) for the expansion of all field or current x profiles:

$$F(x) = \sum_{n=1}^{\infty} \frac{\langle f_n^\dagger, F \rangle}{\langle f_n^\dagger, f_n \rangle} f_n(x), \quad \Phi(x) = \sum_{n=1}^{\infty} \frac{\langle \varphi_n^\dagger, F \rangle_M}{\langle \varphi_n^\dagger, \varphi_n \rangle_M} \varphi_n(x). \quad (\text{A9})$$

The form of the characteristic equation (14) or (31) for the absorptive grating is exactly the same as that for the lossless grating. However, its solutions for the TE- or TM-eigenmode propagation constants squared, k_{zn}^2 , are not real-valued anymore (as they were in Figs. 3 and 4), but acquire the complex values for the complex-valued permittivities $\varepsilon_{1,2}$. The methods for the solution of the explicit transcendental characteristic equation (14) or (31) have been studied in numerous works since the pioneering work of Kronig and Penney [88] and are well known. In particular, in the optics of absorptive lamellar gratings, two efficient numerical methods, which are different but both systematically find all eigenvalues in a prescribed region of the complex plane, have been developed in [105–107]. Modern general-purpose software, like MATHEMATICA, allows one to find and plot the graphs of these eigenvalues (such as Figs. 3 and 4) in practically no time. For most dielectric gratings with relatively small absorption, like GaAs ones, the complex eigenvalues may be found analytically by means of the perturbation theory with the real eigenvalues taken as the zeroth-order approximation.

As expected, it may be shown [11] that all eigenvalues k_{zn}^2 may be found by following their continuous paths in the complex plane with the increase of the imaginary parts $\text{Im}\varepsilon_{1,2}$ of the permittivities, starting from the real eigenvalues occurring at the zero imaginary parts.

APPENDIX B: DERIVATION OF THE MODE-COUPLING EQUATIONS (86) AND (87)

We start with the TE-wave case and represent the electric field E_y inside a grating as a superposition of the eigenmodes $f_m(x, z)$ with amplitudes $F_m(z)$ as per Eq. (85). Then we plug this ansatz into the homogeneous version of Maxwell equation (4), i.e., the Helmholtz equation

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + \frac{\omega^2}{c^2} \varepsilon E_y = 0, \quad (\text{B1})$$

and employ the TE-eigenmode equation

$$\frac{d^2 f_m}{dx^2} + \varepsilon \frac{\omega^2}{c^2} f_m = k_{zm}^2 f_m \quad (\text{B2})$$

from Sec. II A as well as an explicit formula for a derivative of the geometrical optics factor (71) from Sec. V, $d\Phi_m/dz = -\psi_{mm} \Phi_m$, in the following derivatives:

$$\frac{\partial}{\partial z} \Phi_m f_m = -\psi_{mm} \Phi_m f_m + \Phi_m \frac{\partial f_m}{\partial z},$$

$$\frac{\partial^2}{\partial z^2} \Phi_m f_m = \Phi_m \left[\psi_{mm}^2 - \frac{d\psi_{mm}}{dz} - 2\psi_{mm} \frac{\partial}{\partial z} + \frac{\partial^2}{\partial z^2} \right] f_m. \quad (\text{B3})$$

In this way, the Maxwell equation (B1) acquires the form

$$\sum_m \Phi_m \left\{ f_m \left[\frac{d^2 F_m}{dz^2} + \kappa_{zm}^2 F_m \right] + 2 \frac{dF_m}{dz} \left[\frac{\partial f_m}{\partial z} - \psi_{mm} f_m \right] + F_m \left[\frac{\partial^2 f_m}{\partial z^2} - 2\psi_{mm} \frac{\partial f_m}{\partial z} \right] \right\} = 0, \quad (\text{B4})$$

where $\kappa_{zm}^2 = k_{zm}^2 + \psi_{mm}^2 - \frac{d\psi_{mm}}{dz}$.

Next, we project Eq. (B4) onto each n -eigenmode x profile by calculating a scalar product $\langle f_n^\dagger, \dots \rangle$ and using the biorthogonality of the eigenfunction basis, Eq. (A7). This path immediately leads us to the system of coupled second-order differential equations (86).

Now we rewrite Eq. (86) in terms of the amplitudes (76) of the counterpropagating eigenmodes F_n^\pm and their first derivatives. In order to accomplish this goal, we first calculate the second derivative of the amplitude F_n by taking a derivative of the second equation in Eqs. (76),

$$\frac{d^2 F_n}{dz^2} = \frac{\tilde{k}'_{zn}}{2\tilde{k}_{zn}} \frac{dF_n}{dz} - \tilde{k}_{zn}^2 F_n + i\tilde{k}_{zn}^{1/2} \left[e^{i\tilde{\phi}_n} \frac{dF_n^+}{dz} - e^{-i\tilde{\phi}_n} \frac{dF_n^-}{dz} \right], \quad (\text{B5})$$

where $\tilde{k}'_{zn} = d\tilde{k}_{zn}/dz$, plug F_n and $\frac{dF_n}{dz}$ from Eqs. (76) into its right-hand side, and then plug the result for $\frac{d^2 F_n}{dz^2}$ into Eq. (86). The result is the following equation:

$$\begin{aligned} & e^{i\tilde{\phi}_n} \frac{dF_n^+}{dz} - e^{-i\tilde{\phi}_n} \frac{dF_n^-}{dz} \\ &= \frac{\tilde{k}'_{zn}}{2\tilde{k}_{zn}} [e^{-i\tilde{\phi}_n} F_n^- - e^{i\tilde{\phi}_n} F_n^+] \\ &+ \sum_{m \neq n} \frac{\Phi_m}{\Phi_n} \left[2\psi_{nm} \frac{\tilde{k}'_{zm}}{\tilde{k}_{zn}^{1/2}} (e^{-i\tilde{\phi}_m} F_m^- - e^{i\tilde{\phi}_m} F_m^+) \right. \\ &\left. + i \frac{\eta_{nm} - 2\psi_{nm}\psi_{nm}}{(\tilde{k}_{zn}\tilde{k}_{zm})^{1/2}} (e^{-i\tilde{\phi}_m} F_m^- + e^{i\tilde{\phi}_m} F_m^+) \right]. \quad (\text{B6}) \end{aligned}$$

Second, we calculate $\frac{dF_n}{dz}$ by taking the derivative of the first line in Eqs. (76) and plug it into the second equation of Eqs. (76), which gives the other equation,

$$e^{i\tilde{\phi}_n} \frac{dF_n^+}{dz} + e^{-i\tilde{\phi}_n} \frac{dF_n^-}{dz} = \frac{\tilde{k}'_{zn}}{2\tilde{k}_{zn}} [e^{-i\tilde{\phi}_n} F_n^- + e^{i\tilde{\phi}_n} F_n^+]. \quad (\text{B7})$$

Finally, we solve the system of the two linear algebraic equations (B6) and (B7) for the first derivatives of the amplitudes of the counterpropagating eigenmodes $\frac{dF_n^+}{dz}$ and $\frac{dF_n^-}{dz}$. As a result, we obtain a system of first-order differential equations describing the linear mode coupling of the eigenmodes in the form stated in Eq. (87).

The derivation of the mode-coupling equations (86) and (87) in the TM-wave case is very similar, with just a slight modification due to a presence of an extra factor $1/\varepsilon$ in the TM-wave Maxwell equation (20) as compared to the TE-wave Maxwell equation (4). Again, we represent the magnetic field H_y inside a grating as a superposition of the eigenmodes

$\varphi_m(x, z)$ with amplitudes $F_m(z)$ as per Eq. (85). Then we plug this ansatz into the homogeneous version of Maxwell equation (20), i.e., the Helmholtz equation

$$\frac{\partial}{\partial z} \left[\frac{\partial H_y}{\varepsilon \partial z} \right] + \frac{\partial}{\partial x} \left[\frac{\partial H_y}{\varepsilon \partial x} \right] + \frac{\omega^2}{c^2} H_y = 0, \quad (\text{B8})$$

and, using the TM-eigenmode equation

$$\varepsilon \frac{d}{dx} \left(\frac{1}{\varepsilon} \frac{d\varphi_m}{dx} \right) + \frac{\varepsilon \omega^2}{c^2} \varphi_m = k_{zm}^2 \varphi_m \quad (\text{B9})$$

from Sec. II B, rewrite it in the following form:

$$\frac{1}{\varepsilon} \frac{\partial^2 H_y}{\partial z^2} + \frac{\partial(1/\varepsilon)}{\partial z} \frac{\partial H_y}{\partial z} + \sum_m \frac{k_{zm}^2}{\varepsilon} F_m \Phi_m \varphi_m = 0. \quad (\text{B10})$$

For the eigenmode superposition in Eq. (85), we reduce this TM-wave Maxwell equation, by means of the analog of Eqs. (B3) with the substitution $f_m \rightarrow \varphi_m$, to the following form, which is similar to the TE-wave equation (B4):

$$\begin{aligned} & \sum_m \frac{\Phi_m}{\varepsilon} \left\{ \varphi_m \left[\frac{d^2 F_m}{dz^2} + \kappa_{zm}^2 F_m \right] \right. \\ &+ \frac{dF_m}{dz} \left[2 \left(\frac{\partial \varphi_m}{\partial z} - \psi_{mm} \varphi_m \right) + \varepsilon \frac{\partial(1/\varepsilon)}{\partial z} \varphi_m \right] \\ &\left. + F_m \left[\varepsilon \frac{\partial}{\partial z} \left(\frac{1}{\varepsilon} \frac{\partial \varphi_m}{\partial z} \right) - 2\psi_{mm} \left(\frac{\partial \varphi_m}{\partial z} - \frac{\varphi_m}{2\varepsilon} \frac{\partial \varepsilon}{\partial z} \right) \right] \right\} = 0, \quad (\text{B11}) \end{aligned}$$

where again $\kappa_{zm}^2 = k_{zm}^2 + \psi_{mm}^2 - \frac{d\psi_{mm}}{dz}$.

Then, we project Eq. (B10) onto each n -eigenmode x profile by calculating a scalar product $\langle \varphi_n^\dagger, \dots \rangle_M$ and using the biorthogonality of the eigenfunction basis, Eq. (A7). Surely, for the TM-wave case as opposed to the TE-wave case, we now have to employ the modified inner product in Eq. (33) as is explained in Sec. II B. Remarkably, this path immediately leads us to the exact same system of the coupled second-order differential Eqs. (86) as the one obtained for the TE-wave case. Of course, the parameters ψ_{nm} , η_{nm} , and \tilde{k}_{zn}^2 defined in Eqs. (71) and (75) are different for the TE and TM polarizations.

This important result on the TE and TM universality of the equations for the eigenmode amplitudes means that the subsequent derivation of the mode-coupling equations (87) for the TM polarization is identical to the derivation described above in Eqs. (B5), (B6), and (B7) for the TE polarization. This completes the derivation of the TM-wave equations (86) and (87).

- [1] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).
- [2] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart, and Winston, New York, 1976).
- [3] P. Marcos and C. M. Soukoulis, *Wave Propagation: From Electrons to Photonic Crystals and Left-Handed Materials* (Princeton University Press, Princeton, 2008).

- [4] E. Popov (ed.), *Gratings: Theory and Numeric Applications* (Institut Fresnel, CNRS, AMU, 2012).
- [5] Y. K. Sirenko and S. Strom, editors, *Modern Theory of Gratings* (Springer, New York, 2010).
- [6] R. Petit (ed.), *Electromagnetic Theory of Gratings*, Topics in Current Physics, Vol. 22 (Springer-Verlag, Berlin, Heidelberg, 1980).
- [7] M. G. Moharam, E. B. Grann, D. A. Pommet, and T. K. Gaylord, *J. Opt. Soc. Am. A* **12**, 1068 (1995).

- [8] G. Granet and B. Guizal, *J. Opt. Soc. Am. A* **13**, 1019 (1996).
- [9] P. Lalanne and G. M. Morris, *J. Opt. Soc. Am. A* **13**, 779 (1996).
- [10] I. C. Botten, M. S. Craig, R. C. McPhedran, J. L. Adams, and J. R. Andrewartha, *Opt. Acta* **28**, 413 (1981).
- [11] I. C. Botten, M. S. Craig, R. C. McPhedran, J. L. Adams, and J. R. Andrewartha, *Opt. Acta* **28**, 1087 (1981).
- [12] I. C. Botten, M. S. Craig, and R. C. McPhedran, *Opt. Acta* **28**, 1103 (1981).
- [13] L. Li, *J. Mod. Optics* **40**, 553 (1993).
- [14] L. Li, *J. Opt. Soc. Am. A* **14**, 2758 (1997).
- [15] L. Novotny and B. Hecht, *Principles of Nano-Optics*, 2nd ed. (Cambridge University Press, New York, 2012).
- [16] K. Buscha, G. von Freymann, S. Lindenb, S. F. Mingaleeva, L. Tkeshelashvilia, and M. Wegener, *Phys. Rep.* **444**, 101 (2007).
- [17] K. Sakoda, *Optical Properties of Photonic Crystals* (Springer, Berlin, 2005).
- [18] D. M. Whittaker and I. S. Culshaw, *Phys. Rev. B* **60**, 2610 (1999).
- [19] S. G. Tikhodeev, A. L. Yablonskii, E. A. Muljarov, N. A. Gippius, and T. Ishihara, *Phys. Rev. B* **66**, 045102 (2002).
- [20] M. I. Hussein, *Proc. R. Soc. A* **465**, 2825 (2009).
- [21] G. A. Evans and J. M. Hammer (eds.), *Surface Emitting Semiconductor Lasers and Arrays* (Academic Press, Boston, 1993).
- [22] Y. Jin, L. Gao, J. Chen, C. Wu, J. L. Reno, and S. Kumar, *Nat. Commun.* **9**, 1407 (2018).
- [23] C. Sirtori, S. Barbieri, and R. Collombelli, *Nat. Photon.* **7**, 691 (2013).
- [24] V. V. Zheleznyakov, V. V. Kocharovskiy, and V. V. Kocharovskiy, *Sov. Phys. Usp.* **26**, 877 (1983).
- [25] A. A. Belyanin, F. Capasso, V. V. Kocharovskiy, V. V. Kocharovskiy, and M. O. Scully, *Phys. Rev. A* **63**, 053803 (2001).
- [26] A. A. Belyanin, V. V. Kocharovskiy, V. V. Kocharovskiy, and M. O. Scully, *Phys. Rev. A* **65**, 053824 (2002).
- [27] N. Owschimikow, C. Gmachl, A. Belyanin, V. Kocharovskiy, D. L. Sivco, R. Colombelli, F. Capasso, and A. Y. Cho, *Phys. Rev. Lett.* **90**, 043902 (2003).
- [28] S. M. Nekorkin *et al.*, *Appl. Phys. Lett.* **90**, 171106 (2007).
- [29] B. N. Zvonkov *et al.*, *Appl. Phys. Lett.* **92**, 021122 (2008).
- [30] M. A. Belkin *et al.*, *Nat. Photon.* **1**, 288 (2007).
- [31] M. A. Belkin *et al.*, *Appl. Phys. Lett.* **92**, 201101 (2008).
- [32] C. Pflugl *et al.*, *Appl. Phys. Lett.* **93**, 161110 (2008).
- [33] S. Jung *et al.*, *IEEE J. Sel. Top. Quantum Electron.* **21**, 1200710 (2015).
- [34] M. A. Belkin and F. Capasso, *Phys. Scr.* **90**, 118002 (2015).
- [35] K. Fujita *et al.*, *Opt. Express* **24**, 16357 (2016).
- [36] K. Fujita *et al.*, *Appl. Phys. Express* **10**, 082102 (2017).
- [37] J. N. Winn, Y. Fink, S. Fan, and J. D. Joannopoulos, *Opt. Lett.* **23**, 1573 (1998).
- [38] Y. Fink *et al.*, *Science* **282**, 1679 (1998).
- [39] C. Hooijer, D. Lenstra, and A. Lagendijk, *Opt. Lett.* **25**, 1666 (2000).
- [40] G. P. Agrawal, *Nonlinear Fiber Optics* (Academic Press, London, 2001).
- [41] G. I. Stegeman and M. Segev, *Science* **286**, 1518 (1999).
- [42] D. M. Christodoulidis, F. Lederer, and Y. Silberberg, *Nature (London)* **424**, 817 (2003).
- [43] T. Pertsch, T. Zentgraf, U. Peschel, A. Brauer, and F. Lederer, *Phys. Rev. Lett.* **88**, 093901 (2002).
- [44] L. J. Wu, M. Mazilu, T. Karle, and T. F. Krauss, *IEEE J. Quantum Electron.* **38**, 915 (2002).
- [45] T. F. Krauss, *Phys. Status Solidi A* **197**, 688 (2003).
- [46] S. Noda, A. Chutinan, and M. Imada, *Nature (London)* **407**, 608 (2000).
- [47] S. Ogawa *et al.*, *Science* **305**, 227 (2004).
- [48] D. H. Wu and M. Razeghi, *APL Mater.* **5**, 035505 (2017).
- [49] G. Liang, T. Liu, and Q. J. Wang, *IEEE J. Sel. Top. Quantum Electron.* **23**, 1200118 (2017).
- [50] C. Sigler *et al.*, *Appl. Phys. Lett.* **104**, 131108 (2014).
- [51] C. Boyle *et al.*, *Appl. Phys. Lett.* **108**, 121107 (2016).
- [52] G. Xu *et al.*, *Nat. Commun.* **3**, 952 (2012).
- [53] L. Mahler *et al.*, *Appl. Phys. Lett.* **96**, 191109 (2010).
- [54] R. J. E. Taylor *et al.*, *IEEE J. Sel. Top. Quantum Electron.* **21**, 4900307 (2015).
- [55] K. Hirose *et al.*, *Nat. Photon.* **8**, 406 (2014).
- [56] S.-L. Chua *et al.*, *Opt. Lett.* **39**, 2072 (2014).
- [57] E. Miyai *et al.*, *Nature (London)* **441**, 946 (2006).
- [58] E. Miyai and S. Noda, *Appl. Phys. Lett.* **86**, 111113 (2005).
- [59] D. Ohnishi, T. Okano, M. Imada, and S. Noda, *Opt. Express* **12**, 1562 (2004).
- [60] M. Imada *et al.*, *Appl. Phys. Lett.* **75**, 316 (1999).
- [61] J. Liu *et al.*, in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (online) (Optical Society of America, 2018), https://doi.org/10.1364/CLEO_SI.2018.SW3Q.5.
- [62] D. Zhao *et al.*, *Sci. Rep.* **6**, 18860 (2016).
- [63] Y. Kurosaka *et al.*, *Nat. Photon.* **4**, 447 (2010).
- [64] Y. Chassagneux *et al.*, *Nature (London)* **457**, 174 (2009).
- [65] H. Matsubara *et al.*, *Science* **319**, 445 (2008).
- [66] X. K. Sun and A. Yariv, *Opt. Express* **16**, 9155 (2008).
- [67] M. Imada, A. Chutinan, S. Noda, and M. Mochizuki, *Phys. Rev. B* **65**, 195306 (2002).
- [68] D. Zhou *et al.*, in *Proceedings of SPIE*, edited by C. Lei and K. D. Choquette (SPIE, Bellingham, WA, 2015), Vol. 9381, p. 93810B.
- [69] D. Hofstetter, J. Faist, M. Beck, and U. Oesterle, *Appl. Phys. Lett.* **75**, 3769 (1999).
- [70] W. Schrenk *et al.*, *Appl. Phys. Lett.* **77**, 2086 (2000).
- [71] O. Demichel *et al.*, *Opt. Express* **14**, 5335 (2006).
- [72] J. A. Fan *et al.*, *Opt. Express* **14**, 11672 (2006).
- [73] S. Kumar *et al.*, *Opt. Express* **15**, 113 (2007).
- [74] R. Kazarinov and C. Henry, *IEEE J. Quantum Electron.* **21**, 144 (1985).
- [75] R. J. Noll and S. H. Macomber, *IEEE J. Quantum Electron.* **26**, 456 (1990).
- [76] S. H. Macomber *et al.*, *Proc. SPIE* **3001**, 42 (1997).
- [77] M. Kasraian and D. Botez, *Appl. Phys. Lett.* **69**, 2795 (1996).
- [78] M. Kasraian, J. Lopez, and D. Botez, *IEEE Photonics Technol. Lett.* **10**, 27 (1998).
- [79] S. Li, G. Witjaksono, S. Macomber, and D. Botez, *IEEE J. Sel. Top. Quantum Electron.* **9**, 1153 (2003).
- [80] S. Li and D. Botez, *IEEE J. Quantum Electron.* **43**, 655 (2007).
- [81] A. M. Shams-Zadeh-Amiri, W. Li, and X. Li, *IEEE J. Quantum Electron.* **43**, 31 (2007).
- [82] V. Liu and S. Fan, *Comput. Phys. Commun.* **183**, 2233 (2012).
- [83] A. F. Oskooi *et al.*, *Comput. Phys. Commun.* **181**, 687 (2010).
- [84] C. Wu, S. Khanal, J. L. Reno, and S. Kumar, *Optica* **3**, 734 (2016).

- [85] J. S. Jensen *et al.*, *IEEE Photonics Technol. Lett.* **17**, 1202 (2005).
- [86] M. Burger, S. J. Osher, and E. Yablonovitch, *IEICE Trans. Electron.* **E87-C**, 258 (2004).
- [87] M. P. Bends and O. Sigmund, *Topology Optimization: Theory, Methods, and Applications* (Springer, Berlin, 2003).
- [88] R. Kronig and W. G. Penney, *Proc. R. Soc. A* **130**, 499 (1931).
- [89] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media*, 2nd ed. (Elsevier, Amsterdam, 2004).
- [90] B. Z. Katsenelenbaum, *High-Frequency Electrodynamics* (Wiley-VCH, Weinheim, 2006).
- [91] H. Kogelnik, in *Integrated Optics*, Topics in Applied Physics, Vol. 7, edited by T. Tamir (Springer, Berlin, 1975).
- [92] V. V. Kocharovsky, V. V. Kocharovsky, and S. Tasaki, in *Advances in Chemical Physics. Resonances, Instability, and Irreversibility*, edited by I. Prigogine and Stuart A. Rice (John Wiley & Sons, New York, 1997), Vol. XCIX, p. 333.
- [93] N. A. Sinitsyn and V. Y. Chernyak, *J. Phys. A* **50**, 255203 (2017).
- [94] H. Nakamura, *Nonadiabatic Transition: Concepts, Basic Theories, and Applications*, 2nd ed. (World Scientific, 2012).
- [95] D. V. Evans and R. Porter, *Q. J. Mech. Appl. Math.* **55**, 481 (2002).
- [96] J. Li and N. Engheta, *Phys. Rev. B* **74**, 115125 (2006).
- [97] J. M. Foley, S. M. Young, and J. D. Phillips, *Phys. Rev. B* **89**, 165111 (2014).
- [98] X. Gao *et al.*, *Sci. Rep.* **6**, 31908 (2016).
- [99] X. Cui *et al.*, *Sci. Rep.* **6**, 36066 (2016).
- [100] Z. Sadrieva and A. Bogdanov, *J. Phys.: Conf. Ser.* **741**, 012122 (2016).
- [101] E. N. Bulgakov and A. F. Sadreev, *Phys. Rev. A* **96**, 013841 (2017).
- [102] L. Ni *et al.*, *Opt. Express* **25**, 5580 (2017).
- [103] F. Monticone and A. Alu, *New J. Phys.* **19**, 093011 (2017).
- [104] S.-G. Lee and R. Magnusson, *Phys. Rev. B* **99**, 045304 (2019).
- [105] L. C. Botten, M. S. Craig, and R. C. McPhedran, *Comput. Phys. Commun.* **29**, 245 (1983).
- [106] J. Y. Suratteau, M. Cadilhac, and R. Petit, *J. Optics (Paris)* **14**, 273 (1983).
- [107] G. Tayeb and R. Petit, *Opt. Acta* **31**, 1361 (1984).