Geometrical aspects in optical wave-packet dynamics

Masaru Onoda,[1,2,*] Shuichi Murakami,[2,3,†] and Naoto Nagaosa[1,2,3,‡]
[1]*Correlated Electron Research Center (CERC), National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 4, Tsukuba 305-8562, Japan*
[2]*CREST, Japan Science and Technology Corporation (JST), Saitama, 332-0012, Japan*
[3]*Department of Applied Physics, University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*
(Received 20 June 2006; published 22 December 2006)

We construct a semiclassical theory for propagation of an optical wave packet in a nonconducting medium with a periodic structure of dielectric permittivity and magnetic permeability, i.e., a nonconducting photonic crystal. We employ a quantum-mechanical formalism in order to clarify its link to those of electronic systems. It involves the geometrical phase, i.e., Berry's phase, in a natural way, and describes an interplay between orbital motion and internal rotation. Based on the above theory, we discuss the geometrical aspects of the optical Hall effect. We also consider a reduction of the theory to a system without periodic structure and apply it to the transverse shift of an optical beam at an interface reflection or refraction. For a generic incident beam with an arbitrary polarization, an identical result for the transverse shift of each reflected or transmitted beam is given by the following different approaches: (i) analytic evaluation of wave-packet dynamics, (ii) total angular momentum (TAM) conservation *for individual photons*, and (iii) numerical simulation of wave-packet dynamics. It is consistent with a result by classical electrodynamics. This means that the TAM conservation for individual photons is already taken into account in wave optics, i.e., classical electrodynamics. Finally, we show an application of our theory to a two-dimensional photonic crystal, and propose an optimal design for the enhancement of the optical Hall effect in photonic crystals.

DOI: 10.1103/PhysRevE.74.066610 PACS number(s): 42.15.Eq, 42.70.Qs, 03.65.Vf

I. INTRODUCTION

The geometrical phase known as the Berry phase [1] has been attracting extensive interest in various fields, e.g., optics, molecular physics, nuclear physics, and condensed matter physics [2,3]. In particular, in condensed matter physics, important roles of the geometrical phase in electronic transport phenomena have been intensively studied in the past several years, and great strides has been made both in theoretical and experimental researches. Although a hint of the Berry phase was recognized a long time ago as the anomalous velocity in ferromagnets which leads to the anomalous Hall effect [4,5], it was only after the discovery of the quantum Hall effect that the role of the Berry phase in electron transport began to be recognized. In the quantum Hall system under strong magnetic field, the Hall conductance was related to the topological integer, i.e., the Chern number [6–8]. A recent development is the finding that the Berry phase structure is a fundamental characteristic of Bloch wave functions even in ordinary systems. From this viewpoint, the similarity between the anomalous and quantum Hall effects has been revealed [9–11]. The spin Hall effect based on the geometrical mechanism has also been proposed recently and has opened a new stage of spintronics [12,13]. All of these effects are understood from the concept of the generalized anomalous velocity due to the Berry phase. In other words, a trajectory of an electron is affected by the Berry phase.

In optics, one can also find some phenomena related to the Berry phase. A change of polarization of light during propagation in an inhomogeneous medium has been theoretically studied in the early days [14–16]. This effect has been clearly demonstrated in a helically wound optical fiber and related with the Berry phase [17–19]. Its influence on the trajectory of light has been studied recently by deriving a set of semiclassical equations of motion [20]. From this viewpoint, there are several optical phenomena which are now interpreted as a change of light trajectory due to the Berry phase. One is the transverse shift in reflection and refraction at an interface between two homogeneous media [21–28] (this effect in the case of internal total reflection is called an Imbert-Fedorov shift). The other is a rotation of the beam inside an optical fiber, which is sometimes called an optical Magnus effect [29–31]. These phenomena can be coined the optical Hall effect, because of the similarity to the topological Hall effects [6–13] in electronic systems.

Attribution of these optical phenomena to the Berry phase is not merely a re-interpretation, but also can open a new frontier for novel phenomena. The present authors [20] proposed that in photonic crystals, i.e., systems with periodic structures of dielectric permittivity or magnetic permeability [32,33], the optical Hall effect is enhanced by an order of magnitude more than the above-mentioned examples. It is inspired by its electronic counterpart; the topological Hall effects are known to be enhanced by periodic potentials, particularly when two bands come close in energy. Thus by designing a photonic crystal to have near-degenerate bands, the predicted shift of a light beam is large enough to be observable in experiments [20]. To calculate and design such photonic crystals in a quantitative way, a semiclassical Berry-phase theory applicable to photonic crystals is called for. For this purpose, we take the approach by the variational principle [34–37] in our previous [20] and present papers. This is because it is rather difficult to fully incorporate the

*Electronic address: m.onoda@aist.go.jp
†Electronic address: murakami@appi.t.u-tokyo.ac.jp
‡Electronic address: nagaosa@appi.t.u-tokyo.ac.jp

vectorial nature of electromagnetic waves for generic cases like photonic crystals by other approaches [30,31] which use an eikonal approximation.

In the previous work [20], we have briefly reported the essence of the optical Hall effect and the mechanism of its enhancement in photonic crystals. In the present paper, we construct a semiclassical theory of an optical wave packet (or a photon wave packet) in full detail by keeping its close connection to a theory of an electron wave packet. It incorporates the Berry phase in a natural way. The main focus of the present paper is to present basics of the extended geometrical optics applicable to photonic crystals which are attracting great interest as new optical materials. Photonic crystals can be designed to have desired band structures, with an aid of first-principle numerical calculations, which enables the control of many novel properties of lights [32,33]. To serve this purpose, our theory is presented in a transparent way suitable for the first-principle numerical calculations. As is briefly presented in our previous work [20], the effect of the geometrical phase on an optical wave packet can be incorporated in the same manner as that in electronic systems. These generalized equations of motion correctly describe the interplay between the orbital motion and the internal rotation, e.g., polarization, of a wave packet. The optical Hall effect, i.e., the effect similar to the Hall effects in electronic systems, is derived as a consequence of the equations of motion in photonic systems with periodic structures. Indeed, our equations of motion are analogous to the semiclassical equations of motion for electron wave packets in solids [36,37]. However, the latter basically considered spinless electrons. In the case of optics, the polarization degrees of freedom has to be taken into account and the Berry connection is non-Abelian in general. In this sense, an optical wave packet is more similar to a spinful electron wave packet.

Below, for simplicity, we focus on a light propagating in a nonconducting medium in which there is neither electric nor magnetic order, i.e., the dielectric permittivity $\overleftrightarrow{\epsilon}(\boldsymbol{r})$ and the magnetic permeability $\overleftrightarrow{\mu}(\boldsymbol{r})$ are symmetric tensors. Also their frequency dependences are neglected for simplicity. These conditions ensure the equation of continuity of electromagnetic energy [38] and we can construct the unitary theory of electromagnetic field. Based on this theory, the semiclassical equations of motion can be derived on an equal footing with electronic systems in which the semiclassical equations of motion are derived from quantum mechanics. In order to stress the analogy between electronic and photonic systems, we formulate a theory for Bloch states of electromagnetic field in a quantum-mechanical formalism. Although we focus on the unitary theory in this paper, its extension to a nonunitary version for systems with electric or magnetic order and conducting systems would give some insights to the interesting phenomena and proposals, e.g., the photonic Hall effect in a scattering media subject to an external magnetic field [39,40], the magnetically induced deflection due to the Pitaevskii magnetization [41–46], the one-way waveguide of edge states in magnetic photonic crystals [47,48], and Lorentz force on the light due to the toroidal moment [49].

The reduction to a system without periodic structure is straightforward. Indeed, in the previous work [20], we have presented a simple application of our theory to the transverse shift in the reflection or refraction at an interface, and found that this shift is governed by the conservation of total angular momentum (TAM) *for reflected and refracted photons individually*. We have also numerically demonstrated the validity of our theory for the case of an incident wave packet with circular polarization. However, very recently, the transverse shift evaluated by the conservation of TAM has been questioned [50] for the cases of incident wave packets with elliptic polarizations. In this paper, we present an additional way of estimating the transverse shift from the asymptotic form of a wave packet, and also numerical calculations for generic polarized states. As an important consequence from the study on this issue, we find that an identical result for the shift of each beam is given by (i) analytic evaluation of wave-packet dynamics, (ii) TAM conservation for individual photons in Ref. [20], and (iii) numerically exact simulation of wave-packet dynamics. This agreement in different approaches supports the validity of the present theory claiming that the transverse shift is governed by the conservation of TAM for individual photons. This identical result obtained by the above different approaches is also consistent with that obtained by a more conventional approach based on classical electrodynamics [26,27]. In other words, the TAM conservation for individual photons is already taken into account in wave optics, i.e., classical electrodynamics.

For a broad readership, we divide the main contents into two sections; Sec. II is devoted to formalisms, explaining in full detail the derivations of the theory and the resulting formulas, while in Sec. III we focus on two applications of the theory: the transverse shift in interface reflection or refraction, and the optical Hall effect in a two-dimensional photonic crystal. Readers who are mainly interested in the applications can skip Sec. II and jump to Sec. III. For this purpose we make Sec. III self-consistent.

The plan of this paper is as follows. In Sec. II A, an electromagnetic field in a nonconducting medium is quantized in the Hamilton-Jacobi formalism by introducing the Dirac bracket for the constrained system. Some quantum operators for physical observables are also presented. Eigenstates in a periodic system are discussed in Sec. II B for the application to a photonic crystal. In Sec. II C, we consider a perturbed modulation superimposed on a background periodic structure and discuss corrections for the eigenstates and the expectation values of physical observables for an optical wave packet. The equations of motion are derived taking into account the Berry phase and the perturbed modulation. An application of our theory to reflection or refraction problem at a flat interface is discussed in Sec. III A by reducing the theory to the case without periodic structure. Recently some criticisms are raised against our approach to this reflection or refraction problem [50]. Remarks on these criticisms are presented in Sec. III B. In Sec. III C, we apply our theory to a two-dimensional photonic crystal and present some examples of Berry curvatures and internal rotations.

Section IV is devoted to the discussion on the implications of the present work to a wider range of phenomena in physics. Related previous works are mentioned here.

II. FORMALISMS

A. Electromagnetic field in a nonconducting medium

We consider an electromagnetic field in a nonconducting medium with a generic modulation but without electric or magnetic orders, and begin with the following Lagrangian:

$$L = \frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r},t)\cdot\boldsymbol{D}(\boldsymbol{r},t) - \boldsymbol{H}(\boldsymbol{r},t)\cdot\boldsymbol{B}(\boldsymbol{r},t)], \tag{1}$$

where

$$\boldsymbol{D}(\boldsymbol{r},t) = \overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{E}(\boldsymbol{r},t) = \overleftrightarrow{\epsilon}(\boldsymbol{r})[-\partial_t\boldsymbol{A}(\boldsymbol{r},t) - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi(\boldsymbol{r},t)], \tag{2a}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{H}(\boldsymbol{r},t) = \boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{A}(\boldsymbol{r},t). \tag{2b}$$

We take the unit in which $\hbar = c = 1$ where c is the speed of light in a vacuum. The medium where light propagates is treated as an insulating material, which is characterized by the dielectric permittivity $\overleftrightarrow{\epsilon}(\boldsymbol{r})$ and the magnetic permeability $\overleftrightarrow{\mu}(\boldsymbol{r})$. These are assumed to be locally symmetric and real-valued tensors, and their frequency dependences are neglected. As mentioned in Sec. I, the equation of continuity for the electromagnetic energy holds under these conditions [38]. It should be noted that they are the sufficient conditions but not the necessary conditions. The functional derivatives with respect to $\partial_t\boldsymbol{A}(\boldsymbol{r},t)$ and $\partial_t\phi(\boldsymbol{r},t)$ determine the canonical momenta $\boldsymbol{\pi}(\boldsymbol{r},t)$ for $\boldsymbol{A}(\boldsymbol{r},t)$ and $\pi_\phi(\boldsymbol{r},t)$ for $\phi(\boldsymbol{r},t)$, respectively. The former gives the canonical definition of $\boldsymbol{\pi}(\boldsymbol{r},t)$ as $\boldsymbol{\pi}(\boldsymbol{r},t) = -\boldsymbol{D}(\boldsymbol{r},t)$, while the latter gives the constraint $\pi_\phi(\boldsymbol{r},t) = 0$. Here we introduce the Lagrange multiplier $\lambda_\phi(\boldsymbol{r},t)$ for this constraint, and the Hamiltonian is given by

$$\begin{aligned} H &= \int d\boldsymbol{r}\,\boldsymbol{\pi}(\boldsymbol{r},t)\cdot\partial_t\boldsymbol{A}(\boldsymbol{r},t) - L + \int d\boldsymbol{r}\lambda_\phi(\boldsymbol{r},t)\pi_\phi(\boldsymbol{r},t) \\ &= H_0 + \int d\boldsymbol{r}[-\pi(\boldsymbol{r})\cdot\boldsymbol{\nabla}_{\boldsymbol{r}}\phi(\boldsymbol{r}) + \lambda_\phi(\boldsymbol{r})\pi_\phi(\boldsymbol{r})], \end{aligned} \tag{3a}$$

$$\begin{aligned} H_0 = \frac{1}{2}\int d\boldsymbol{r}\{&\boldsymbol{\pi}(\boldsymbol{r})\overleftrightarrow{\epsilon}^{-1}(\boldsymbol{r})\boldsymbol{\pi}(\boldsymbol{r}) \\ &+ [\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{A}(\boldsymbol{r})]\overleftrightarrow{\mu}^{-1}(\boldsymbol{r})[\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{A}(\boldsymbol{r})]\}. \end{aligned} \tag{3b}$$

In order for the constraint $\pi_\phi \approx 0$ to be consistently satisfied, the following additional constraint is required:

$$\{\pi_\phi(\boldsymbol{r}), H\}_{\mathrm{P}} = -\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{\pi}(\boldsymbol{r}) \approx 0, \tag{4}$$

where $\{\cdots\}_{\mathrm{P}}$ is the Poisson bracket. The symbol "$\approx$" means "weak equality" which is satisfied when all constraints are imposed [51]. When the Poisson brackets among a set of constraints and a Hamiltonian vanish on a constrained subspace, these constraints are called first-class constraints by definition. On the other hand, when the Poisson brackets of these constraints among themselves do not vanish even on the constrained subspace, we call them second-class constraints. In the present case, $\pi_\phi(\boldsymbol{r})$ and $\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{\pi}(\boldsymbol{r})$ commute with each other, and the commutation relation between $\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{\pi}(\boldsymbol{r})$ and the Hamiltonian generates no additional constraint. So the present system has two first-class constraints,

$$\chi_1(\boldsymbol{r}) \equiv \pi_\phi(\boldsymbol{r}) \approx 0, \tag{5a}$$

$$\chi_2(\boldsymbol{r}) \equiv \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{\pi}(\boldsymbol{r}) \approx 0. \tag{5b}$$

In order to make a canonical formalism for such a constrained system, all the first-class constraints are transformed to be second class by introducing gauge fixing conditions. Here we take the following gauge conditions:

$$\chi_3(\boldsymbol{r}) \equiv \phi(\boldsymbol{r}) \approx 0, \tag{6a}$$

$$\chi_4(\boldsymbol{r}) \equiv \boldsymbol{\nabla}_{\boldsymbol{r}}\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{A}(\boldsymbol{r}) \approx 0. \tag{6b}$$

Then the commutation relations between the original constraints and the gauge conditions are represented by

$$\overleftrightarrow{C}(\boldsymbol{r},\boldsymbol{r}') = \{\chi_\alpha(\boldsymbol{r}),\chi_\beta(\boldsymbol{r}')\}_{\mathrm{P}} = \begin{pmatrix} 0 & -\overleftrightarrow{C}(\boldsymbol{r},\boldsymbol{r}') \\ \overleftrightarrow{C}(\boldsymbol{r},\boldsymbol{r}') & 0 \end{pmatrix}, \tag{7a}$$

$$\overleftrightarrow{C}(\boldsymbol{r},\boldsymbol{r}') = \begin{pmatrix} \delta(\boldsymbol{r}-\boldsymbol{r}') & 0 \\ 0 & -\boldsymbol{\nabla}_{\boldsymbol{r}}\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\nabla}_{\boldsymbol{r}}\delta(\boldsymbol{r}-\boldsymbol{r}') \end{pmatrix}. \tag{7b}$$

Introducing Lagrange multipliers for the constraints including the gauge conditions, we redefine the Hamiltonian as

$$H = H_0 + \int d\boldsymbol{r}\boldsymbol{\lambda}(\boldsymbol{r})\cdot\boldsymbol{\chi}(\boldsymbol{r}), \tag{8}$$

where $\boldsymbol{\lambda}(\boldsymbol{r}) = [\lambda_1(\boldsymbol{r}),\lambda_2(\boldsymbol{r}),\lambda_3(\boldsymbol{r}),\lambda_4(\boldsymbol{r})]$ and $\boldsymbol{\chi}(\boldsymbol{r}) = [\chi_1(\boldsymbol{r}),\chi_2(\boldsymbol{r}),\chi_3(\boldsymbol{r}),\chi_4(\boldsymbol{r})]$. The Lagrange multipliers are determined by the conditions $\{\boldsymbol{\chi}(\boldsymbol{r}),H\}_{\mathrm{P}} \approx 0$ and given by

$$\boldsymbol{\lambda}(\boldsymbol{r}) = -\int d\boldsymbol{r}'\overleftrightarrow{C}^{-1}(\boldsymbol{r},\boldsymbol{r}')\{\boldsymbol{\chi}(\boldsymbol{r}'),H_0\}_{\mathrm{P}}, \tag{9}$$

where

$$\overleftrightarrow{C}^{-1}(\boldsymbol{r},\boldsymbol{r}') = \begin{pmatrix} 0 & \overleftrightarrow{C}^{-1}(\boldsymbol{r},\boldsymbol{r}') \\ -\overleftrightarrow{C}^{-1}(\boldsymbol{r},\boldsymbol{r}') & 0 \end{pmatrix}, \tag{10a}$$

$$\overleftrightarrow{C}^{-1}(\boldsymbol{r},\boldsymbol{r}') = \begin{pmatrix} \delta(\boldsymbol{r}-\boldsymbol{r}') & 0 \\ 0 & g(\boldsymbol{r},\boldsymbol{r}') \end{pmatrix}, \tag{10b}$$

and $g(\boldsymbol{r},\boldsymbol{r}')$ satisfies

$$\boldsymbol{\nabla}_{\boldsymbol{r}}\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\nabla}_{\boldsymbol{r}}g(\boldsymbol{r},\boldsymbol{r}') = \boldsymbol{\nabla}_{\boldsymbol{r}'}\overleftrightarrow{\epsilon}^{T}(\boldsymbol{r}')\boldsymbol{\nabla}_{\boldsymbol{r}'}g(\boldsymbol{r},\boldsymbol{r}') = -\delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{11}$$

As a preparation for the quantum theory of this system, we introduce the Dirac bracket defined by

$$\{F,G\}_{\mathrm{D}} = \{F,G\}_{\mathrm{P}} - \int d\boldsymbol{r}d\boldsymbol{r}'\{F,\boldsymbol{\chi}(\boldsymbol{r})\}_{\mathrm{P}}\overleftrightarrow{C}^{-1}(\boldsymbol{r},\boldsymbol{r}')\{\boldsymbol{\chi}(\boldsymbol{r}'),G\}_{\mathrm{P}}. \tag{12}$$

Especially for $\boldsymbol{A}(\boldsymbol{r})$ and $\boldsymbol{\pi}(\boldsymbol{r})$, we obtain the following relation:

$$\{A^i(\boldsymbol{r}),\pi_j(\boldsymbol{r}')\}_{\mathrm{D}} = \delta^i_j\delta(\boldsymbol{r}-\boldsymbol{r}') - \sum_k \nabla^i_{\boldsymbol{r}}\overleftrightarrow{\epsilon}^T_{jk}(\boldsymbol{r}')\nabla^k_{\boldsymbol{r}'}g(\boldsymbol{r},\boldsymbol{r}'), \tag{13}$$

and this leads to the relation between the physical observables,

$$\{B_i(\boldsymbol{r}),D_j(\boldsymbol{r}')\}_{\mathrm{D}} = \sum_k \epsilon_{ijk}\nabla^k_{\boldsymbol{r}}\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{14}$$

where ϵ_{ijk} is the completely antisymmetric tensor defined using $\epsilon_{xyz}=1$. It is noted that this relation is the same as that in the vacuum, while that between $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{H}(\boldsymbol{r})$ is not the case. The equations of motion are derived as

$$\partial_t \boldsymbol{A}(\boldsymbol{r},t) = \{\boldsymbol{A}(\boldsymbol{r},t),H\}_{\mathrm{D}} \approx \overleftrightarrow{\epsilon}^{-1}(\boldsymbol{r})\boldsymbol{\pi}(\boldsymbol{r},t), \tag{15a}$$

$$\partial_t \boldsymbol{\pi}(\boldsymbol{r},t) = \{\boldsymbol{\pi}(\boldsymbol{r},t),H\}_{\mathrm{D}} = -\boldsymbol{\nabla}_{\boldsymbol{r}} \times \{\overleftrightarrow{\mu}^{-1}(\boldsymbol{r})[\boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{A}(\boldsymbol{r},t)]\}. \tag{15b}$$

The above equations are equivalent to the Maxwell equations

$$\partial_t \boldsymbol{D}(\boldsymbol{r},t) = \boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{H}(\boldsymbol{r},t), \tag{16a}$$

$$\partial_t \boldsymbol{B}(\boldsymbol{r},t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{E}(\boldsymbol{r},t), \tag{16b}$$

$$\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \boldsymbol{D}(\boldsymbol{r}) = \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \boldsymbol{B}(\boldsymbol{r}) = 0. \tag{16c}$$

The present system is straightforwardly quantized by the identification as $i\{F,G\}_{\mathrm{D}} \to [F,G]$. Especially, the basic commutation relation, Eq. (14), is quantized as follows:

$$[B_i(\boldsymbol{r}),D_j(\boldsymbol{r}')] = i\sum_k \epsilon_{ijk}\nabla^k_{\boldsymbol{r}}\delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{17}$$

Here we introduce some quantum operators which are useful to check the property of a wave packet. They are the Hamiltonian H, the center of the position $\mathcal{R}$ weighted by energy density, the energy current (the Poynting vector) $\mathcal{P}$, and the rotation of energy current $\mathcal{J}$, which are respectively defined by

$$H = \frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r})\cdot\boldsymbol{D}(\boldsymbol{r}) + \boldsymbol{H}(\boldsymbol{r})\cdot\boldsymbol{B}(\boldsymbol{r})], \tag{18a}$$

$$\mathcal{R} = \frac{1}{2}\int d\boldsymbol{r}\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r})\cdot\boldsymbol{D}(\boldsymbol{r}) + \boldsymbol{H}(\boldsymbol{r})\cdot\boldsymbol{B}(\boldsymbol{r})], \tag{18b}$$

$$\mathcal{P} = \frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r})\times\boldsymbol{H}(\boldsymbol{r}) - \boldsymbol{H}(\boldsymbol{r})\times\boldsymbol{E}(\boldsymbol{r})], \tag{18c}$$

$$\mathcal{J} = \frac{1}{2}\int d\boldsymbol{r}\boldsymbol{r}\times[\boldsymbol{E}(\boldsymbol{r})\times\boldsymbol{H}(\boldsymbol{r}) - \boldsymbol{H}(\boldsymbol{r})\times\boldsymbol{E}(\boldsymbol{r})]. \tag{18d}$$

It should be noted that the last two operators are different from the momentum and angular momentum operators defined by

$$\boldsymbol{P} = \frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{D}(\boldsymbol{r})\times\boldsymbol{B}(\boldsymbol{r}) - \boldsymbol{B}(\boldsymbol{r})\times\boldsymbol{D}(\boldsymbol{r})], \tag{19a}$$

$$\boldsymbol{J} = \frac{1}{2}\int d\boldsymbol{r}\boldsymbol{r}\times[\boldsymbol{D}(\boldsymbol{r})\times\boldsymbol{B}(\boldsymbol{r}) - \boldsymbol{B}(\boldsymbol{r})\times\boldsymbol{D}(\boldsymbol{r})], \tag{19b}$$

while $\mathcal{P}$ and $\mathcal{J}$ are conceptually close to $\boldsymbol{P}$ and $\boldsymbol{J}$, respectively. This is because $\mathcal{P}$ and $\mathcal{J}$ are not necessarily proportional to $\boldsymbol{P}$ and $\boldsymbol{J}$. In other words, $\mathcal{P}$ and $\mathcal{J}$ do not necessarily satisfy the algebra of the momentum and the angular momentum. Therefore in a system with translational and rotational symmetries, what should be conserved are $\boldsymbol{P}$ and $\boldsymbol{J}$, rather than $\mathcal{P}$ and $\mathcal{J}$. Actually it has been experimentally confirmed that $\boldsymbol{J}$ is conserved in a dielectric medium with rotational symmetry [52]. In spite of these shortcomings of $\mathcal{P}$ and $\mathcal{J}$, we focus on $\mathcal{P}$ and $\mathcal{J}$ when a system has no continuous translational nor rotational symmetry. This is because $\mathcal{P}$ and $\mathcal{J}$ have relatively simple expressions even in a periodic system as shown in Appendix C. Especially, a part of $\mathcal{J}$ suggests a close relation between the internal rotation and the Berry curvature in a photonic system, as well as in the quantum Hall system where the internal rotation of an spinless electron is originated by the cyclotron motion [36]. However, when a system has continuous translational and rotational symmetries, we focus on $\boldsymbol{P}$ and $\boldsymbol{J}$. This is the case in the reflection or refraction problem at a flat interface in Sec. III A. The list of physical observables including the above operators both in electronic and photonic systems are given in Table I for comparison.

Finally, it should be noted that the optical Hall effect comes from the particle-wave duality of an optical wave packet and the geometrical or topological property of a wave function. Therefore this effect can be observed in a macroscopic wave packet of light described by classical electrodynamics, when a wave packet under consideration is approximately coherent. In this sense, the second quantization is not always necessary. The second quantization is adopted, for convenience, to calculate a motion of a wave packet on an equal footing with that of an electronic system, as shown in Sec. II C. As long as we consider an approximately coherent wave packet in a single particle approximation of quantum theory of photon or in a linear approximation of classical electrodynamics, results obtained by both formalisms coincide with each other as shown in Sec. III A. Detailed remarks on the relation between quantum and classical pictures of the optical Hall effect are given in Appendix A.

B. Eigenfunctions in a periodic system

Here we introduce eigenfunctions in a periodic system,

$$\boldsymbol{\Phi}^F_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t) = e^{-iE_{n\boldsymbol{k}}t}\boldsymbol{\Phi}^F_{n\lambda\boldsymbol{k}}(\boldsymbol{r}) = \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}-iE_{n\boldsymbol{k}}t}}{\sqrt{2E_{n\boldsymbol{k}}}}\boldsymbol{U}^F_{n\lambda\boldsymbol{k}}(\boldsymbol{r}), \tag{20}$$

where $F=E$ or H. The symbols n, λ, and $\boldsymbol{k}$ represent the band index, the index for degenerate modes in the nth band, and the lattice momentum, respectively, and $E_{n\boldsymbol{k}}$ is the energy eigenvalue of the nth band, which may be degenerate. It should be noted that the band index n is not needed in locally

TABLE I. Operators relevant to wave-packet dynamics. H: Hamiltonian; $\hat{H}(\boldsymbol{r})$: first-quantized Hamiltonian; $\boldsymbol{R}$: position (dipole moment); $\boldsymbol{P}$: (pseudo) momentum; $\boldsymbol{J}$: angular momentum; $\hat{s}$: spin matrix; $\boldsymbol{I}$: charge current; $\boldsymbol{M}$: magnetic moment; $\hat{v}(\boldsymbol{r})$: first-quantized velocity; $\boldsymbol{\mathcal{R}}$: position weighted by energy density; $\boldsymbol{\mathcal{P}}$: energy current; $\boldsymbol{\mathcal{J}}$: rotation of energy current.

	Electronic system	Photonic system
H	$\int d\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})\hat{H}(\boldsymbol{r})\psi(\boldsymbol{r})$	$\frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r})\cdot\boldsymbol{D}(\boldsymbol{r})+\boldsymbol{H}(\boldsymbol{r})\cdot\boldsymbol{B}(\boldsymbol{r})]$
$\boldsymbol{R}$	$\int d\boldsymbol{r}\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})\psi(\boldsymbol{r})$	undefined
$\boldsymbol{P}$	$\int d\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})[-i\boldsymbol{\nabla}_{\boldsymbol{r}}-e\boldsymbol{A}(\boldsymbol{r})+e\boldsymbol{B}\times\boldsymbol{r}]\psi(\boldsymbol{r})$	$\frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{D}(\boldsymbol{r})\times\boldsymbol{B}(\boldsymbol{r})-\boldsymbol{B}(\boldsymbol{r})\times\boldsymbol{D}(\boldsymbol{r})]$
$\boldsymbol{J}$	$\int d\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})[\boldsymbol{r}\times(-i\boldsymbol{\nabla}_{\boldsymbol{r}})+\hat{s}]\psi(\boldsymbol{r})$	$\frac{1}{2}\int d\boldsymbol{r}\boldsymbol{r}\times[\boldsymbol{D}(\boldsymbol{r})\times\boldsymbol{B}(\boldsymbol{r})-\boldsymbol{B}(\boldsymbol{r})\times\boldsymbol{D}(\boldsymbol{r})]$
$\boldsymbol{I}$	$\int d\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})e\hat{v}(\boldsymbol{r})\psi(\boldsymbol{r})$	undefined
$\boldsymbol{M}$	$\int d\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})\left[\frac{e}{2}\boldsymbol{r}\times\hat{v}(\boldsymbol{r})+g\mu_B\hat{s}\right]\psi(\boldsymbol{r})$	undefined
$\boldsymbol{\mathcal{R}}$	$\frac{1}{2}\int d\boldsymbol{r}\psi^{\dagger}(\boldsymbol{r})\{\hat{H}(\boldsymbol{r}),\boldsymbol{r}\}\psi(\boldsymbol{r})$	$\frac{1}{2}\int d\boldsymbol{r}\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r})\cdot\boldsymbol{D}(\boldsymbol{r})+\boldsymbol{H}(\boldsymbol{r})\cdot\boldsymbol{B}(\boldsymbol{r})]$
$\boldsymbol{\mathcal{P}}$	$\frac{i}{2}\int d\boldsymbol{r}[H,\psi^{\dagger}(\boldsymbol{r})\{\hat{H}(\boldsymbol{r}),\boldsymbol{r}\}\psi(\boldsymbol{r})]$	$\frac{1}{2}\int d\boldsymbol{r}[\boldsymbol{E}(\boldsymbol{r})\times\boldsymbol{H}(\boldsymbol{r})-\boldsymbol{H}(\boldsymbol{r})\times\boldsymbol{E}(\boldsymbol{r})]$
$\boldsymbol{\mathcal{J}}$	$\frac{i}{2}\int d\boldsymbol{r}\boldsymbol{r}\times[H,\psi^{\dagger}(\boldsymbol{r})\{\hat{H}(\boldsymbol{r}),\boldsymbol{r}\}\psi(\boldsymbol{r})]$	$\frac{1}{2}\int d\boldsymbol{r}\boldsymbol{r}\times[\boldsymbol{E}(\boldsymbol{r})\times\boldsymbol{H}(\boldsymbol{r})-\boldsymbol{H}(\boldsymbol{r})\times\boldsymbol{E}(\boldsymbol{r})]$

isotropic systems without periodic structure, but we must keep the index λ to distinguish different polarization states. $\boldsymbol{U}^{E}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})$ and $\boldsymbol{U}^{H}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})$ are Bloch functions for electric field and magnetic field, respectively. It should be noted that the lattice momentum $\boldsymbol{k}$ will be restricted to the first Brillouin zone in the rest of this paper. The eigenfunctions satisfy the Maxwell equations

$$\overleftrightarrow{\epsilon}(\boldsymbol{r})\partial_t\boldsymbol{\Phi}^{E}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)=\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^{H}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t), \tag{21a}$$

$$\overleftrightarrow{\mu}(\boldsymbol{r})\partial_t\boldsymbol{\Phi}^{H}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)=-\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^{E}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t), \tag{21b}$$

$$\boldsymbol{\nabla}_{\boldsymbol{r}}\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^{E}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)=\boldsymbol{\nabla}_{\boldsymbol{r}}\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^{H}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)=0, \tag{21c}$$

and they lead to the following eigen equations:

$$\boldsymbol{\nabla}_{\boldsymbol{r}}\times[\overleftrightarrow{\mu}^{-1}(\boldsymbol{r})\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^{E}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})]=\overleftrightarrow{\epsilon}(\boldsymbol{r})E^{2}_{n\boldsymbol{k}}\boldsymbol{\Phi}^{E}_{n\lambda\boldsymbol{k}}(\boldsymbol{r}), \tag{22a}$$

$$\boldsymbol{\nabla}_{\boldsymbol{r}}\times[\overleftrightarrow{\epsilon}^{-1}(\boldsymbol{r})\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^{H}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})]=\overleftrightarrow{\mu}(\boldsymbol{r})E^{2}_{n\boldsymbol{k}}\boldsymbol{\Phi}^{H}_{n\lambda\boldsymbol{k}}(\boldsymbol{r}). \tag{22b}$$

In the case of $\overleftrightarrow{\epsilon}^{T}(\boldsymbol{r})=\overleftrightarrow{\epsilon}(\boldsymbol{r})$ and $\overleftrightarrow{\mu}^{T}(\boldsymbol{r})=\overleftrightarrow{\mu}(\boldsymbol{r})$, we can orthonormalize the Bloch functions with the same lattice momentum $\boldsymbol{k}$ as

$$\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}\boldsymbol{U}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{U}^{E}_{n'\lambda'\boldsymbol{k}}(\boldsymbol{r})=\delta_{nn'}\delta_{\lambda\lambda'}, \tag{23a}$$

$$\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}\boldsymbol{U}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{U}^{H}_{n'\lambda'\boldsymbol{k}}(\boldsymbol{r})=\delta_{nn'}\delta_{\lambda\lambda'}, \tag{23b}$$

where the domain of integration is the unit cell with the volume v_{WS}. The orthonormality for the eigenfunctions will be discussed later.

We introduce the Fourier transformation,

$$\boldsymbol{U}^{F}_{n\lambda\boldsymbol{k}}(\boldsymbol{G})=\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}e^{-i\boldsymbol{G}\cdot\boldsymbol{r}}\boldsymbol{U}^{F}_{n\lambda\boldsymbol{k}}(\boldsymbol{r}), \tag{24a}$$

$$\overleftrightarrow{\epsilon}(\boldsymbol{G},\boldsymbol{G}')=\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}e^{-i(\boldsymbol{G}-\boldsymbol{G}')\cdot\boldsymbol{r}}\overleftrightarrow{\epsilon}(\boldsymbol{r}), \tag{24b}$$

$$\overleftrightarrow{\mu}(\boldsymbol{G},\boldsymbol{G}')=\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}e^{-i(\boldsymbol{G}-\boldsymbol{G}')\cdot\boldsymbol{r}}\overleftrightarrow{\mu}(\boldsymbol{r}), \tag{24c}$$

where $F=E$ or H. $\boldsymbol{G}$ represents a reciprocal lattice vector. In terms of the above representations in the Fourier space, we introduce the following compact notations for the latter convenience:

$$|U\rangle=[\boldsymbol{U}(\boldsymbol{G}_0),\boldsymbol{U}(\boldsymbol{G}_1),\boldsymbol{U}(\boldsymbol{G}_2),\cdots], \tag{25}$$

and the tensors,

$$\boldsymbol{P}_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}')=(\boldsymbol{k}+\boldsymbol{G})\delta(\boldsymbol{G},\boldsymbol{G}')=\boldsymbol{K}\delta(\boldsymbol{G},\boldsymbol{G}'), \tag{26a}$$

$$[S_i]_{jk}(\boldsymbol{G},\boldsymbol{G}')=-i\epsilon_{ijk}\delta(\boldsymbol{G},\boldsymbol{G}'), \tag{26b}$$

$$\Xi^{E}_{\boldsymbol{k}}=\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\overleftrightarrow{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}, \tag{26c}$$

$$\Xi^{H}_{\boldsymbol{k}}=\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\overleftrightarrow{\epsilon}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}, \tag{26d}$$

where we have introduced the abbreviation $\boldsymbol{K}=\boldsymbol{k}+\boldsymbol{G}$. The inner product of the Bloch functions and the algebra of the above tensors are represented as

$$\langle U|V\rangle=\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}\boldsymbol{U}^{*}(\boldsymbol{r})\cdot\boldsymbol{V}(\boldsymbol{r}), \tag{27a}$$

$$\langle U|i\boldsymbol{S}|V\rangle=\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}\boldsymbol{U}^{*}(\boldsymbol{r})\times\boldsymbol{V}(\boldsymbol{r}), \tag{27b}$$

$$-i\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}|U\rangle=[\boldsymbol{K}_0\times\boldsymbol{U}(\boldsymbol{G}_0),\boldsymbol{K}_1\times\boldsymbol{U}(\boldsymbol{G}_1),\cdots]. \tag{27c}$$

Thus the orthonormality is rewritten as,

$$\langle\boldsymbol{U}^{E}_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|\boldsymbol{U}^{E}_{n'\lambda'\boldsymbol{k}}\rangle=\delta_{nn'}\delta_{\lambda\lambda'}, \tag{28a}$$

$$\langle\boldsymbol{U}^{H}_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\mu}|\boldsymbol{U}^{H}_{n'\lambda'\boldsymbol{k}}\rangle=\delta_{nn'}\delta_{\lambda\lambda'}. \tag{28b}$$

By this notation convention, the Maxwell equations for the Bloch functions are represented in the following compact forms:

$$\vec{\epsilon} E_{nk} |U^E_{n\lambda k}\rangle = i \boldsymbol{P}_{\boldsymbol{k}} \cdot \boldsymbol{S} |U^H_{n\lambda k}\rangle, \tag{29a}$$

$$\vec{\mu} E_{nk} |U^H_{n\lambda k}\rangle = -i \boldsymbol{P}_{\boldsymbol{k}} \cdot \boldsymbol{S} |U^E_{n\lambda k}\rangle, \tag{29b}$$

$$\langle K | \vec{\epsilon} | U^E_{n\lambda k}\rangle = \langle K | \vec{\mu} | U^H_{n\lambda k}\rangle = 0, \tag{29c}$$

where

$$|K\rangle = [0, \cdots, 0, \boldsymbol{k} + \boldsymbol{G}, 0, \cdots]. \tag{30}$$

From Eqs. (29a) and (29b), we can easily derive the following equations:

$$\Xi^E_{\boldsymbol{k}} |U^E_{n\lambda k}\rangle = \vec{\epsilon} E^2_{nk} |U^E_{n\lambda k}\rangle, \tag{31a}$$

$$\Xi^H_{\boldsymbol{k}} |U^H_{n\lambda k}\rangle = \vec{\mu} E^2_{nk} |U^H_{n\lambda k}\rangle. \tag{31b}$$

In relativistic systems, the orthonormality for the eigenfunctions are conventionally represented in terms of the inner product defined by

$$(f|g) = i \int d\boldsymbol{r} [\boldsymbol{f}^*(\boldsymbol{r},t) \cdot [\partial_t \boldsymbol{g}(\boldsymbol{r},t)] - [\partial_t \boldsymbol{f}^*(\boldsymbol{r},t)] \cdot \boldsymbol{g}(\boldsymbol{r},t)], \tag{32}$$

and we obtain the following orthonormality relation as shown in Appendix B:

$$(\Phi^E_{n\lambda k} | \vec{\epsilon} | \Phi^E_{n'\lambda' k'}) = \delta_{nn'} \delta_{\lambda\lambda'} \tilde{\delta}(\boldsymbol{k} - \boldsymbol{k}'), \tag{33a}$$

$$(\Phi^H_{n\lambda k} | \vec{\mu} | \Phi^H_{n'\lambda' k'}) = \delta_{nn'} \delta_{\lambda\lambda'} \tilde{\delta}(\boldsymbol{k} - \boldsymbol{k}'), \tag{33b}$$

$$(\Phi^{E*}_{n\lambda k} | \vec{\epsilon} | \Phi^E_{n'\lambda' k'}) = (\Phi^{H*}_{n\lambda k} | \vec{\mu} | \Phi^H_{n'\lambda' k'}) = 0. \tag{33c}$$

where $\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}') = (2\pi)^3 \delta(\boldsymbol{k}-\boldsymbol{k}')$. It should be noted that we have used $\vec{\epsilon}^T(\boldsymbol{r}) = \vec{\epsilon}(\boldsymbol{r})$ and $\vec{\mu}^T(\boldsymbol{r}) = \vec{\mu}(\boldsymbol{r})$ in the derivation of the above relations. This orthonormality is required to expand the electric and magnetic fields in terms of the eigenfunctions as follows:

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_{n,\lambda} \int_{\mathrm{BZ}} d\boldsymbol{k} E_{nk} [\boldsymbol{\Phi}^E_{n\lambda k}(\boldsymbol{r},t) a_{n\lambda k} + \boldsymbol{\Phi}^{E*}_{n\lambda k}(\boldsymbol{r},t) a^\dagger_{n\lambda k}], \tag{34a}$$

$$\boldsymbol{H}(\boldsymbol{r},t) = \sum_{n,\lambda} \int_{\mathrm{BZ}} d\boldsymbol{k} E_{nk} [\boldsymbol{\Phi}^H_{n\lambda k}(\boldsymbol{r},t) a_{n\lambda k} + \boldsymbol{\Phi}^{H*}_{n\lambda k}(\boldsymbol{r},t) a^\dagger_{n\lambda k}], \tag{34b}$$

where the $\boldsymbol{k}$ integration is over the first Brillouin zone, i.e.,

$$\int_{\mathrm{BZ}} d\boldsymbol{k} = \int_{\boldsymbol{k} \in 1\mathrm{stBZ}} \frac{d\boldsymbol{k}}{(2\pi)^3}. \tag{35}$$

The operators $a_{n\lambda k}$ and $a^\dagger_{n\lambda k}$ are defined by

$$a_{n\lambda k} = \frac{1}{E_{nk}} (\Phi^E_{n\lambda k} | \vec{\epsilon} | E) = \frac{1}{E_{nk}} (\Phi^H_{n\lambda k} | \vec{\mu} | H) = \int d\boldsymbol{r} [\boldsymbol{\Phi}^{E*}_{n\lambda k}(\boldsymbol{r}) \cdot \boldsymbol{D}(\boldsymbol{r}) + \boldsymbol{\Phi}^{H*}_{n\lambda k}(\boldsymbol{r}) \cdot \boldsymbol{B}(\boldsymbol{r})], \tag{36a}$$

$$a^\dagger_{n\lambda k} = \frac{1}{E_{nk}} (E | \vec{\epsilon} | \Phi^E_{n\lambda k}) = \frac{1}{E_{nk}} (H | \vec{\mu} | \Phi^H_{n\lambda k}) = \int d\boldsymbol{r} [\boldsymbol{D}(\boldsymbol{r}) \cdot \boldsymbol{\Phi}^E_{n\lambda k}(\boldsymbol{r}) + \boldsymbol{B}(\boldsymbol{r}) \cdot \boldsymbol{\Phi}^H_{n\lambda k}(\boldsymbol{r})]. , \tag{36b}$$

By using Eq. (17), the following commutation relation is obtained:

$$[a_{n\lambda k}, a^\dagger_{n'\lambda' k'}] = \delta_{nn'} \delta_{\lambda\lambda'} \tilde{\delta}(\boldsymbol{k} - \boldsymbol{k}'). \tag{37}$$

C. Equations of motion

In order to see the effect of geometrical phase on the trajectory of a wave packet, a driving force is needed [53]. This is because the geometrical effect is given by the vector product between the driving force and the Berry curvature as we shall see later. In an electronic system, a driving force is most conventionally produced by the gradient of electric potential. The counterpart in a photonic system is given by the gradient of $\vec{\epsilon}(\boldsymbol{r})$ or $\vec{\mu}(\boldsymbol{r})$. Of course, a periodic structure itself gives a gradient. However, this effect is exactly taken into account by considering optical Bloch states as shown in Appendix C where readers can find details about the basic features of an optical wave packet in a periodic system. Thus we regard the deviation from a periodic structure as a driving force for the Bloch states and treat it perturbatively. Here, the perturbation is introduced as a modulation superimposed onto the periodic structure of $\vec{\epsilon}(\boldsymbol{r})$ and $\vec{\mu}(\boldsymbol{r})$ as

$$\vec{\epsilon}^{-1}(\boldsymbol{r}) \to \gamma^2_\epsilon(\boldsymbol{r}) \vec{\epsilon}^{-1}(\boldsymbol{r}), \quad \vec{\mu}^{-1}(\boldsymbol{r}) \to \gamma^2_\mu(\boldsymbol{r}) \vec{\mu}^{-1}(\boldsymbol{r}), \tag{38}$$

where $\gamma_\epsilon(\boldsymbol{r})$ and $\gamma_\mu(\boldsymbol{r})$ are scalar functions, and we call them "the modulation functions" hereafter. This kind of modulation does not change the local symmetries of $\vec{\epsilon}(\boldsymbol{r})$ and $\vec{\mu}(\boldsymbol{r})$, and does not violate the energy conservation. We summarize the definitions of Berry connection and curvature in Table II and the main results obtained here, i.e., the equations of motion for an optical wave packet, in Table III. Appendix D supplements details about the commutation relations and expectation values of various operators which are needed to derive the equations of motion in a modulated system.

Now we derive the equations of motion for the dynamics of an optical wave packet. An exact wave function $|\Psi\rangle$ satisfies the Schrödinger equation

$$i \frac{d}{dt} |\Psi\rangle = H |\Psi\rangle. \tag{39}$$

This equation is derived by applying the variational principle to the quantity [34]

$$L = \langle \Psi | i \frac{d}{dt} - H | \Psi \rangle. \tag{40}$$

It is natural to consider that the trajectory of the wave packet is determined in terms of the effective Lagrangian which is given by replacing $|\Psi\rangle$ with a variational wave packet $|W\rangle$, characterized mainly by centers of position and wave vector, $\boldsymbol{r}_c$ and $\boldsymbol{k}_c$ [35]. Although $|W\rangle$ can be brought closer to $|\Psi\rangle$ by enlarging the number of variational parameters,

TABLE II. Berry connection and curvature.

	Electronic system	Photonic system
Bloch function	$\|U_{n\lambda \boldsymbol{k}}\rangle$	$\|U^E_{n\lambda \boldsymbol{k}}\rangle$, $\|U^H_{n\lambda \boldsymbol{k}}\rangle$
Normalization	$\langle U_{n\lambda \boldsymbol{k}}\|U_{n'\lambda' \boldsymbol{k}}\rangle=\delta_{nn'}\delta_{\lambda\lambda'}$	$\langle U^E_{n\lambda \boldsymbol{k}}\|\vec{\epsilon}\|U^E_{n'\lambda' \boldsymbol{k}}\rangle=\langle U^H_{n\lambda \boldsymbol{k}}\|\vec{\mu}\|U^H_{n'\lambda' \boldsymbol{k}}\rangle=\delta_{nn'}\delta_{\lambda\lambda'}$
Berry connection	$[\boldsymbol{\Lambda}_{n\boldsymbol{k}}]_{\lambda\lambda'}=-i\langle U_{n\lambda \boldsymbol{k}}\|\boldsymbol{\nabla}_{\boldsymbol{k}}U_{n\lambda' \boldsymbol{k}}\rangle$	$\boldsymbol{\Lambda}_{n\boldsymbol{k}}=\frac{1}{2}[\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}+\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}]$
		$[\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}]_{\lambda\lambda'}=-i\langle U^E_{n\lambda \boldsymbol{k}}\|\vec{\epsilon}\|\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{n\lambda' \boldsymbol{k}}\rangle$
		$[\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}]_{\lambda\lambda'}=-i\langle U^H_{n\lambda \boldsymbol{k}}\|\vec{\mu}\|\boldsymbol{\nabla}_{\boldsymbol{k}}U^H_{n\lambda' \boldsymbol{k}}\rangle$
Berry curvature	$\boldsymbol{\Omega}_{n\boldsymbol{k}}=\boldsymbol{\nabla}_{\boldsymbol{k}}\times\boldsymbol{\Lambda}_{n\boldsymbol{k}}+i\boldsymbol{\Lambda}_{n\boldsymbol{k}}\times\boldsymbol{\Lambda}_{n\boldsymbol{k}}$	$\boldsymbol{\Omega}_{n\boldsymbol{k}}=\boldsymbol{\nabla}_{\boldsymbol{k}}\times\boldsymbol{\Lambda}_{n\boldsymbol{k}}+i\boldsymbol{\Lambda}_{n\boldsymbol{k}}\times\boldsymbol{\Lambda}_{n\boldsymbol{k}}$

those concerning the details of the wave packet are neglected here. This approximation is justified when a modulation is weak and slowly varying, and has been successfully applied to the quantum Hall system and gives the semiclassical understanding of the motion of magnetic Bloch states [36,37,54].

In general, the modulation may mix the creation and annihilation operators. Here we consider the situation in which the modulation is sufficiently weak and time-independent, and this mixing is negligible. Therefore the approximated wave packet can be constructed as

$$|W\rangle=\int_{\mathrm{BZ}}d\boldsymbol{k}w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)\sum_{\lambda}z_{c\lambda}a^{\dagger}_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}|0\rangle,$$

$$w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)=w_r(\boldsymbol{k}-\boldsymbol{k}_c)e^{-i\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)},\tag{41}$$

where $w_r(\boldsymbol{k}-\boldsymbol{k}_c)$ is a real function, and $w_r(\boldsymbol{k}-\boldsymbol{k}_c)$ and $z_{c\lambda}$ satisfy the normalization conditions, $\int_{\mathrm{BZ}}d\boldsymbol{k}w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)=1$ and $\Sigma_\lambda|z_{c\lambda}|^2=1$, respectively. We assume $w_r(\boldsymbol{k}-\boldsymbol{k}_c)$ has a sharp peak around $\boldsymbol{k}_c=\int_{\mathrm{BZ}}d\boldsymbol{k}w_r^2(\mathbf{k}-\mathbf{k}_c)\boldsymbol{k}$. Here we require that the center of wave packet, $\boldsymbol{r}_c$, is self-consistently determined by

$$\boldsymbol{r}_c=\int_{\mathrm{BZ}}d\boldsymbol{k}w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)-(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}}|z_c)],\tag{42}$$

where $\boldsymbol{\Lambda}_{n\boldsymbol{k}}$ is the Berry connection defined by

$$\boldsymbol{\Lambda}_{n\boldsymbol{k}}=\frac{1}{2}[\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}+\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}],\tag{43a}$$

$$[\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}]_{\lambda\lambda'}=-i\langle U^E_{n\lambda \boldsymbol{k}}|\vec{\epsilon}|\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{n\lambda' \boldsymbol{k}}\rangle,\tag{43b}$$

$$[\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}]_{\lambda\lambda'}=-i\langle U^H_{n\lambda \boldsymbol{k}}|\vec{\mu}|\boldsymbol{\nabla}_{\boldsymbol{k}}U^H_{n\lambda' \boldsymbol{k}}\rangle,\tag{43c}$$

and we introduced the abbreviation

$$(z_c|M|z_c)=\sum_{\lambda,\lambda'}z^*_{c\lambda}M_{\lambda\lambda'}z_{c\lambda'}.\tag{44}$$

The annihilation and creation operators of approximated eigenmodes are defined by

$$a_{n\lambda \boldsymbol{k};\boldsymbol{r}_c}=\int d\boldsymbol{r}\left[\sqrt{\frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}}\boldsymbol{\Phi}^{E*}_{n\lambda \boldsymbol{k}}(\boldsymbol{r})\cdot\boldsymbol{D}(\boldsymbol{r})+\sqrt{\frac{\gamma_\mu(\boldsymbol{r}_c)}{\gamma_\epsilon(\boldsymbol{r}_c)}}\boldsymbol{\Phi}^{H*}_{n\lambda \boldsymbol{k}}(\boldsymbol{r})\cdot\boldsymbol{B}(\boldsymbol{r})\right],\tag{45a}$$

$$a^{\dagger}_{n\lambda \boldsymbol{k};\boldsymbol{r}_c}=\int d\boldsymbol{r}\left[\sqrt{\frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}}\boldsymbol{D}(\boldsymbol{r})\cdot\boldsymbol{\Phi}^{E}_{n\lambda \boldsymbol{k}}(\boldsymbol{r})+\sqrt{\frac{\gamma_\mu(\boldsymbol{r}_c)}{\gamma_\epsilon(\boldsymbol{r}_c)}}\boldsymbol{B}(\boldsymbol{r})\cdot\boldsymbol{\Phi}^{H}_{n\lambda \boldsymbol{k}}(\boldsymbol{r})\right].\tag{45b}$$

These operators satisfy the same commutation relation as that in a periodic system,

$$[a_{n\lambda \boldsymbol{k};\boldsymbol{r}_c},a^{\dagger}_{n'\lambda'\boldsymbol{k}';\boldsymbol{r}_c}]=\delta_{nn'}\delta_{\lambda\lambda'}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}').\tag{46}$$

The approximated eigenmodes depend on the variable $\boldsymbol{r}_c$. Thus we must consider also the operator $\boldsymbol{\nabla}_{\boldsymbol{r}_c}a^{\dagger}_{n\lambda \boldsymbol{k};\boldsymbol{r}_c}$ when estimating the effective Lagrangian. However, we can show that the contribution from $\boldsymbol{\nabla}_{\boldsymbol{r}_c}a^{\dagger}_{n\lambda \boldsymbol{k};\boldsymbol{r}_c}$ vanishes, by using $[a_{n\lambda \boldsymbol{k};\boldsymbol{r}_c},\boldsymbol{\nabla}_{\boldsymbol{r}_c}a^{\dagger}_{n'\lambda'\boldsymbol{k}';\boldsymbol{r}_c}]=0$ and $a_{n\lambda \boldsymbol{k};\boldsymbol{r}_c}|0\rangle=0$.

As was mentioned above, we consider the situation in which the mixing of the approximated annihilation and creation operators is negligible and assume that the relation

TABLE III. Equations of motion of optical wave packet.

Modulation	$\boldsymbol{\epsilon}^{-1}(\boldsymbol{r})\to\gamma^2_\epsilon(\boldsymbol{r})\boldsymbol{\epsilon}^{-1}(\boldsymbol{r})$, $\boldsymbol{\mu}^{-1}(\boldsymbol{r})\to\gamma^2_\mu(\boldsymbol{r})\boldsymbol{\mu}^{-1}(\boldsymbol{r})$
Effective Lagrangian	$L_{\mathrm{eff}}\cong\boldsymbol{k}_c\cdot\dot{\boldsymbol{r}}_c-\dot{\boldsymbol{k}}_c\cdot(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)+i(z_c|\dot{z}_c)-\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}$
Equations of motion	$\dot{\boldsymbol{r}}_c=\boldsymbol{\nabla}_{\boldsymbol{k}_c}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}+\dot{\boldsymbol{k}}_c\times(z_c|\boldsymbol{\Omega}_{n\boldsymbol{k}_c}|z_c)-i(z_c|[\boldsymbol{f}^{\Delta}_c\cdot\boldsymbol{\Delta}_{n\boldsymbol{k}_c},\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}]|z_c)$
	$\dot{\boldsymbol{k}}_c=-\boldsymbol{\nabla}_{\boldsymbol{r}_c}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}$, $(\dot{z}_c|=-i[\dot{\boldsymbol{k}}_c\cdot\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}+\boldsymbol{f}^{\Delta}_c\cdot\boldsymbol{\Delta}_{n\boldsymbol{k}_c}]|z_c)$
Perturbed energy	$\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}=E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}+\boldsymbol{f}^{\Delta}_c\cdot(z_c|\boldsymbol{\Delta}_{n\boldsymbol{k}_c}|z_c)$
	$E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}=\gamma_\epsilon(\boldsymbol{r}_c)\gamma_\mu(\boldsymbol{r}_c)E_{n\boldsymbol{k}_c}$, $\boldsymbol{f}^{\Delta}_c=-\left[\nabla_{\boldsymbol{r}_c}\ln\frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right]E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}$, $\boldsymbol{\Delta}_{n\boldsymbol{k}_c}=\frac{1}{2}[\boldsymbol{\Lambda}^E_{n\boldsymbol{k}_c}-\boldsymbol{\Lambda}^H_{n\boldsymbol{k}_c}]$
Energy conservation	$\frac{d}{dt}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}=0$

$a_{n\lambda \boldsymbol{k};\boldsymbol{r}_c}|0\rangle=0$ holds. The expectation values of H and $\mathcal{R}$ are estimated by the derivative expansion with respect to $\gamma_\epsilon(\boldsymbol{r})$ and $\gamma_\mu(\boldsymbol{r})$ as

$$\langle W|H|W\rangle \cong \mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}, \tag{47a}$$

$$\langle W|\mathcal{R}|W\rangle \cong E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}[\nabla_{\boldsymbol{k}_c}\vartheta(\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)-(z_c|\Lambda_{n\boldsymbol{k}_c}|z_c)], \tag{47b}$$

where $E_{n\boldsymbol{k};\boldsymbol{r}_c}=\gamma_\epsilon(\boldsymbol{r}_c)\gamma_\mu(\boldsymbol{r}_c)E_{n\boldsymbol{k}}$,

$$\frac{\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}}{E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}}=1-\left[\nabla_{\boldsymbol{r}_c}\ln\frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right]\cdot(z_c|\boldsymbol{\Delta}_{n\boldsymbol{k}_c}|z_c), \tag{48}$$

and

$$\boldsymbol{\Delta}_{n\boldsymbol{k}}=\frac{1}{2}[\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}-\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}]. \tag{49}$$

This function $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ is a difference between the Berry connections of the electric and magnetic parts. Although $(z_c|\boldsymbol{\Lambda}^{E,H}_{n\boldsymbol{k}}|z_c)$ depends on the representations of eigenmodes, $(z_c|\boldsymbol{\Delta}_{n\boldsymbol{k}}|z_c)$ does not. This issue is related to the gauge transformation in the $\boldsymbol{k}$ space given in Appendix C2. The property of $(z_c|\boldsymbol{\Delta}_{n\boldsymbol{k}}|z_c)$ is also discussed in Appendix E.

In the above evaluations, we assumed that the shape of $w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)$ is sufficiently sharp compared to the slow variations of $E_{n\boldsymbol{k};\boldsymbol{r}_c}$ and $\boldsymbol{\Lambda}_{n\boldsymbol{k}}(\boldsymbol{k})$ around $\boldsymbol{k}_c$, and neglected terms which depend on the shape of $w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)$. Therefore even with the perturbative modulation, we can regard $\boldsymbol{r}_c$ defined by Eq. (42) as the center of gravity $\langle W|\mathcal{R}|W\rangle/\langle W|H|W\rangle$. In the present approximation, we assume that the modulation is so weak and smooth that we can neglect the second-order derivatives of the modulation functions in the effective Lagrangian and the equations of motion, which we shall derive later. In the equations of motion, the difference between $\boldsymbol{r}_c$ and the center of gravity appears as higher derivatives than the original derivatives. Therefore we may neglect the difference between $\boldsymbol{r}_c$ and the center of gravity due to the derivatives of the modulation functions, while we cannot neglect the correction due to the first derivatives in $\langle W|H|W\rangle$.

Here we introduce the effective Lagrangian in order to derive the equations of motion of a wave packet.

$$L_{\text{eff}}=\langle W|i\frac{d}{dt}-H|W\rangle. \tag{50}$$

The first term on the right-hand side of Eq. (50) is calculated as follows:

$$\langle W|i\frac{d}{dt}|W\rangle \cong \boldsymbol{k}_c\cdot\dot{\boldsymbol{r}}_c-\dot{\boldsymbol{k}}_c\cdot(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)+i(z_c|\dot{z}_c) + \frac{d}{dt}\left[\int_{\text{BZ}}d\boldsymbol{k}w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)-\boldsymbol{k}_c\cdot\boldsymbol{r}_c\right]. \tag{51}$$

Neglecting the total time derivative, we obtain the final form of the effective Lagrangian,

$$L_{\text{eff}}\cong \boldsymbol{k}_c\cdot\dot{\boldsymbol{r}}_c-\dot{\boldsymbol{k}}_c\cdot(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)+i(z_c|\dot{z}_c)-\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}. \tag{52}$$

From this Lagrangian, the equations of motion are derived as follows:

$$\dot{\boldsymbol{r}}_c=\nabla_{\boldsymbol{k}_c}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}+\dot{\boldsymbol{k}}_c\times(z_c|\boldsymbol{\Omega}_{n\boldsymbol{k}_c}|z_c)-i(z_c|[\boldsymbol{f}_c^\Delta\cdot\boldsymbol{\Delta}_{n\boldsymbol{k}_c},\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}]|z_c), \tag{53a}$$

$$\dot{\boldsymbol{k}}_c=-\nabla_{\boldsymbol{r}_c}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}, \tag{53b}$$

$$|\dot{z}_c)=-i[\dot{\boldsymbol{k}}_c\cdot\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}+\boldsymbol{f}_c^\Delta\cdot\boldsymbol{\Delta}_{n\boldsymbol{k}_c}]|z_c), \tag{53c}$$

where

$$\boldsymbol{f}_c^\Delta=-\left[\nabla_{\boldsymbol{r}_c}\ln\frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right]E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}, \tag{54}$$

and $\boldsymbol{\Omega}_{n\boldsymbol{k}}$ is the Berry curvature defined by

$$\boldsymbol{\Omega}_{n\boldsymbol{k}}=\nabla_{\boldsymbol{k}}\times\boldsymbol{\Lambda}_{n\boldsymbol{k}}+i\boldsymbol{\Lambda}_{n\boldsymbol{k}}\times\boldsymbol{\Lambda}_{n\boldsymbol{k}}. \tag{55}$$

It should be noted that the above equations of motion satisfy the energy conservation, i.e.,

$$\frac{d}{dt}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}=0. \tag{56}$$

Rigorously speaking, a Lagrange multiplier is needed for the constraint $(z_c|z_c)=\Sigma_\lambda|z_{c\lambda}|^2=1$ in the derivation of the equations of motion, while the constraint is implicitly imposed here. Therefore in the above equations of motion, we should consider that this constraint is always imposed.

In generic cases with periodic structures, we cannot analytically evaluate $\boldsymbol{\Lambda}_{n\boldsymbol{k}}$, $\boldsymbol{\Omega}_{n\boldsymbol{k}}$, or $\boldsymbol{\Delta}_{n\boldsymbol{k}}$. However, in principle, we can numerically calculate them and thus solve the equations of motion of a wave packet subject to a modulation superimposed onto a periodic structure. Appendix E presents formulae for $\boldsymbol{\Omega}_{n\boldsymbol{k}}$ and $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ as well as the internal rotation of a wave packet, which are useful for numerical calculations. Although the relation $(z_c|\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}|z_c)=(z_c|\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}|z_c)$ is not generally proved, this can be confirmed at least for some systems with locally isotropic $\boldsymbol{\epsilon}(\boldsymbol{r})$ and $\boldsymbol{\mu}(\boldsymbol{r})$, e.g., for the elliptically polarized light in systems without periodic structure and for the TM modes in two-dimensional photonic crystals where μ is constant. In this case with $\boldsymbol{\Delta}_{n\boldsymbol{k}}=0$, we can replace the perturbed energy as $\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c;z_c}\to E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}=\gamma_\epsilon(\boldsymbol{r}_c)\gamma_\mu(\boldsymbol{r}_c)E_{n\boldsymbol{k}_c}$.

Finally, it is noted that the optical Hall effect is originated by the second term on the right-hand side of Eq. (53a), which is sometimes called "anomalous velocity." The anomalous velocity is the vector product of the Berry curvature $(z_c|\boldsymbol{\Omega}_{n\boldsymbol{k}_c}|z_c)$ and the driving force $\dot{\boldsymbol{k}}_c$, i.e., the gradient of a superimposed modulation. Therefore both of the driving force and the Berry curvature are needed for this phenomenon. In an electronic system under a strong magnetic field, it is pointed out that the Berry curvature is closely related to the internal rotation of an electron wave packet [36]. Thus a similar relation is also expected in a photonic system. Actually, in Appendixes C and E, we can see the close relation between the Berry curvature and the internal rotation of an optical wave packet.

III. APPLICATIONS

A. Transverse shift in reflection and refraction

As an application of the theoretical framework developed in Sec. II, we consider the case with locally isotropic dielectric permittivity $\epsilon(\boldsymbol{r})$ and magnetic permeability $\mu(\boldsymbol{r})$. (See Tables II and III where the main results in Sec. II are summarized.) In this case, we can write down the equations of motion (53a)–(53c) in simple forms. For this purpose, we first calculate the Berry connection $\boldsymbol{\Lambda}_{\boldsymbol{k}}$ and the Berry curvature $\boldsymbol{\Omega}_{\boldsymbol{k}}$. In a helicity basis, the eigenvectors for the right (+)- and left(−)-circular polarizations can be written as

$$\boldsymbol{U}^{E}_{\pm,\boldsymbol{k}} = \frac{1}{\sqrt{2\epsilon}}(\boldsymbol{e}_{\theta} \pm i\boldsymbol{e}_{\phi}), \quad \boldsymbol{U}^{H}_{\pm,\boldsymbol{k}} = \frac{1}{\sqrt{2\mu}}(\boldsymbol{e}_{\phi} \mp i\boldsymbol{e}_{\theta}), \tag{57}$$

where $\boldsymbol{e}_{\theta,\phi}$ with $\boldsymbol{e}_{k}=\boldsymbol{k}/k$ are the orthogonal unit vectors in the spherical coordinate of the $\boldsymbol{k}$ space. After some calculations we obtain

$$\boldsymbol{\Lambda}^{E}_{\boldsymbol{k}} = \boldsymbol{\Lambda}^{H}_{\boldsymbol{k}} = -\frac{\cos\theta}{k\sin\theta}\sigma_3\boldsymbol{e}_{\phi}, \tag{58}$$

which then yields [55]

$$\boldsymbol{\Lambda}_{\boldsymbol{k}} = -\frac{\cos\theta}{k\sin\theta}\sigma_3\boldsymbol{e}_{\phi}, \quad \boldsymbol{\Omega}_{\boldsymbol{k}} = \frac{\boldsymbol{k}}{k^3}\sigma_3, \quad \boldsymbol{\Delta}_{\boldsymbol{k}} = 0. \tag{59}$$

Thus the equations of motion are simplified as follows:

$$\dot{\boldsymbol{r}}_c = v(\boldsymbol{r}_c)\frac{\boldsymbol{k}_c}{k_c} + \dot{\boldsymbol{k}}_c \times (z_c|\boldsymbol{\Omega}_{\boldsymbol{k}_c}|z_c), \tag{60a}$$

$$\dot{\boldsymbol{k}}_c = -[\boldsymbol{\nabla}_{\boldsymbol{r}_c} v(\boldsymbol{r}_c)]k_c, \tag{60b}$$

$$|\dot{z}_c) = -i\dot{\boldsymbol{k}}_c \cdot \boldsymbol{\Lambda}_{\boldsymbol{k}_c}|z_c), \tag{60c}$$

where $\boldsymbol{r}_c$, $\boldsymbol{k}_c$, and $|z_c)=[z_+,z_-]$ ($|z_+|^2+|z_-|^2=1$) are the position, the momentum, and the polarization state of an optical wave packet, respectively, and $v(\boldsymbol{r})=1/\sqrt{\epsilon(\boldsymbol{r})\mu(\boldsymbol{r})}$. As is analogous to the Hall effects in electronic systems, the second term on the right-hand side of Eq. (60a) describes the optical Hall effect induced by a modulation of refractive index [20]. The equation for $|z_c)$, Eq. (60c), describes a phase shift by the directional change of propagation discussed in Refs. [17–19]. This equation gives the solution $|z^{\text{out}}_c) = [e^{-i\Theta}z^{\text{in}}_+, e^{i\Theta}z^{\text{in}}_-]$, where $|z^{\text{in}}_c) = [z^{\text{in}}_+, z^{\text{in}}_-]$ is the initial state of polarization. Θ is a solid angle made by the trajectory of momentum: $\Theta = \oint d\boldsymbol{k}\cdot[\boldsymbol{\Lambda}_{\boldsymbol{k}}]_{++} = \int_S d\boldsymbol{S}_{\boldsymbol{k}}\cdot[\boldsymbol{\Omega}_{\boldsymbol{k}}]_{++}$ where $d\boldsymbol{S}_{\boldsymbol{k}}$ is the surface element in the $\boldsymbol{k}$ space and S is a surface surrounded by the trajectory. Our approach can be easily generalized to treat systems with periodic structures on the same footing, and offers a powerful tool for applications compared with the eikonal approximation [30].

The simplest example of the optical Hall effect is realized as the transverse shift at the interface refraction and reflection. There have been a number of studies on the shifts within and out of the incident plane at the total reflection. The former is well known as the Goos-Hänchen effect [56,57] and has been explained in terms of the evanescent wave penetrating into the forbidden region. The latter one, which is referred to as the Imbert-Fedorov shift, was interpreted by Fedorov [21] as an analog of the Goos-Hänchen effect and was observed experimentally by Imbert using multiple total reflections [22], followed by a number of theoretical approaches [23–25]. Furthermore, it was pointed out that the shift out of the incident plane could also occur in partial reflection and refraction [25–27]. However, some of the theoretical predictions for the amount of shift contradict each other. One reason is an experimental difficulty for a measurement of the tiny shift, as the shift is only a fraction of a wavelength. It was only recent that the Imbert-Fedorov shift is measured for a single total reflection [28]. Thus the physical mechanism for the transverse shift is still controversial.

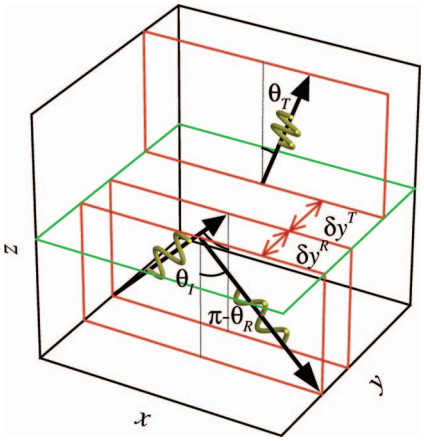

FIG. 1. (Color) Transverse shift of light beams in the refraction and reflection at an interface.

In our previous paper [20], we calculated this transverse shift by using the conservation of the z component of TAM *for individual photons*, which follows from the equations of motion (60a)–(60c) applied to this interface problem. Bliokh *et al.* [50] then questioned the result, claiming that it does not match their result for elliptically polarized Gaussian beams. Here we show that, in each case of generic polarizations, an identical result for the transverse shift of each beam is given by the following different approaches: (i) analytic evaluation of wave-packet dynamics, (ii) TAM conservation for individual photons in Ref. [20], and (iii) numerically exact simulation of wave-packet dynamics. It agrees with a result by classical electrodynamics, as presented in Appendix F. In Sec. III B, we shall resolve the inconsistency between the identical result by the approaches (i)–(iii) and that given in Ref. [50]. Thereby the validity of our theory presented here is completely guaranteed.

Our equations of motion are not directly applicable to the refraction and reflection problem at a sharp interface since they require the slowly varying conditions $|\boldsymbol{\nabla}\ln\epsilon|, |\boldsymbol{\nabla}\ln\mu| \ll k$. Indeed, our equations of motion do not correctly describe a splitting of an incident wave packet into reflected and transmitted wave packets or changes of their polarization states at the interface. However, in a case with a flat interface, a simple extension of our theory works well, as will be explained in the following. Here, as shown in Fig. 1, we consider the case where the incident beam comes from the

region with $x<0$ and $z<0$ along the plane of $y=$const, and the interface is the $z=0$ plane.

1. Analytic evaluation

Far from the interface, $\boldsymbol{r}_c$ in Eq. (42) is easily estimated as

$$\boldsymbol{r}_c|_{t\to\pm\infty} \cong \boldsymbol{\nabla}_{\boldsymbol{k}^A}\vartheta^A(\boldsymbol{k}^A, t\to\pm\infty) - (z^A|\boldsymbol{\Lambda}_{\boldsymbol{k}^A}|z^A), \quad (61)$$

where $A=I$ for $t\to-\infty$, $A=T$ or R for $t\to\infty$. The momenta $\boldsymbol{k}^{I,T,R}$ and the polarization states $|z^{I,T,R}\rangle$ are those of the incident (I), transmitted (T), and reflected (R) beams, respectively. Due to the wave-packet splitting at the interface, our semiclassical equations of motion do not tell us the values of $|z^A\rangle$ and ϑ^A. Hence for these variables, we borrow the results of reflection or refraction of a plane wave. With a natural choice of a wave packet presented in Appendix F, the y component of the first term in Eq. (61) is unchanged at the interface. Thus the transverse shift comes only from the second term as

$$\delta y^A = -(z^A|\boldsymbol{\Lambda}_{\boldsymbol{k}^A}|z^A) + (z^I|\boldsymbol{\Lambda}_{\boldsymbol{k}^I}|z^I), \quad (62)$$

where $A=T$ or R. Substituting Eq. (59) in each Berry connection $(z^A|\boldsymbol{\Lambda}_{\boldsymbol{k}^A}|z^A)$ $(A=I,T,R)$, we obtain the following equation for the transverse shift:

$$\delta y^A = \frac{1}{k^I \sin\theta_I}[(z^A|\sigma_3|z^A)\cos\theta_A - (z^I|\sigma_3|z^I)\cos\theta_I], \quad (63)$$

where $A=T$ or R, $\theta_{I,T,R}$ are the angles between the positive z axis and the propagating directions of the incident, transmitted, and reflected beams, respectively.

2. Total angular momentum conservation

Then, what is the physical meaning of the above result? First, it should be noted that $(z|\sigma_3|z)=|z_+|^2-|z_-|^2$ represents the magnitude of spin polarization in the direction of $\boldsymbol{k}_c$, i.e., $(z|\sigma_3|z)=\pm1$ for right- or left-circular polarizations, $(z|\sigma_3|z)=0$ for linear polarizations, and $|(z|\sigma_3|z)|<1$ for elliptic polarizations. Therefore it is intuitively expected that this phenomenon is closely related to the angular momentum of a wave packet. For a system with rotational symmetry around the z axis, the equations of motion lead to the conservation of the z component of the following TAM:

$$\boldsymbol{j}_c = \boldsymbol{r}_c \times \boldsymbol{k}_c + (z_c|\sigma_3|z_c)\frac{\boldsymbol{k}_c}{k_c}. \quad (64)$$

This conservation is expected to hold even in the case of a sharp interface, because it is based on the rotational symmetry around the z axis. Actually, from Eq. (63) for the transverse shift and Eq. (64) for the TAM, we can reach the conservation of the z component of the TAM *for each of individual photons*,

$$j_z^I = j_z^T, \quad j_z^I = j_z^R, \quad (65)$$

where $\boldsymbol{j}^{I,T,R}$ are the TAM of incident, transmitted, and reflected beams, respectively. It just makes sense that the incident beam is regarded as a collection of photons; each photon is reflected or transmitted stochastically at the interface. As we shall see in Sec. III B (and Appendix F in detail), Eq. (63) for the transverse shift is consistent with the result derived in classical electrodynamics. In this sense, this photon picture is implicitly incorporated already in classical electrodynamics.

Inversely, assuming the conservation of TAM for individual photons, we can derive the transverse shift as Eq. (63). This is what we have done in our previous paper [20]. This derivation of the transverse shift is akin to the derivation of Snell's law based on the particle picture of light in which the refracted and reflected angles are obtained from the conservation of energy and momentum (parallel to the interface) for individual photons.

3. Numerical simulation

In order to verify our theory quantitatively, we check the property of the transverse shift in more detail. To obtain $(z^A|\sigma_3|z^A)$, we decompose the incident wave as

$$|z^I) = \frac{z_+^I + z_-^I}{\sqrt{2}}|p) + \frac{i(z_+^I - z_-^I)}{\sqrt{2}}|s), \quad (66)$$

where $|p)=1/\sqrt{2}[1,1]$ and $|s)=1/\sqrt{2}[-i,i]$ represent the p- and s-polarized states. Straightforward calculation yields

$$(z^A|\sigma_3|z^A) = \frac{2[(z^I|\sigma_3|z^I)\mathrm{Re}(A_p^* A_s) + (z^I|\sigma_2|z^I)\mathrm{Im}(A_p^* A_s)]}{[1+(z^I|\sigma_1|z^I)]|A_p|^2 + [1-(z^I|\sigma_1|z^I)]|A_s|^2}, \quad (67)$$

with $A=T$ or R, and T_p and T_s (R_p and R_s) are the amplitude transmission (reflection) coefficients for p and s polarization, respectively.

When we focus on the partial reflection and refraction, A_p and A_s are real, and Eq. (63) is rewritten as follows,

$$\delta y^A = \frac{(z^I|\sigma_3|z^I)}{k^I \tan\theta_I}\left[\frac{2A_p A_s \cos\theta_A/\cos\theta_I}{[1+(z^I|\sigma_1|z^I)]A_p^2 + [1-(z^I|\sigma_1|z^I)]A_s^2} - 1\right], \quad (68)$$

where $A=R$ or T. This means that the incident beams with $|z)=[e^{\pm i(\phi/2)}\cos\theta/2, e^{\pm i(\phi/2)}\sin\theta/2]$, where θ and ϕ represent the spherical coordinate of the Poincaré sphere, cause the shift of the same magnitude and the same direction with each other, i.e., the shift independent of the sign of ϕ. In addition, the incident beams with $|z)=[e^{\mp i(\phi/2)}\sin\theta/2, e^{\pm i(\phi/2)}\cos\theta/2]$ cause the shift of the same magnitude as the above beams, but of the opposite direction to them. In the partial reflection and refraction, no shift is observed when an incident beam is linearly polarized.

On the other hand, in the total reflection, we have $|R_p|=|R_s|=1$, and the shift of the reflected beam is represented by

$$\delta y^R = -\frac{1}{k^I \tan\theta_I}[(z^I|\sigma_3|z^I)[\mathrm{Re}(R_p^* R_s)+1] + (z^I|\sigma_2|z^I)\mathrm{Im}(R_p^* R_s)]. \quad (69)$$

In particular, for the incident beam with linear polarization $|z)=[e^{-i(\phi/2)}/\sqrt{2}, e^{i(\phi/2)}/\sqrt{2}]$, the shift is the same magnitude

and the same direction for $\phi=\alpha$ and $\pi-\alpha$. The direction is reversed by the replacement $\phi\rightarrow-\phi$ without change of the magnitude.

We have confirmed all the above features quantitatively by numerically solving Maxwell equations for wave packets. In Ref. [20], we have presented the results only for the incident beam with right-circular polarization. Here, to complete the argument, we present the results of the numerical simulations for more generic cases. Figure 2 shows the shifts for the incident beam with the elliptical polarization $z_+^I/z_-^I=2$ at the interfaces with relative refractive indices (a) $n=2.0$, (b) $n=0.8$, and (c) for the incident beam with linear polarization $z_+^I/z_-^I=i$ at the interface with $n=0.5$. (We take the value of magnetic permeability common in both media upper and lower the interface, i.e., $\mu_1=\mu_2$, in these simulations.) The solid and dashed lines represent the analytic results Eq. (63) for transmitted and reflected beams, respectively. The filled circles and squares are the results of simulations for transmitted and reflected beams. We note that, in Fig. 2(c), the shift for the linearly polarized beam is nonzero only for a region of total reflection, in accordance with our analytic result. In all cases, the numerical results agree excellently with Eq. (63), thus verifying our theory. (We have confirmed this consistency also in cases where both of permittivity and permeability are different in two media upper and lower the interface, i.e., $\epsilon_1\neq\epsilon_2$ and $\mu_1\neq\mu_2$.)

Finally we should comment on our constitution method of a set of incident, transmitted, and reflected wave packets, which is an exact solution of the Maxwell equations. In each numerical simulation, we have constructed an elliptically polarized incident wave packet as a superposition of plane waves with a common polarization state, i.e., $|z^I\rangle$ which is independent of $\boldsymbol{k}$. This is a natural definition of incident wave packet. Otherwise, the concept "an elliptically polarized incident wave packet" gets fuzzy, and the linear composition from and decomposition to different orthonormal bases of incident wave packets are violated. This is because incident wave packets with different functions of $\{|z^I(\boldsymbol{k})\rangle\}$'s for constituent plane waves can have the same mean polarization state $|z^I(\boldsymbol{k}^I)\rangle$. Imposing the exact boundary conditions, transmitted and reflected wave packets are automatically generated. For partial reflection, a single incident wave packet split into reflected and transmitted wave packets after reflection or refraction at the interface. The position of each wave packet is estimated when each wave packet is far from the interface. It should be noted that the numerical simulations exactly take into account the changes of shapes of wave packets, while the analytic evaluation assumes the sharpness of a weight function for the superposition.

B. Remarks on other theories

Recently, Bliokh *et al.* calculated the shift for an elliptic Gaussian incident beam in classical electrodynamics, and their result disagrees with that obtained from our theory [50]. They attributed the difference to a "fallacy" in our TAM conservation for individual photons. We explain below in detail that our theory is totally free from the criticism. To prove this, it is enough to show that Eq. (63) is equivalent to the transverse shift evaluated in classical electrodynamics, i.e., the result by Fedoseev [26,27]. His procedure of calculation is as follows. First construct a wave packet by a linear superposition of plane waves. By taking into account the

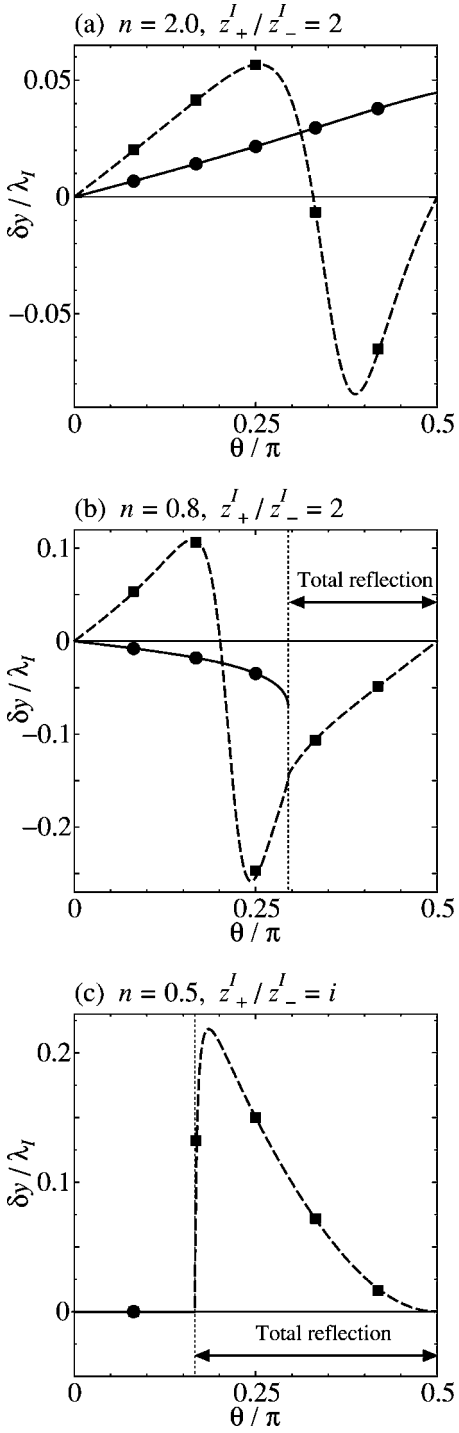

FIG. 2. Shifts of reflected and transmitted beams. n is the relative refractive index of the upper medium with respect to the lower medium. λ_I is the wavelength of incident light in the lower medium. The solid and dashed lines represent the analytic results Eq. (63) for transmitted and reflected beams, respectively. The filled circles and squares are the results of simulations for transmitted and reflected beams.

exact boundary conditions of electromagnetic fields [Eqs. (70a) and (70b)] at a flat interface, as in the textbooks of optics or classical electrodynamics [38,57], one can construct transmitted and reflected wave packets as a set of them is an exact solution of the Maxwell equations. The center of each wave packet is defined as an average position weighted by each energy density. The result of this calculation by classical electrodynamics [26,27] is identical with that of our theory [Eqs. (68) and (69) derived from Eq. (63)] in the second quantized formalism. The details are presented in Appendix F. Hence we checked that the following three approaches give the same transverse shift for each wave packet: (i) analytic evaluation of wave-packet dynamics both in classical and quantum-mechanical formalisms, (ii) TAM conservation for individual photons [Eq. (65)], (iii) numerically exact simulation of wave-packet dynamics.

There remains an inconsistency between the identical result obtained by (i)–(iii) and one by Bliokh *et al.* in Ref. [50]. One reason is an inappropriate boundary condition for a set of wave packets in their paraxial approximation. This boundary condition is different from the correct one:

$$\boldsymbol{t}\cdot[\boldsymbol{E}^I(\boldsymbol{r},t)+\boldsymbol{E}^R(\boldsymbol{r},t)]=\boldsymbol{t}\cdot\boldsymbol{E}^T(\boldsymbol{r},t), \tag{70a}$$

$$\boldsymbol{t}\cdot[\boldsymbol{H}^I(\boldsymbol{r},t)+\boldsymbol{H}^R(\boldsymbol{r},t)]=\boldsymbol{t}\cdot\boldsymbol{H}^T(\boldsymbol{r},t), \tag{70b}$$

where $\boldsymbol{t}$ is an arbitrary unit vector parallel to the interface, $\boldsymbol{E}^A$ and $\boldsymbol{H}^A$ $(A=I,T,R)$ are electric and magnetic fields of incident (I), transmitted (T), and reflected (R) beams, respectively. Another reason for this contradiction comes from the definition of a center of wave packet in Ref. [50]. The methods (i)–(iii) commonly use the position averaged with a weight of energy density. In Ref. [50], on the other hand, the center is defined as a center of the wave packet *projected onto* its mean polarization state. The center of the wave packet in the former definition, i.e., the position averaged by the energy density, can be easily measured by photon counting, as employed in two measurements on the Imbert-Fedorov shift [22,28], while the latter definition requires counting of photons projected onto a specified polarization. The agreement between totally different approaches, i.e., (i)–(iii) and classical electrodynamics, suggests that our definition is a natural one.

Finally we should comment on the relation between the conservation laws of TAM in the wave and particle pictures of light. In Ref. [50], it is claimed that, for an incident beam with an elliptic polarization, the conservation of TAM for individual photons [Eq. (65)] is inconsistent with the conservation of TAM for whole beams,

$$j_z^I=R^2j_z^R+T^2\frac{n_2\mu_1\cos\theta^T}{n_1\mu_2\cos\theta^I}j_z^T. \tag{71}$$

However, as we shall show below, Eq. (65) is a sufficient condition for Eq. (71). From Fresnel formulas, we have

$$1=R_p^2+T_p^2\frac{n_2\mu_1\cos\theta_T}{n_1\mu_2\cos\theta_I}, \tag{72a}$$

$$1=R_s^2+T_s^2\frac{n_2\mu_1\cos\theta_T}{n_1\mu_2\cos\theta_I}, \tag{72b}$$

for the p- and s-polarized beams, and also

$$1=R^2+T^2\frac{n_2\mu_1\cos\theta_T}{n_1\mu_2\cos\theta_I}, \tag{73}$$

for a beam with an arbitrary polarization, where

$$R^2=\frac{R_p^2+|m|^2R_s^2}{1+|m|^2},\quad T^2=\frac{T_p^2+|m|^2T_s^2}{1+|m|^2}, \tag{74}$$

and $m=z_s/z_p$, $|z)=z_p|p)+z_s|s)$. The above formula represents the conservation of energy flow or equivalently the conservation of the number of photons. One can easily see that the above formula and Eq. (65) yields Eq. (71). To summarize, the TAM conservation for individual photons [Eq. (65)] has neither contradiction nor inconsistency with other theories.

C. Two-dimensional photonic crystal

We consider a two-dimensional photonic crystal, where $\vec{\epsilon}(\boldsymbol{r})$ and $\vec{\mu}(\boldsymbol{r})$ are periodically modulated in the xy plane and uniform along the z direction. For simplicity, $\vec{\epsilon}(\boldsymbol{r})$ and $\vec{\mu}(\boldsymbol{r})$ are assumed to be locally isotropic and replaced by scalar variables $\epsilon(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$. It is noted that in general there may appear ordinary degeneracies at symmetric points and accidental degeneracies at some specific points in the Brillouin zone. Around these points, the semiclassical argument based on the adiabaticity would not be a good approximation, and it is needed to seriously incorporate wave-packet dynamics. Here we restrict ourselves to bands without degeneracy. In this case, the inversion symmetry of the periodic structure must be broken in order for a band to have nonzero Berry curvature. This is because the Fourier transformation of $\epsilon(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$ are real-valued, when a system has the inversion symmetry.

For a wave packet constructed from a nondegenerate band, the equations of motion in Eqs. (53a)–(53c) are reduced to the following ones:

$$\dot{\boldsymbol{r}}_c=\boldsymbol{\nabla}_{\boldsymbol{k}_c}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c}+\dot{\boldsymbol{k}}_c\times\boldsymbol{\Omega}_{n\boldsymbol{k}_c}, \tag{75a}$$

$$\dot{\boldsymbol{k}}_c=-\boldsymbol{\nabla}_{\boldsymbol{r}_c}\mathcal{E}_{n\boldsymbol{k}_c;\boldsymbol{r}_c}, \tag{75b}$$

$$\dot{z}_c=-i[\dot{\boldsymbol{k}}_c\cdot\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}+\boldsymbol{f}_c^{\Delta}\cdot\boldsymbol{\Delta}_{n\boldsymbol{k}_c}]z_c, \tag{75c}$$

where $\boldsymbol{r}_c$ and $\boldsymbol{k}_c$ are the position and the momentum of an optical wave packet. The parameter z_c is a simple complex number and just represents a phase shift. The definition of other variables in the above equations are given in Tables II and III. In the above equations of motion, the most important and controllable quantity is the second term on the right-hand side of Eq. (75a), i.e., the Berry curvature $\boldsymbol{\Omega}_{n\boldsymbol{k}_c}$. The anomalous velocity $\dot{\boldsymbol{k}}_c\times\boldsymbol{\Omega}_{n\boldsymbol{k}_c}$ of the optical wave packet leads to the optical Hall effect. Compared with this term, the other corrections due to $\boldsymbol{\Delta}_{n\boldsymbol{k}_c}$ are small as shown below (see also Appendix H). Thus an optimal design for the enhance-

ment of the optical Hall effect is equivalent to the enhancement of the magnitude of the Berry curvature. In the present case, the Berry curvature comes from an interband effect due to a periodic structure without inversion symmetry, and roughly scales as the inverse square of a band splitting [see Eqs. (E3a) and (E3b)]. Therefore we can expect the enhancement of the optical Hall effect for wave packets constructed from Bloch waves around nearly degenerate points in the Brillouin zone. In two-dimensional photonic crystals, Bloch waves propagating along the xy plane ($k_z=0$) are classified into the transverse magnetic (TM) and the transverse electric (TE) modes. In other words, the Maxwell equations for the Bloch functions (29a)–(29c) decouple into two sets of equations, one for the TM and the other for the TE modes, and the problem of wave-packet dynamics can be more simplified. Appendix G gives useful formulas for the Berry curvature $\mathbf{\Omega}_{nk}$ and $\mathbf{\Delta}_{nk}$ in such modes. As for other modes and a more generic case with degenerate bands, we must use formulas given in Appendix E.

We present examples of the Berry curvatures and the internal rotations of nondegenerate bands in the two-dimensional photonic crystal with $\mu=\mu_0$ and

$$\epsilon^{-1}(\boldsymbol{r}) = \frac{4}{3(5+12|\xi|+8\xi^2)}\sum_{i=1}^{3}\left\{\left[\xi-\cos\left(\boldsymbol{b}_i\cdot\boldsymbol{r}+\frac{2\pi}{3}\right)\right]^2 + \left[\xi+\cos\left(\boldsymbol{b}_i\cdot\boldsymbol{r}-\frac{2\pi}{3}\right)\right]^2\right\}, \tag{76}$$

where $\boldsymbol{b}_1=\left(\frac{2\pi\sqrt{3}}{3a},-\frac{2\pi}{3a}\right)$, $\boldsymbol{b}_2=\left(0,\frac{4\pi}{3a}\right)$, $\boldsymbol{b}_3=-\boldsymbol{b}_1-\boldsymbol{b}_2$, and a is the lattice constant. It is noted that, for $0<|\xi|<1$, ξ represents the degree of inversion-symmetry breaking. The spatial distribution of $\epsilon(\boldsymbol{r})$ and the band structure of TM and TE modes are shown in Figs. 3(a) and 3(b), respectively.

Figure 4 shows the Berry curvatures and the internal rotations of the first and second bands of TM and TE modes. The internal rotation of an optical wave packet is defined by $(z_c|\boldsymbol{\mathcal{S}}_{nk_c}|z_c)=\langle W|\boldsymbol{\mathcal{J}}|W\rangle-\boldsymbol{r}_c\times E_{nk_c}\boldsymbol{\nabla}_{k_c}E_{nk_c}$, where $\langle W|\boldsymbol{\mathcal{J}}|W\rangle$ is the total rotation of energy current and the second term represents the rotation of the center of gravity (see Appendixes C, E, and G). We can clearly see the correlation between them in each band except for their relative sign. The relative signs are roughly determined by a factor $\delta E=(E_{\mathrm{TM(TE)}n\boldsymbol{k}}-E_{\mathrm{TM(TE)}m\boldsymbol{k}})$ at nearly degenerate points $\boldsymbol{k}$, where n and m represent band indices of nearly degenerate bands. This is because the Berry curvature and the internal rotation are proportional to $1/\delta E^2$ and $1/\delta E$, respectively (see Appendix G). It is expected that this internal rotation is closely related to a physical angular momentum. Therefore these results suggest that we can generate a photonic mode with angular momentum by using a photonic crystal without inversion symmetry.

Before considering the motion of wave packets in this photonic crystal, we should comment on $\mathbf{\Delta}_{\mathrm{TM}nk}$ and $\mathbf{\Delta}_{\mathrm{TE}nk}$, which give corrections to energy dispersions and group velocities of TM and TE modes. Because $\mu(\boldsymbol{r})=\mu_0$ in the present case, it follows from Eq. (G8a) in Appendix G that $\mathbf{\Delta}_{\mathrm{TM}nk}=0$. Thus when a modulation is applied only to the

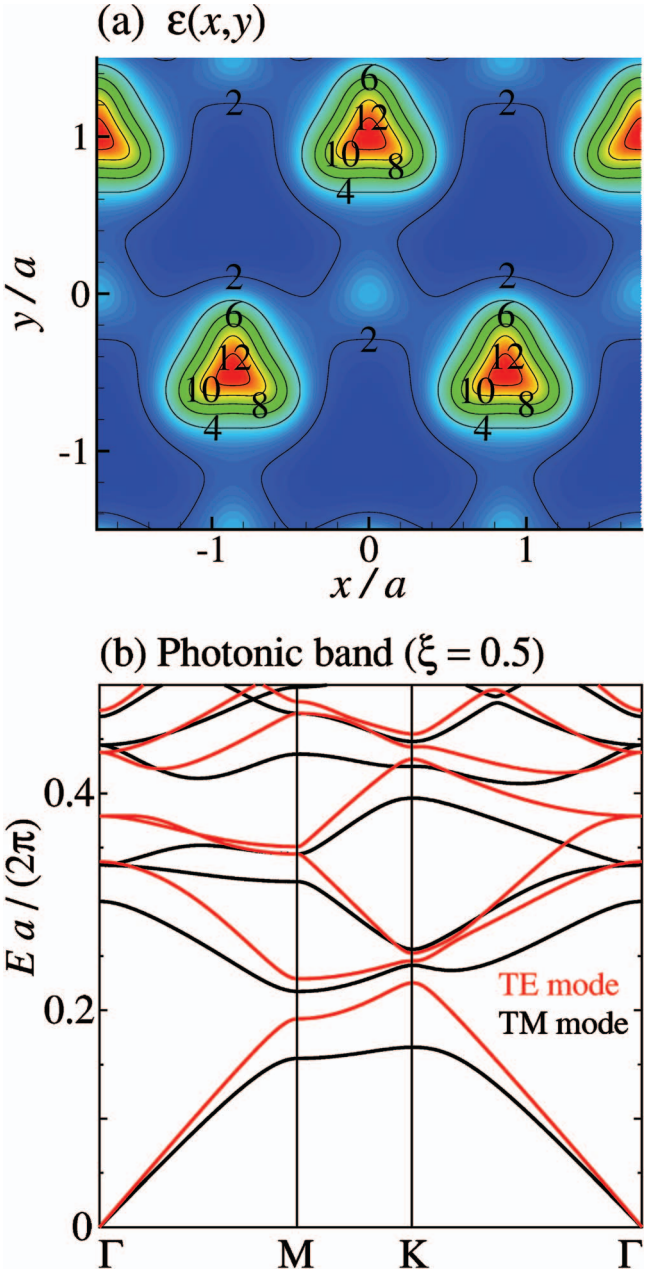

FIG. 3. (Color) (a) Dielectric function and (b) band structure of a two-dimensional photonic crystal. The Brillouin zone is shown in Figs. 4 and 5(b).

dielectric permittivity as $1/\epsilon(\boldsymbol{r})\to\gamma_\epsilon^2(\boldsymbol{r})/\epsilon(\boldsymbol{r})$, the energy of the TM mode is just rescaled by the factor $\gamma_\epsilon(\boldsymbol{r}_c)$, i.e., $E_{\mathrm{TM}nk_c}\to E_{\mathrm{TM}nk_c;\boldsymbol{r}_c}=\gamma_\epsilon(\boldsymbol{r}_c)E_{\mathrm{TM}nk_c}$. On the other hand, $\mathbf{\Delta}_{\mathrm{TE}nk}$ is nonzero. From Eq. (48), additional corrections appear in the energy dispersions of TE modes. However, as shown in Appendix H, we can see $|\mathbf{\Delta}_{\mathrm{TE}nk}|\lesssim 0.1a$. Thus these corrections are estimated to be at most a few percent as long as the modulation is sufficiently weak, i.e., $|a\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln\gamma_\epsilon(\boldsymbol{r}_c)|\ll 1$. In the similar argument, we can also neglect corrections to the group velocities of TE modes compared to their anomalous velocities, at least in the present photonic crystal. All the details of this issue are given in Appendix H.

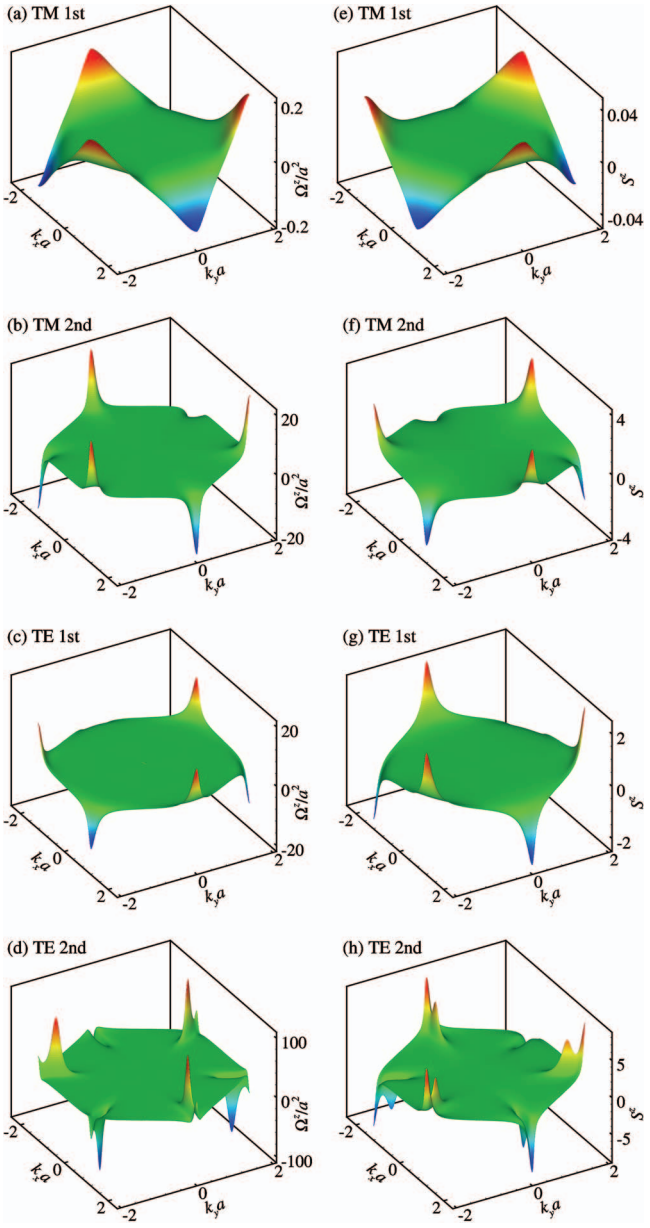

FIG. 4. (Color) Berry curvatures (a)–(d) and the internal rotations (e)–(h) of the first and second bands of TM and TE modes in the two-dimensional photonic crystal (ξ=0.5).

Now we consider the motions of wave packets constructed from TM and TE modes. It is noted that these wave packets are extended in the z direction, becaus e the z components of their momentum are fixed as k_z=0. From Fig. 4, we can see that $\mathbf{\Omega}_{\boldsymbol{k}}$ is strongly enhanced near the corners of the Brillouin zone. This enhancement is interpreted as a two-dimensional cut of the monopole structure in an extended space including parameters [9], e.g., ξ in the present case. Therefore we set the initial $\boldsymbol{k}_c$ near the corners of the Brillouin zone in order to make the effect of anomalous velocity prominent. We superimpose the following modulation onto the periodic structure, $1/\epsilon(\boldsymbol{r}) \to \gamma_\epsilon^2(\boldsymbol{r})/\epsilon(\boldsymbol{r})$,

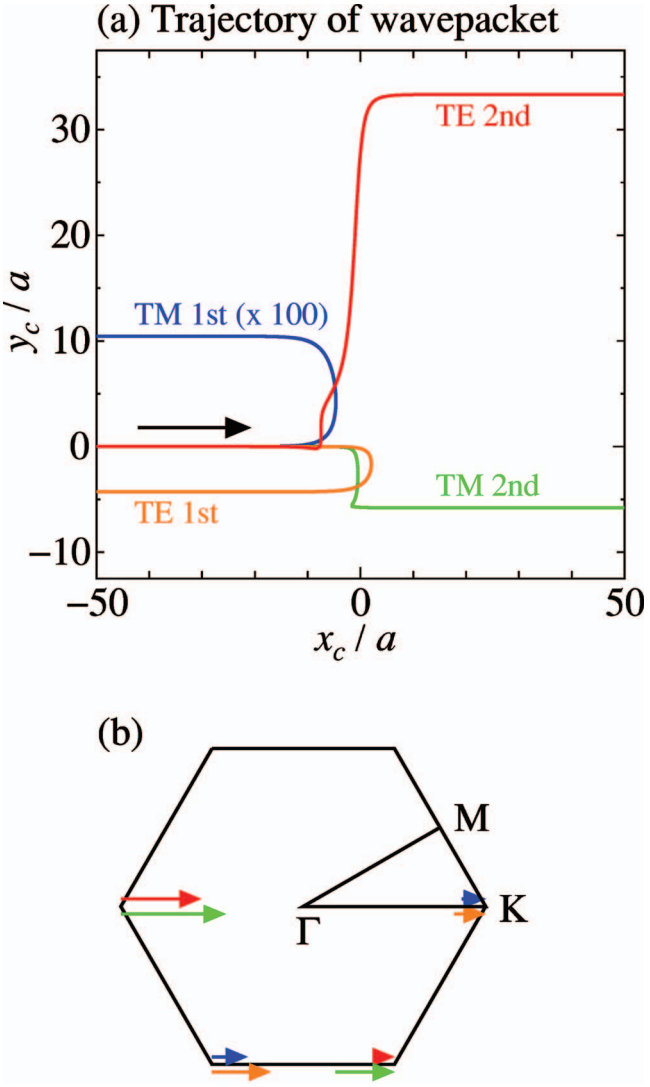

FIG. 5. (Color) Trajectories of wave packets in (a) real and (b) momentum spaces. The color of each arrow in (b) corresponds to that of each line in (a). The momentum-space trajectories in the figure are drawn with appropriate shifts from their actual ones which are on the line of k_y=0 or the horizontal Brillouin-zone boundary.

$$\frac{1}{\gamma_\epsilon(\boldsymbol{r})} = \frac{1}{2}\left[(\tilde{n}+1) + (\tilde{n}-1)\tanh\frac{x}{w}\right], \tag{77}$$

where $\tilde{n}>0$ represents a relative refractive index multiplied to the periodic structure in the region $x\to\infty$, and w is the mean width of the modulation. Here we take $\tilde{n}$=1.2 and w=5a which satisfy the condition of weak and slowly varying modulation. The obtained trajectories are shown in Fig. 5. It is found that the shift of $\boldsymbol{r}_c$ reaches to dozens of times the lattice constant especially for the wave packet constructed from the TE second band.

Finally we note that, also in more generic systems than discussed above, this effect can be enhanced considerably by designing crystal structures. The Berry curvature around a nearly degenerate point is determined mostly by the splitting, $2|m_g|$, between neighboring bands. Note that the sign of m_g

depends on details of wave functions around the nearly degenerate point, while its magnitude is determined only by the splitting. Suppose that at $\boldsymbol{k}=\boldsymbol{k}_0$ another band comes very close in energy to the one considered. The Berry curvature around $\boldsymbol{k}_0$ is evaluated as

$$\Omega_z \sim \frac{v^2 m_g}{(v^2|\boldsymbol{k}-\boldsymbol{k}_0|^2+m_g^2)^{3/2}}, \tag{78}$$

where v is a nominal velocity around $\boldsymbol{k}_0$. Thus when the light traverses near $\boldsymbol{k}_0$, the shift is estimated as

$$\delta y_c \sim \frac{v m_g}{v^2\kappa^2+m_g^2}\,\mathrm{sgn}[\nabla_{x_c}\gamma(x_c)], \tag{79}$$

where κ is the minimum value of $|\boldsymbol{k}-\boldsymbol{k}_0|$ when $\boldsymbol{k}$ traverses near $\boldsymbol{k}_0$. Therefore in the case of $\kappa=0$, the shift is larger for smaller $|m_g|$. This argument gives an intelligent explanation of a relation among relative magnitudes and signs of the shifts in Fig. 5.

IV. DISCUSSION

We have presented in detail the derivation of the equations of motion for an optical wave packet within the unitary theory. In our formalism, the equations are derived in the same fashion as those of electronic systems, thereby the similarities between them are evident. This suggests that the topological Hall effect driven by the geometrical mechanism is a broad concept ranging over various areas of physics such as electronic, photonic, acoustic, hydrodynamic, and relativistic phenomena. For example, in Refs. [29–31], the optical Hall effect in an optical fiber is referred to as "the optical Magnus effect" in analogy with the Magnus effect, which is a transverse aerodynamic effect on rotating objects. On the other hand, the relation of this effect to the geometrical effect on a spinning particle in general relativity was recently pointed out in Ref. [58], where equations of motion similar to ours are derived by considering the motion of a spinning particle in a space with a metric $g_{ij}(\boldsymbol{r})=n^2(\boldsymbol{r})\delta_{ij}$. This argument reminds us that, in the early stage of the study on general relativity, Einstein had tried to formulate the theory by generalizing the speed of light in vacuum. This issue might be related to the deep question about the dual nature between force and velocity in relativistic dynamics.

Here we should mention the effects called the photonic Hall effect [39,40] and the magnetically induced deflection due to the Pitaevskii magnetization [41–46], both of which are observed in Faraday-active media subject to external magnetic fields. The former effect takes place in a random medium, and is theoretically interpreted by the magnetically induced off-diagonal components of a diffusion tensor [39] and experimentally proved to be due to the magnetically induced changes in the optical properties of scatterers [40]. The latter effect is observed in a homogeneous medium [42], and is interpreted by the magnetically induced change in the dispersion relation of each mode due to the Pitaevskii magnetization [41]. (Additional remarks on these effects are given in Appendix I.) On the other hand, the optical Hall effect is caused by the anomalous velocity due to the geometrical propriety of a wave packet, which appears without external magnetic field nor scatterers.

As shown in Sec. III C, there is a close relation between the Berry curvature and the internal rotation. (See also Appendix E for details.) It is physically expected that an internal rotation can be related to an internal angular momentum. From this viewpoint, Laguerre-Gauss beams, which have internal orbital angular momenta [59], are of a particular interest. The Imbert-Fedorov effect is expected for these beams as well as for circularly polarized beams, and theoretical and experimental investigations on this problem has been done recently [60–62]. In Sec. III C, we have shown that there appear photonic modes with internal rotations in a two-dimensional photonic crystal without inversion symmetry. An angular momentum corresponding to this kind of internal rotation would be detected by measuring a torque working on a photonic crystal when we inject a linearly polarized light into the crystal (through a buffer layer if needed). In addition, when a photonic crystal is composed of Faraday-active media and subject to an external magnetic field, there would take place the magnetically induced deflection due to the Pitaevskii magnetization caused by this kind of generic internal rotation, rather than by the spin of circular polarization.

Last, we make a remark on the relevance of the quantum nature in the geometrical and topological properties discussed in this paper. Although we have formulated the theory of an optical wave packet in the quantum-mechanical formalism in order to clarify its connection to that of an electron wave packet, the phenomenon itself is based on the duality between position and momentum which is common in wave dynamics. Therefore the topological Hall effect is a generic one in both quantum mechanics and classical wave dynamics (see the argument in Appendix A). Actually, we can extend the argument on the enhancement of the optical Hall effect to other kinds of wave-packet dynamics, e.g., dynamics of the sonic wave packet which is mentioned in Ref. [30]. The sonic Hall effect in phononic crystals [63,64] could be enhanced in the same manner as the optical Hall effect in photonic crystals. It should be noted that the spin of a constituent particle is not always necessary; even a scalar wave can also have an internal rotation and a Berry curvature due to a periodic structure breaking inversion and/or time-reversal symmetries.

APPENDIX A: QUANTUM OR CLASSICAL?

In this paper, we employ a quantum-mechanical formalism in order to formulate our theory of a photonic system on an equal footing with that of an electronic system. The photonic system is described by an effective model where a dielectric medium is regarded as a classical object. This quantization procedure of photons is corresponding to that of electrons described by an effective model, e.g., a model with an effective mass-matrix and/or an effective (periodic) potential. In an electronic system, an electron is treated as a quantum object, and the equations of motion for a semiclassical wave packet are derived from an effective Lagrangian. In a photonic system, the counterpart of this effective Lagrangian

is most naturally represented in a second-quantized formalism with keeping its close connection to that in an electronic system. This is because, in the Maxwell theory, we cannot define a positive-definite probability density, while we can define a positive-definite energy density. (The second quantization is adopted to define the quantum wave function of a photon in Sec. II C, which cannot be directly represented by the field strength of an electromagnetic field.) However, as long as we consider an approximately coherent wave packet, the center of wave packet coincides with the center of gravity as shown in Sec. II C. This fact enables us to link our quantum-mechanical theory and classical electrodynamics.

One may wonder whether the optical Hall effect is quantum or classical. We cannot answer this question in a single sentence, but give some remarks on it as follows. This effect comes from the particle-wave duality of an optical wave packet and the geometrical or topological property of a wave function. Therefore similar geometrical or topological effects are expected in various kinds of quantum or classical and microscopic or macroscopic wave dynamics, when wave packets under consideration are approximately coherent. Indeed, we can formulate a theory for this class of phenomena based on a classical wave dynamics of a macroscopic system, while its direct connection to electronic systems is not necessarily clear. The confusion represented by the above question is mainly due to the situation in which we sometimes refer to wave equations for photonic systems as classical Maxwell equations and those for electronic systems as quantum Schrödinger equations, while both photons and electrons are quantum objects. The origin of this situation comes from the following two advantages of photonic systems which lead to the success of classical Maxwell theory. In contrast to electronic systems, (i) the statistics of photons is bosonic, and (ii) effective self-interactions between photons are usually very weak. However, as long as we treat a photon and an electron in a single particle approximation, we can formulate both theories on an equal footing.

It is beyond the scope of this paper to fix the terminologies "quantum" and "classical" common in photonic and electronic systems. Although we treat a photon as a quantum object through to the end of this paper, we refer to results obtained purely by wave dynamics of light as those obtained by classical electrodynamics. As long as we consider an approximately coherent wave packet in a single particle approximation of quantum theory of a photon or in a linear approximation of classical electrodynamics, results obtained by both formalisms coincide with each other as shown in Sec. III A.

APPENDIX B: ORTHONORMALITY OF EIGENFUNCTIONS

The orthonormality, Eqs. (33a) and (33b), is approved with the orthonomality of Bloch functions, Eqs. (23a) and (23b). For example, Eq. (33a) can be shown by using the following relation:

$$
\begin{aligned}
&(\Phi^E_{n\lambda \boldsymbol{k}}|\vec{\epsilon}|\Phi^E_{n'\lambda'\boldsymbol{k}'}) \\
&= (E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}'}) \int d\boldsymbol{r}\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\vec{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t) \\
&= \frac{E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}'}}{2\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}'}}} e^{i(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'})t} \\
&\quad \times \int d\boldsymbol{r} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} \boldsymbol{U}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\vec{\epsilon}(\boldsymbol{r})\boldsymbol{U}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r}) \\
&= \frac{E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}'}}{2\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}'}}} e^{i(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'})t} \\
&\quad \times \sum_{\boldsymbol{a}} \int_{\mathrm{WS}} d\boldsymbol{r} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{a}+\boldsymbol{r})} \boldsymbol{U}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\vec{\epsilon}(\boldsymbol{r})\boldsymbol{U}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r}) \\
&= \frac{E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}'}}{2\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}'}}} e^{i(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'})t} \sum_{\boldsymbol{G}} \tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G}) \\
&\quad \times \int_{\mathrm{WS}} \frac{d\boldsymbol{r}}{v_{\mathrm{WS}}} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} \boldsymbol{U}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\vec{\epsilon}(\boldsymbol{r})\boldsymbol{U}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r}), \qquad \text{(B1)}
\end{aligned}
$$

where $\boldsymbol{a}$ represents an arbitrary lattice vector. Since the lattice momentum $\boldsymbol{k}$ and $\boldsymbol{k}'$ are in the first Brillouin zone, we can reach the result

$$
\begin{aligned}
&(\Phi^E_{n\lambda\boldsymbol{k}}|\vec{\epsilon}|\Phi^E_{n'\lambda'\boldsymbol{k}'}) \\
&= \frac{E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}}}{2\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}}}} e^{i(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}})t} \tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')\langle U^E_{n\lambda\boldsymbol{k}}|\vec{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}}\rangle \\
&= \delta_{nn'}\delta_{\lambda\lambda'}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'), \qquad \text{(B2)}
\end{aligned}
$$

where we have used Eq. (23a). In the same manner, Eq. (33b), can be also proved by using Eq. (23b).

Next we prove the orthogonality, Eq. (33c). From the definition of the inner product, Eq. (32), we can show

$$
\begin{aligned}
(\Phi^{E*}_{n\lambda\boldsymbol{k}}|\vec{\epsilon}|\Phi^E_{n'\lambda'\boldsymbol{k}'}) &= (E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}}) \\
&\quad \times \int d\boldsymbol{r}\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\vec{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t), \qquad \text{(B3a)}
\end{aligned}
$$

$$
\begin{aligned}
(\Phi^{H*}_{n\lambda\boldsymbol{k}}|\vec{\mu}|\Phi^H_{n'\lambda'\boldsymbol{k}'}) &= (E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}}) \\
&\quad \times \int d\boldsymbol{r}\boldsymbol{\Phi}^H_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\vec{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t). \qquad \text{(B3b)}
\end{aligned}
$$

In the case of $E_{n\boldsymbol{k}}=E_{n'\boldsymbol{k}'}$, it is clear that Eq. (33c) is approved. Thus in what follows, we consider the case of $E_{n\boldsymbol{k}} \neq E_{n'\boldsymbol{k}'}$. In this case, we can easily show $\int d\boldsymbol{r}\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\vec{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t)=0$ and $\int d\boldsymbol{r}\boldsymbol{\Phi}^H_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\vec{\mu}(\boldsymbol{r}) \times \boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t)=0$ from the relations

$$(E^2_{n'k'} - E^2_{nk}) \int d\boldsymbol{r}\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t) = 0, \tag{B4a}$$

$$(E^2_{n'k'} - E^2_{nk}) \int d\boldsymbol{r}\boldsymbol{\Phi}^H_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t) = 0. \tag{B4b}$$

Consequently, the orthogonality, Eq. (33c) is approved in all cases.

The above relations are derived from the eigenequations, Eqs. (22a) and (22b). For example, the relation for $\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)$ is proved as

$$\begin{aligned} E^2_{n'k'} \int d\boldsymbol{r}\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t) &= \int d\boldsymbol{r}\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\cdot\{\boldsymbol{\nabla}_{\boldsymbol{r}}\times[\overleftrightarrow{\mu}^{-1}(\boldsymbol{r})\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t)]\} \\ &= \int d\boldsymbol{r}\{\boldsymbol{\nabla}_{\boldsymbol{r}}\times[\overleftrightarrow{\mu}^{-1}(\boldsymbol{r})\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)]\}\cdot\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t) \\ &= E^2_{nk}\int d\boldsymbol{r}\boldsymbol{\Phi}^E_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r},t), \end{aligned} \tag{B5}$$

where $\overleftrightarrow{\mu}^T(\boldsymbol{r}) = \overleftrightarrow{\mu}(\boldsymbol{r})$ is used in the transformation from the second line to the third line, and $\overleftrightarrow{\epsilon}^T(\boldsymbol{r}) = \overleftrightarrow{\epsilon}(\boldsymbol{r})$ is used in the transformation from the third line to the fourth line. A similar relation can be derived also for $\boldsymbol{\Phi}^H_{n\lambda\boldsymbol{k}}(\boldsymbol{r},t)$.

APPENDIX C: WAVE PACKET IN A PERIODIC SYSTEM

Here we present details about an optical wave packet in a periodic system. Basic features of the wave packet are discussed in Appendix C 1. These features are helpful to understand the effect of an additional modulation superimposed onto a periodic structure, which is discussed in Sec. II C. Some comments on a gauge transformation in momentum space are given in Appendix C 2. In Appendix C 3, we present detailed procedures to evaluate expectation values which appear in Appendix C 1.

1. Wave packet

We begin with the wave packet defined by

$$|W\rangle = \int_{\mathrm{BZ}} d\boldsymbol{k} w(\boldsymbol{k},\boldsymbol{k}_c,t)\sum_{\lambda} z_{c\lambda} a^{\dagger}_{n\lambda\boldsymbol{k}}|0\rangle, \tag{C1a}$$

$$w(\boldsymbol{k},\boldsymbol{k}_c,t) = w_r(\boldsymbol{k}-\boldsymbol{k}_c)e^{-i\vartheta(\boldsymbol{k},t)}, \tag{C1b}$$

where $w_r(\boldsymbol{k}-\boldsymbol{k}_c)$ is a real function, and $w_r(\boldsymbol{k}-\boldsymbol{k}_c)$ and $z_{c\lambda}$ satisfy the normalization conditions, $\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)=1$ and $\Sigma_\lambda |z_{c\lambda}|^2=1$, respectively. We assume $w_r(\boldsymbol{k}-\boldsymbol{k}_c)$ has a sharp peak around $\boldsymbol{k}_c = \int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\boldsymbol{k}$. It should be noted that, rigorously speaking, we need to replace this single photon wave packet with a coherent (or squeezed) state wave packet when we apply the present formalism to a light beam with a macroscopic number of photons. However, from the linearity of the Maxwell equations, the equations of motion for the single photon wave packet is applicable also to the macroscopic coherent beam.

In a fermionic system, we can define the position operator as the center of the probability density of a fermion, which is positive-definite both in nonrelativistic and relativistic cases. However, for a relativistic boson, the definition of its position is nontrivial. In order to find an appropriate definition for the position of wave packet, we first consider the energy and the position weighted by the energy density evaluated as follows:

$$\langle W|H|W\rangle \cong E_{nk_c}, \tag{C2a}$$

$$\langle W|\boldsymbol{\mathcal{R}}|W\rangle \cong E_{nk_c}[\boldsymbol{\nabla}_{\boldsymbol{k}_c}\vartheta(\boldsymbol{k}_c,t) - (z_c|\boldsymbol{\Lambda}_{nk_c}|z_c)]. \tag{C2b}$$

It should be noted that $\cong$ in Eqs. (C2a) and (C2b) means that the above expectation values are evaluated under the assumption that the shape of $w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)$ is sufficiently sharp compared to the variations of E_{nk} and $\boldsymbol{\Lambda}_{nk}(\boldsymbol{k})$ around $\boldsymbol{k}_c$, and we neglected terms which depend on the shape of $w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)$.

From Eqs. (C2a) and (C2b), the center of gravity is estimated as

$$\frac{\langle W|\boldsymbol{\mathcal{R}}|W\rangle}{\langle W|H|W\rangle} \cong \boldsymbol{\nabla}_{\boldsymbol{k}_c}\vartheta(\boldsymbol{k}_c,t) - (z_c|\boldsymbol{\Lambda}_{nk_c}|z_c). \tag{C3}$$

Comparing this result with the naive definition for the position of wave packet, $\int_{\mathrm{BZ}} \frac{d\boldsymbol{k}}{(2\pi)^3} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},t)$, we can reach the appropriate definition for the position of wave packet,

$$\boldsymbol{r}_c = \int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},t) - (z_c|\boldsymbol{\Lambda}_{nk}|z_c)]. \tag{C4}$$

In order to check the property of the wave packet, we consider the expectation values of physical observables. As shown in Appendix C 3, the energy current and the rotation of energy current are evaluated as

$$\langle W|\boldsymbol{\mathcal{P}}|W\rangle \cong E_{nk_c}\boldsymbol{\nabla}_{\boldsymbol{k}_c}E_{nk_c}, \tag{C5a}$$

$$\langle W|\boldsymbol{\mathcal{J}}|W\rangle \cong \boldsymbol{r}_c \times E_{nk_c}\boldsymbol{\nabla}_{\boldsymbol{k}_c}E_{nk_c} + (z_c|\boldsymbol{\mathcal{S}}_{nk_c}|z_c), \tag{C5b}$$

where

$$\boldsymbol{\mathcal{S}}_{nk} = \frac{1}{2}[\boldsymbol{\mathcal{S}}^E_{nk} + \boldsymbol{\mathcal{S}}^H_{nk}], \tag{C6a}$$

$$\begin{aligned} [\boldsymbol{\mathcal{S}}^E_{nk}]_{\lambda\lambda'} = -\frac{i}{2}[&\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{n\lambda\boldsymbol{k}}|\times(\overleftrightarrow{\epsilon}E^2_{nk} - \Xi^E_{\boldsymbol{k}})|\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{n\lambda'\boldsymbol{k}}\rangle \\ &+ \langle U^E_{n\lambda\boldsymbol{k}}|\boldsymbol{S}\times\overleftrightarrow{\mu}^{-1}\boldsymbol{S}|U^E_{n\lambda'\boldsymbol{k}}\rangle], \end{aligned} \tag{C6b}$$

$$[\boldsymbol{\mathcal{S}}^H_{nk}]_{\lambda\lambda'} = -\frac{i}{2}[\langle \boldsymbol{\nabla}_k U^H_{n\lambda k}| \times (\overleftrightarrow{\mu} E^2_{nk} - \Xi^H_k)|\boldsymbol{\nabla}_k U^H_{n\lambda' k}\rangle + \langle U^H_{n\lambda k}|\boldsymbol{S} \times \overleftrightarrow{\epsilon}^{-1}\boldsymbol{S}|U^H_{n\lambda' k}\rangle]. \quad \text{(C6c)}$$

It is noted that the first term of $\langle W|\boldsymbol{\mathcal{J}}|W\rangle$ is interpreted as the orbital rotational motion, i.e., the rotation of the center of gravity, and the second term as the internal one, i.e., the rotation around the center of gravity. Especially for the locally isotropic system in which $\overleftrightarrow{\epsilon}(\boldsymbol{r})$ and $\overleftrightarrow{\mu}(\boldsymbol{r})$ are scalar variables, $\epsilon(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$, the contribution from the second terms in $\boldsymbol{\mathcal{S}}^E_{nk}$ and $\boldsymbol{\mathcal{S}}^H_{nk}$ are rewritten by using $\boldsymbol{S}\times\boldsymbol{S}=i\boldsymbol{S}$ as

$$-\frac{i}{4}[\langle U^E_{n\lambda k}|\boldsymbol{S} \times \overleftrightarrow{\mu}^{-1}\boldsymbol{S}|U^E_{n\lambda' k}\rangle + \langle U^H_{n\lambda k}|\boldsymbol{S} \times \overleftrightarrow{\epsilon}^{-1}\boldsymbol{S}|U^H_{n\lambda' k}\rangle] \to \frac{1}{4}[\langle U^E_{n\lambda k}|\mu^{-1}\boldsymbol{S}|U^E_{n\lambda' k}\rangle + \langle U^H_{n\lambda k}|\epsilon^{-1}\boldsymbol{S}|U^H_{n\lambda' k}\rangle]. \quad \text{(C7)}$$

This suggests that the internal rotation correctly includes the spin of a constituent particle, i.e., the polarization of a light in the present case. However, it should be noted that the second terms on the right-hand side of Eqs. (C6b) and (C6c) are not the whole contributions of spin. Actually, when we consider the circularly polarized light in isotropic homogeneous media, all terms of the internal rotation give the same contribution and totally represent the rotation originated by the polarization. In addition, the internal rotation defined above contains the internal orbital one and the spin one generally.

2. Gauge transformation

When a system has a symmetry represented by the unitary matrix $[M_{nk}]_{\lambda\lambda'}$, Maxwell equations are invariant under the transformation

$$|\tilde{U}^F_{n\lambda k}\rangle = \sum_{\lambda'} [M_{nk}]_{\lambda'\lambda}|U^F_{n\lambda' k}\rangle, \quad \text{(C8)}$$

where $F=E$ or H. (Here we consider the case in which there are degeneracies indexed by the subscript λ or λ'.) For the sake of convenience, we call this transformation as the gauge transformation in $\boldsymbol{k}$ space. By this gauge transformation, $\boldsymbol{\Lambda}_{nk}$, $\boldsymbol{\Omega}_{nk}$, $\boldsymbol{\Delta}_{nk}$, and $\boldsymbol{\mathcal{S}}_{nk}$ are transformed as

$$\tilde{\boldsymbol{\Lambda}}_{nk} = M^{-1}_{nk}\boldsymbol{\Lambda}_{nk}M_{nk} - iM^{-1}_{nk}\boldsymbol{\nabla}_k M_{nk}, \quad \text{(C9a)}$$

$$\tilde{\boldsymbol{\Omega}}_{nk} = M^{-1}_{nk}\boldsymbol{\Omega}_{nk}M_{nk}, \quad \text{(C9b)}$$

$$\tilde{\boldsymbol{\Delta}}_{nk} = M^{-1}_{nk}\boldsymbol{\Delta}_{nk}M_{nk}, \quad \text{(C9c)}$$

$$\tilde{\boldsymbol{\mathcal{S}}}_{nk} = M^{-1}_{nk}\boldsymbol{\mathcal{S}}_{nk}M_{nk}. \quad \text{(C9d)}$$

The gauge transformation of Bloch functions in Eq. (C8) is equivalent to that of the corresponding creation operators as

$$\tilde{a}^\dagger_{n\lambda k} = \sum_{\lambda'} [M_{nk}]_{\lambda'\lambda}a^\dagger_{n\lambda' k}. \quad \text{(C10)}$$

In terms of these transformed operators, the wave packet in Eq. (C1a) is represented by

$$|W\rangle = \int_{\mathrm{BZ}} d\boldsymbol{k} w(\boldsymbol{k},\boldsymbol{k}_c,t)\sum_{\lambda,\lambda'} [M^{-1}_{nk}]_{\lambda'\lambda}z_{c\lambda}\tilde{a}^\dagger_{n\lambda' k}|0\rangle. \quad \text{(C11)}$$

It should be noted that we have changed only the representation but not the physical state of the wave packet. Therefore the expectation values of physical observables, e.g., H, $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\mathcal{P}}$, and $\boldsymbol{\mathcal{J}}$, must be gauge invariant. Indeed, we can easily show that the evaluations of H and $\boldsymbol{\mathcal{P}}$ in Appendix C 1 are gauge invariant because of the invariance of E_{nk}. From Eqs. (C9b)–(C9d), and Eq. (C11) we can also show the invariance of $(z_c|\boldsymbol{\Omega}_{nk}|z_c)$, $(z_c|\boldsymbol{\Delta}_{nk}|z_c)$, and $(z_c|\boldsymbol{\mathcal{S}}_{nk}|z_c)$. However, it is not clear whether the evaluations of $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{J}}$ given in Appendix C 1 are also the case. In order to confirm this point, it is enough to check whether the position of wave packet $\boldsymbol{r}_c$ in Eq. (C4) is gauge invariant or not. In the representation of Eq. (C11), the derivative of phase factor $\boldsymbol{\nabla}_k\vartheta(\boldsymbol{k},t)$ and $(z_c|\boldsymbol{\Lambda}_{nk}|z_c)$ in Appendix C 1 are replaced as

$$\boldsymbol{\nabla}_k\vartheta(\boldsymbol{k},t) = i(z_c|e^{i\vartheta(\boldsymbol{k},t)}\boldsymbol{\nabla}_k e^{-i\vartheta(\boldsymbol{k},t)}|z_c) \to i(z_c|[e^{i\vartheta(\boldsymbol{k},t)}M_{nk}]\boldsymbol{\nabla}_k[e^{-i\vartheta(\boldsymbol{k},t)}M^{-1}_{nk}]|z_c), \quad \text{(C12a)}$$

$$(z_c|\boldsymbol{\Lambda}_{nk}|z_c) \to (z_c|M_{nk}\tilde{\boldsymbol{\Lambda}}_{nk}M^{-1}_{nk}|z_c). \quad \text{(C12b)}$$

The above formulas and Eq. (C9a) prove the gauge invariance of $\boldsymbol{r}_c$ as follows:

$$\begin{aligned}\tilde{\boldsymbol{r}}_c &= \int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\{i(z_c|[e^{i\vartheta(\boldsymbol{k},t)}M_{nk}]\boldsymbol{\nabla}_k[e^{-i\vartheta(\boldsymbol{k},t)}M^{-1}_{nk}]|z_c) - (z_c|M_{nk}\tilde{\boldsymbol{\Lambda}}_{nk}M^{-1}_{nk}|z_c)\} \\ &= \boldsymbol{r}_c + \int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c) \times i(z_c|[M_{nk}(\boldsymbol{\nabla}_k M^{-1}_{nk}) + (\boldsymbol{\nabla}_k M_{nk})M^{-1}_{nk}]|z_c) \\ &= \boldsymbol{r}_c. \end{aligned} \quad \text{(C13)}$$

Combining this result and the gauge invariance of E_{nk} and $(z_c|\boldsymbol{\mathcal{S}}_{nk}|z_c)$, we can confirm that the evaluations of $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{J}}$ in Appendix C 1 are also gauge invariant.

3. Expectation values

Here we present the detailed evaluations of the expectation values, i.e., the Hamiltonian H, the position weighted by the energy density $\boldsymbol{\mathcal{R}}$, the energy current $\boldsymbol{\mathcal{P}}$, and the rotation of energy current $\boldsymbol{\mathcal{J}}$, with respect to the optical wave packet $|W\rangle$ in a periodic system. The expectation value of an operator $\mathcal{O}$ is obtained from commutation relations between $\mathcal{O}$ and the creation and annihilation operators,

$$\langle W|\mathcal{O}|W\rangle$$
$$= \int_{\mathrm{BZ}} d\boldsymbol{k}d\boldsymbol{k}' w^*(\boldsymbol{k},\boldsymbol{k}_c,t)w(\boldsymbol{k}',\boldsymbol{k}_c,t)\langle 0|a_{nz_c\boldsymbol{k}}\mathcal{O}a^\dagger_{nz_c\boldsymbol{k}'}|0\rangle$$
$$= \int_{\mathrm{BZ}} d\boldsymbol{k}d\boldsymbol{k}' w^*(\boldsymbol{k},\boldsymbol{k}_c,t)w(\boldsymbol{k}',\boldsymbol{k}_c,t)\langle 0|[a_{nz_c\boldsymbol{k}},[\mathcal{O},a^\dagger_{nz_c\boldsymbol{k}'}]]|0\rangle, \tag{C14}$$

where we have introduced the abbreviation

$$a^{(\dagger)}_{nz_c\boldsymbol{k}} = \sum_\lambda z_{c\lambda} a^{(\dagger)}_{n\lambda\boldsymbol{k}}, \tag{C15}$$

and this will be used also for the Bloch functions as

$$|U^{E,H}_{nz_c\boldsymbol{k}}\rangle = \sum_\lambda z_{c\lambda}|U^{E,H}_{n\lambda\boldsymbol{k}}\rangle. \tag{C16}$$

The basic commutation relation between $\boldsymbol{B}(\boldsymbol{r})$ and $\boldsymbol{D}(\boldsymbol{r})$ can be represented in the following integral form:

$$\int d\boldsymbol{r}d\boldsymbol{r}'[\boldsymbol{\Phi}^*_1(\boldsymbol{r})\cdot\boldsymbol{B}(\boldsymbol{r}),\boldsymbol{D}(\boldsymbol{r}')\cdot\boldsymbol{\Phi}_2(\boldsymbol{r}')]$$
$$= -i\int d\boldsymbol{r}[\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}^*_1(\boldsymbol{r})]\cdot\boldsymbol{\Phi}_2(\boldsymbol{r})$$
$$= -i\int d\boldsymbol{r}\boldsymbol{\Phi}^*_1(\boldsymbol{r})\cdot[\boldsymbol{\nabla}_{\boldsymbol{r}}\times\boldsymbol{\Phi}_2(\boldsymbol{r})]. \tag{C17}$$

In particular, in a periodic system, by the above commutation relation and Eqs. (21a) and (21b), we can easily show that

$$[a_{n\lambda\boldsymbol{k}},[H,a^\dagger_{n'\lambda'\boldsymbol{k}'}]] = E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}'}\int d\boldsymbol{r}[\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})$$
$$+\boldsymbol{\Phi}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})]$$
$$= E_{n\boldsymbol{k}}\delta_{nn'}\delta_{\lambda\lambda'}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'). \tag{C18}$$

In the transformation to the last line, we have used

$$\int d\boldsymbol{r}e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$$
$$= \sum_{\boldsymbol{a}}\int_{\mathrm{WS}} d\boldsymbol{r}e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{a}+\boldsymbol{r})}F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{a}+\boldsymbol{r})$$
$$= \int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}\sum_{\boldsymbol{G}}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r}), \tag{C19}$$

where $\boldsymbol{a}$ represents an arbitrary lattice vector, and $F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$ is a periodic function, i.e., $F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{a}+\boldsymbol{r})=F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$. Since $\boldsymbol{k}$ and $\boldsymbol{k}'$ are in the first Brillouin zone, we have also used the following relation implicitly:

$$\sum_{\boldsymbol{G}}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})\,e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}|_{\boldsymbol{k},\boldsymbol{k}'\in 1\mathrm{st\ BZ}} = \tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'). \tag{C20}$$

Then we obtain the result

$$\langle W|H|W\rangle = \int_{\mathrm{BZ}} d\boldsymbol{k}w^2_r(\boldsymbol{k}-\boldsymbol{k}_c)E_{n\boldsymbol{k}} \cong E_{n\boldsymbol{k}_c}. \tag{C21}$$

In a similar manner, we obtain the following commutation relation which is needed to estimate the expectation value of $\boldsymbol{\mathcal{R}}$:

$$[a_{n\lambda\boldsymbol{k}},[\boldsymbol{\mathcal{R}},a^\dagger_{n'\lambda'\boldsymbol{k}'}]] = E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}'}\int d\boldsymbol{r}\boldsymbol{r}[\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})$$
$$+\boldsymbol{\Phi}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})]$$
$$= \frac{i}{4}\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}'}}[(\boldsymbol{\nabla}_{\boldsymbol{k}}-\boldsymbol{\nabla}_{\boldsymbol{k}'})\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')]$$
$$\times[\langle U^E_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}'}\rangle + \langle U^H_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\mu}|U^H_{n'\lambda'\boldsymbol{k}'}\rangle]. \tag{C22}$$

In the transformation to the last line, we have used the relation

$$\int d\boldsymbol{r}\boldsymbol{r}e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$$
$$= \frac{i}{2}\int d\boldsymbol{r}[(\boldsymbol{\nabla}_{\boldsymbol{k}}-\boldsymbol{\nabla}_{\boldsymbol{k}'})e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}]F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$$
$$= \frac{i}{2}\sum_{\boldsymbol{a}}\int_{\mathrm{WS}} d\boldsymbol{r}[(\boldsymbol{\nabla}_{\boldsymbol{k}}-\boldsymbol{\nabla}_{\boldsymbol{k}'})e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{a}+\boldsymbol{r})}]F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$$
$$= \frac{i}{2}\int_{\mathrm{WS}}\frac{d\boldsymbol{r}}{v_{\mathrm{WS}}}F_{\boldsymbol{k}\boldsymbol{k}'}(\boldsymbol{r})$$
$$\times\left[(\boldsymbol{\nabla}_{\boldsymbol{k}}-\boldsymbol{\nabla}_{\boldsymbol{k}'})\sum_{\boldsymbol{G}}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'+\boldsymbol{G})e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\right], \tag{C23}$$

and Eq. (C20). This commutation relation leads to the result

$$\langle W|\boldsymbol{\mathcal{R}}|W\rangle = \frac{i}{2}\int_{\mathrm{BZ}} d\boldsymbol{k}E_{n\boldsymbol{k}}[w^*(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,t)\boldsymbol{\nabla}_{\boldsymbol{k}}w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,t)$$
$$-[\boldsymbol{\nabla}_{\boldsymbol{k}}w^*(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,t)]w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,t)]$$
$$+\frac{i}{4}\int_{\mathrm{BZ}} d\boldsymbol{k}w^2_r(\boldsymbol{k}-\boldsymbol{k}_c)E_{n\boldsymbol{k}}$$
$$\times[\langle U^E_{nz_c\boldsymbol{k}}|\overleftrightarrow{\epsilon}|\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{nz_c\boldsymbol{k}}\rangle - \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{nz_c\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{nz_c\boldsymbol{k}}\rangle$$
$$+\langle U^H_{nz_c\boldsymbol{k}}|\overleftrightarrow{\mu}|\boldsymbol{\nabla}_{\boldsymbol{k}}U^H_{nz_c\boldsymbol{k}}\rangle - \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^H_{nz_c\boldsymbol{k}}|\overleftrightarrow{\mu}|U^H_{nz_c\boldsymbol{k}}\rangle]$$
$$= \int_{\mathrm{BZ}} d\boldsymbol{k}w^2_r(\boldsymbol{k}-\boldsymbol{k}_c)E_{n\boldsymbol{k}}[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},t) - (z_c|\boldsymbol{\Lambda}_{\boldsymbol{k}}|z_c)]$$
$$\cong E_{n\boldsymbol{k}_c}[\boldsymbol{\nabla}_{\boldsymbol{k}_c}\vartheta(\boldsymbol{k}_c,t) - (z_c|\boldsymbol{\Lambda}_{\boldsymbol{k}_c}|z_c)]. \tag{C24}$$

The expectation value of $\boldsymbol{\mathcal{P}}$ is estimated by using the commutation relation

$$[a_{n\lambda k},[\mathcal{P},a^{\dagger}_{n'\lambda'k'}]] = E_{nk}E_{n'k'}\int d\boldsymbol{r}[\boldsymbol{\Phi}^{E*}_{n\lambda k}(\boldsymbol{r})\times\boldsymbol{\Phi}^{H}_{n'\lambda'k'}(\boldsymbol{r}) - \boldsymbol{\Phi}^{H*}_{n\lambda k}(\boldsymbol{r})\times\boldsymbol{\Phi}^{E}_{n'\lambda'k'}(\boldsymbol{r})]$$
$$= \frac{1}{2}\sqrt{E_{nk}E_{n'k}}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')[\langle U^{E}_{n\lambda k}|i\boldsymbol{S}|U^{H}_{n'\lambda'k}\rangle - \langle U^{H}_{n\lambda k}|i\boldsymbol{S}|U^{E}_{n'\lambda'k}\rangle], \tag{C25}$$

where Eqs. (C19) and (C20) have been used. Combining this commutation relation and the formula

$$E_{nk}[\langle U^{E}_{nz_ck}|i\boldsymbol{S}|U^{H}_{nz_ck}\rangle - \langle U^{H}_{nz_ck}|i\boldsymbol{S}|U^{E}_{nz_ck}\rangle]$$
$$= \langle U^{E}_{nz_ck}|[\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}+\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}]|U^{E}_{nz_ck}\rangle$$
$$= \langle U^{E}_{nz_ck}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^{E}_{\boldsymbol{k}}]|U^{E}_{nz_ck}\rangle = \boldsymbol{\nabla}_{\boldsymbol{k}}E^{2}_{n\boldsymbol{k}}, \tag{C26}$$

we obtain the result

$$\langle W|\mathcal{P}|W\rangle = \frac{1}{2}\int_{\mathrm{BZ}} d\boldsymbol{k}w^{2}_{r}(\boldsymbol{k}-\boldsymbol{k}_c)E_{nk}[\langle U^{E}_{nz_ck}|i\boldsymbol{S}|U^{H}_{nz_ck}\rangle - \langle U^{H}_{nz_ck}|i\boldsymbol{S}|U^{E}_{nz_ck}\rangle]$$
$$= \frac{1}{2}\int_{\mathrm{BZ}} d\boldsymbol{k}w^{2}_{r}(\boldsymbol{k}-\boldsymbol{k}_c)\boldsymbol{\nabla}_{\boldsymbol{k}}E^{2}_{n\boldsymbol{k}} \cong E_{nk_c}\boldsymbol{\nabla}_{\boldsymbol{k}_c}E_{nk_c}. \tag{C27}$$

The expectation value of $\mathcal{J}$ is derived from the commutation relation

$$[a_{n\lambda k},[\mathcal{J},a^{\dagger}_{n'\lambda'k'}]] = E_{nk}E_{n'k'}\int d\boldsymbol{r}\boldsymbol{r}\times[\boldsymbol{\Phi}^{E*}_{n\lambda k}(\boldsymbol{r})\times\boldsymbol{\Phi}^{H}_{n'\lambda'k'}(\boldsymbol{r}) - \boldsymbol{\Phi}^{H*}_{n\lambda k}(\boldsymbol{r})\times\boldsymbol{\Phi}^{E}_{n'\lambda'k'}(\boldsymbol{r})]$$
$$= \frac{i}{4}\sqrt{E_{nk}E_{n'k'}}[(\boldsymbol{\nabla}_{\boldsymbol{k}}-\boldsymbol{\nabla}_{\boldsymbol{k}'})\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')]\times[\langle U^{E}_{n\lambda k}|i\boldsymbol{S}|U^{H}_{n'\lambda'k'}\rangle - \langle U^{H}_{n\lambda k}|i\boldsymbol{S}|U^{E}_{n'\lambda'k'}\rangle], \tag{C28}$$

where we have used Eq. (C23) in the transformation to the last line. It should be noted that, as in the previous commutation relations, the above commutation relation is also restricted to the case in which both $\boldsymbol{k}$ and $\boldsymbol{k}'$ are in the first Brillouin zone. In addition, our wave packet is constructed from degenerate eigenmodes, i.e., eigenmodes with the same band index n. Thus the following formula, which will be proved later, is useful to estimate the expectation value of $\mathcal{J}$ with respect to the wave packet:

$$\frac{E_{nk}}{4}[\langle U^{E}_{nz_ck}|\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times\boldsymbol{S}|U^{H}_{nz_ck}\rangle - \langle U^{H}_{nz_ck}|\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle - \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times\boldsymbol{S}|U^{E}_{nz_ck}\rangle]$$
$$= -(z_c|\boldsymbol{\Lambda}_{nk}|z_c)\times(E_{nk}\boldsymbol{\nabla}_{\boldsymbol{k}}E_{nk}) + (z_c|\boldsymbol{\mathcal{S}}_{nk}|z_c). \tag{C29}$$

Combining the above relation and the commutation relation, we obtain the result

$$\langle W|\mathcal{J}|W\rangle = \int_{\mathrm{BZ}} d\boldsymbol{k}w^{2}_{r}(\boldsymbol{k}-\boldsymbol{k}_c)\Big\{[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},t)]\times(E_{nk}\boldsymbol{\nabla}_{\boldsymbol{k}}E_{nk}) + \frac{E_{nk}}{4}[\langle U^{E}_{nz_ck}|\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times\boldsymbol{S}|U^{H}_{nz_ck}\rangle - \langle U^{H}_{nz_ck}|\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle - \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times\boldsymbol{S}|U^{E}_{nz_ck}\rangle]\Big\}$$
$$= \int_{\mathrm{BZ}} d\boldsymbol{k}w^{2}_{r}(\boldsymbol{k}-\boldsymbol{k}_c)\{[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},t)-(z_c|\boldsymbol{\Lambda}_{nk}|z_c)]\times(E_{nk}\boldsymbol{\nabla}_{\boldsymbol{k}}E_{nk})+(z_c|\boldsymbol{\mathcal{S}}_{nk}|z_c)\}$$
$$\cong \boldsymbol{r}_c\times(E_{nk_c}\boldsymbol{\nabla}_{\boldsymbol{k}_c}E_{nk_c})+(z_c|\boldsymbol{\mathcal{S}}_{nk_c}|z_c). \tag{C30}$$

The proof of Eq. (C29) needs basic but tedious calculations. Here we comment that the formula is confirmed by using Eqs. (29a)–(31b) and the $\boldsymbol{k}$ derivatives of Eqs. (31a) and (31b). The outline of the derivation is given as follows:

$$(z_c|\boldsymbol{\Lambda}_{nk}|z_c)\times(E_{nk}\boldsymbol{\nabla}_{\boldsymbol{k}}E_{nk})$$
$$+\frac{E_{nk}}{4}[\langle U^{E}_{nz_ck}|\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times\boldsymbol{S}|U^{H}_{nz_ck}\rangle - \langle U^{H}_{nz_ck}|\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle - \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times\boldsymbol{S}|U^{E}_{nz_ck}\rangle]$$
$$= \frac{i}{8}[\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times(\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^{E}_{\boldsymbol{k}})|U^{E}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times(\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^{H}_{\boldsymbol{k}})|U^{H}_{nz_ck}\rangle + \langle U^{E}_{nz_ck}|(\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^{E}_{\boldsymbol{k}})\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle + \langle U^{H}_{nz_ck}|(\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^{H}_{\boldsymbol{k}})\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle]$$
$$-\frac{i}{4}[\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times(\vec{\epsilon}E^{2}_{nk}-\Xi^{E}_{\boldsymbol{k}})|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times(\vec{\mu}E^{2}_{nk}-\Xi^{H}_{\boldsymbol{k}})|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle]$$
$$-\frac{i}{4}[\langle U^{H}_{nz_ck}|\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\epsilon}^{-1}\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}|U^{E}_{nz_ck}\rangle + \langle U^{E}_{nz_ck}|\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\epsilon}^{-1}\boldsymbol{S}|U^{H}_{nz_ck}\rangle]$$
$$= -\frac{i}{4}[\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times(\vec{\epsilon}E^{2}_{nk}-\Xi^{E}_{\boldsymbol{k}})|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times(\vec{\mu}E^{2}_{nk}-\Xi^{H}_{\boldsymbol{k}})|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle]$$
$$+\frac{i}{8}[\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}|\times(\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}-\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S})|U^{E}_{nz_ck}\rangle - \langle U^{E}_{nz_ck}|(\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}-\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S})\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{E}_{nz_ck}\rangle + \langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}|\times(\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}-\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S})|U^{H}_{nz_ck}\rangle - \langle U^{H}_{nz_ck}|(\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{S}-\boldsymbol{S}\vec{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}}\cdot\boldsymbol{S})\times|\boldsymbol{\nabla}_{\boldsymbol{k}}U^{H}_{nz_ck}\rangle]$$

$$= (z_c|\boldsymbol{\mathcal{S}}_{n\boldsymbol{k}}|z_c) + \frac{i}{8}\boldsymbol{\nabla}_{\boldsymbol{k}} \times [\langle U^E_{nz_c\boldsymbol{k}}|(\boldsymbol{P}_{\boldsymbol{k}} \cdot \boldsymbol{S}\overleftrightarrow{\mu}^{-1}\boldsymbol{S} - \boldsymbol{S}\overleftrightarrow{\mu}^{-1}\boldsymbol{P}_{\boldsymbol{k}} \cdot \boldsymbol{S})|U^E_{nz_c\boldsymbol{k}}\rangle + \langle U^H_{nz_c\boldsymbol{k}}|(\boldsymbol{P}_{\boldsymbol{k}} \cdot \boldsymbol{S}\overleftrightarrow{\epsilon}^{-1}\boldsymbol{S} - \boldsymbol{S}\overleftrightarrow{\epsilon}^{-1}\boldsymbol{P}_{\boldsymbol{k}} \cdot \boldsymbol{S})|U^H_{nz_c\boldsymbol{k}}\rangle]$$
$$= (z_c|\boldsymbol{\mathcal{S}}_{n\boldsymbol{k}}|z_c). \tag{C31}$$

APPENDIX D: EXPECTATION VALUES IN A MODULATED SYSTEM

When a modulation is introduced into a periodic system, the arguments given in Appendix C are modified. Here we consider the modulation represented by Eq. (38) which is sufficiently weak and smooth. It is noted that the commutation relation Eq. (C17) is not modified even under any modulation. From this commutation relation, the creation and annihilation operators of approximated eigenmodes satisfy the same commutation relation as that in a periodic system as shown below. However, the approximated eigenmodes depend on the variable $\boldsymbol{r}_c$. Thus we must additionally take into account the operator $\boldsymbol{\nabla}_{\boldsymbol{r}_c} a^\dagger_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}$ for the derivation of the effective Lagrangian. Fortunately, we can show that the contribution from $\boldsymbol{\nabla}_{\boldsymbol{r}_c} a^\dagger_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}$ vanishes by the following relation:

$$[a_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}, a^\dagger_{n'\lambda'\boldsymbol{k}';\boldsymbol{r}_c}]$$
$$= i\int d\boldsymbol{r}[\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r}) \cdot [\boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})] - [\boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{\Phi}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})] \cdot \boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})]$$
$$= (E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}'})\int d\boldsymbol{r}\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})$$
$$= \frac{E_{n\boldsymbol{k}} + E_{n'\boldsymbol{k}}}{\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}}}}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')\langle U^E_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}}\rangle$$
$$= \delta_{nn'}\delta_{\lambda\lambda'}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}'), \tag{D1a}$$

$$[a_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}, \boldsymbol{\nabla}_{\boldsymbol{r}_c} a^\dagger_{n'\lambda'\boldsymbol{k}';\boldsymbol{r}_c}]$$
$$= -\frac{i}{2}\left[\boldsymbol{\nabla}_{\boldsymbol{r}_c} \ln \frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right]\int d\boldsymbol{r}[\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r}) \cdot [\boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})] + [\boldsymbol{\nabla}_{\boldsymbol{r}} \times \boldsymbol{\Phi}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})] \cdot \boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})]$$
$$= \frac{1}{2}\left[\boldsymbol{\nabla}_{\boldsymbol{r}_c} \ln \frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right](E_{n\boldsymbol{k}} - E_{n'\boldsymbol{k}'}) \times \int d\boldsymbol{r}\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})$$
$$= \frac{1}{4}\left[\boldsymbol{\nabla}_{\boldsymbol{r}_c} \ln \frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right]\frac{E_{n\boldsymbol{k}} - E_{n'\boldsymbol{k}}}{\sqrt{E_{n\boldsymbol{k}}E_{n'\boldsymbol{k}}}}\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')\langle U^E_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}}\rangle$$
$$= 0, \tag{D1b}$$

where, in each commutation relation, we have used Eqs. (21a) from the first to the second expression, and Eqs. (C19) and (C20) from the second to the third expression.

Next the expectation values of H and $\boldsymbol{\mathcal{R}}$ in a modulated system will be estimated. In the effective Lagrangian, we retain terms up to the first order with respect to the derivative of $\gamma_\epsilon(\boldsymbol{r})$ or $\gamma_\mu(\boldsymbol{r})$. Therefore we need to estimate the expectation value of H up to the first derivatives of the modulation functions. As for the expectation value of $\boldsymbol{\mathcal{R}}$, we may neglect derivative terms as was discussed in Sec. II C. The commutation relations needed to estimate $\langle W|H|W\rangle$ and $\langle W|\boldsymbol{\mathcal{R}}|W\rangle$ are calculated by derivative expansion with respect to the modulation functions. First, the commutation relation for $\langle W|H|W\rangle$ is estimated as follows:

$$[a_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}, [H, a^\dagger_{n'\lambda'\boldsymbol{k}';\boldsymbol{r}_c}]]$$
$$= E_{n\boldsymbol{k};\boldsymbol{r}_c}E_{n'\boldsymbol{k}';\boldsymbol{r}_c}\int d\boldsymbol{r}\left[\frac{\gamma^2_\epsilon(\boldsymbol{r})}{\gamma^2_\epsilon(\boldsymbol{r}_c)}\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r}) + \frac{\gamma^2_\mu(\boldsymbol{r})}{\gamma^2_\mu(\boldsymbol{r}_c)}\boldsymbol{\Phi}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})\right]$$
$$= \frac{1}{2}\sqrt{E_{n\boldsymbol{k};\boldsymbol{r}_c}E_{n'\boldsymbol{k};\boldsymbol{r}_c}}\{\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')[\langle U^E_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}}\rangle + \langle U^H_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\mu}|U^H_{n'\lambda'\boldsymbol{k}}\rangle] + ie^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_c} \times [(\boldsymbol{\nabla}_{\boldsymbol{k}} - \boldsymbol{\nabla}_{\boldsymbol{k}'})\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')] \cdot [[\boldsymbol{\nabla}_{\boldsymbol{r}_c} \ln \gamma_\epsilon(\boldsymbol{r}_c)] \times \langle U^E_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}'}\rangle + [\boldsymbol{\nabla}_{\boldsymbol{r}_c} \ln \gamma_\mu(\boldsymbol{r}_c)]\langle U^H_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\mu}|U^H_{n'\lambda'\boldsymbol{k}'}\rangle]\} + \cdots. \tag{D2}$$

The first and second terms in the second expression come from the terms of zeroth and first order of $(\boldsymbol{r}-\boldsymbol{r}_c)$ respectively. In the transformation to the last expression, we have used Eqs. (C19), (C20), and (C23).

In the same manner, restricting to the case in which both of $\boldsymbol{k}$ and $\boldsymbol{k}'$ are in the first Brillouin zone and using Eq. (C23), the commutation relation for $\langle W|\boldsymbol{\mathcal{R}}|W\rangle$ is estimated as follows:

$$[a_{n\lambda\boldsymbol{k};\boldsymbol{r}_c}, [\boldsymbol{\mathcal{R}}, a^\dagger_{n'\lambda'\boldsymbol{k}';\boldsymbol{r}_c}]]$$
$$= E_{n\boldsymbol{k};\boldsymbol{r}_c}E_{n'\boldsymbol{k}';\boldsymbol{r}_c}\int d\boldsymbol{r}\boldsymbol{r}\left[\frac{\gamma^2_\epsilon(\boldsymbol{r})}{\gamma^2_\epsilon(\boldsymbol{r}_c)}\boldsymbol{\Phi}^{E*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\epsilon}(\boldsymbol{r})\boldsymbol{\Phi}^E_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r}) + \frac{\gamma^2_\mu(\boldsymbol{r})}{\gamma^2_\mu(\boldsymbol{r}_c)}\boldsymbol{\Phi}^{H*}_{n\lambda\boldsymbol{k}}(\boldsymbol{r})\overleftrightarrow{\mu}(\boldsymbol{r})\boldsymbol{\Phi}^H_{n'\lambda'\boldsymbol{k}'}(\boldsymbol{r})\right]$$
$$= \frac{i}{4}\sqrt{E_{n\boldsymbol{k};\boldsymbol{r}_c}E_{n'\boldsymbol{k}';\boldsymbol{r}_c}}[(\boldsymbol{\nabla}_{\boldsymbol{k}} - \boldsymbol{\nabla}_{\boldsymbol{k}'})\tilde{\delta}(\boldsymbol{k}-\boldsymbol{k}')] \times [\langle U^E_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\epsilon}|U^E_{n'\lambda'\boldsymbol{k}'}\rangle + \langle U^H_{n\lambda\boldsymbol{k}}|\overleftrightarrow{\mu}|U^H_{n'\lambda'\boldsymbol{k}'}\rangle] + \cdots. \tag{D3}$$

From these commutation relations, we obtain the following results which leads to the estimation of the center of gravity as $\langle W|\boldsymbol{\mathcal{R}}|W\rangle/\langle W|H|W\rangle \cong \nabla_{\boldsymbol{k}}\vartheta(\boldsymbol{k}_c, \boldsymbol{r}_c, z_c, t) - (z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)$,

$$\langle W|H|W\rangle$$
$$= \int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c) E_{n\boldsymbol{k};\boldsymbol{r}_c}$$
$$\times\{1+[\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln[\gamma_\epsilon(\boldsymbol{r}_c)\gamma_\mu(\boldsymbol{r}_c)]]\cdot[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)-\boldsymbol{r}_c]$$
$$-[\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln\gamma_\epsilon(\boldsymbol{r}_c)]\cdot(z_c|\boldsymbol{\Lambda}^E_{n\boldsymbol{k}}|z_c)$$
$$+[\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln\gamma_\mu(\boldsymbol{r}_c)]\cdot(z_c|\boldsymbol{\Lambda}^H_{n\boldsymbol{k}}|z_c)\}$$
$$+\cdots$$
$$\cong (1+\{\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln[\gamma_\epsilon(\boldsymbol{r}_c)\gamma_\mu(\boldsymbol{r}_c)]\}\cdot(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)$$
$$-[\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln\gamma_\epsilon(\boldsymbol{r}_c)]\cdot(z_c|\boldsymbol{\Lambda}^E_{n\boldsymbol{k}_c}|z_c)$$
$$+[\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln\gamma_\mu(\boldsymbol{r}_c)]\cdot(z_c|\boldsymbol{\Lambda}^H_{n\boldsymbol{k}_c}|z_c))E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}$$
$$=\left\{1-\left[\boldsymbol{\nabla}_{\boldsymbol{r}_c}\ln\frac{\gamma_\epsilon(\boldsymbol{r}_c)}{\gamma_\mu(\boldsymbol{r}_c)}\right]\cdot(z_c|\boldsymbol{\Delta}_{n\boldsymbol{k}_c}|z_c)\right\}E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}, \tag{D4a}$$

$$\langle W|\boldsymbol{\mathcal{R}}|W\rangle$$
$$=\frac{i}{2}\int_{\mathrm{BZ}} d\boldsymbol{k} E_{n\boldsymbol{k};\boldsymbol{r}_c}\{w^*(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)\boldsymbol{\nabla}_{\boldsymbol{k}}w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,t)$$
$$-[\boldsymbol{\nabla}_{\boldsymbol{k}}w^*(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)]w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,t)\}$$
$$+\frac{i}{4}\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)E_{n\boldsymbol{k};\boldsymbol{r}_c}$$
$$\times[\langle U^E_{nz_c\boldsymbol{k}}|\vec{\boldsymbol{\epsilon}}|\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{nz_c\boldsymbol{k}}\rangle-\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^E_{nz_c\boldsymbol{k}}|\vec{\boldsymbol{\epsilon}}|U^E_{nz_c\boldsymbol{k}}\rangle$$
$$+\langle U^H_{nz_c\boldsymbol{k}}|\vec{\boldsymbol{\mu}}|\boldsymbol{\nabla}_{\boldsymbol{k}}U^H_{nz_c\boldsymbol{k}}\rangle-\langle\boldsymbol{\nabla}_{\boldsymbol{k}}U^H_{nz_c\boldsymbol{k}}|\vec{\boldsymbol{\mu}}|U^H_{nz_c\boldsymbol{k}}\rangle]+\cdots$$
$$=\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)E_{n\boldsymbol{k};\boldsymbol{r}_c}[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)-(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}}|z_c)]$$
$$+\cdots$$
$$\cong E_{n\boldsymbol{k}_c;\boldsymbol{r}_c}[\boldsymbol{\nabla}_{\boldsymbol{k}_c}\vartheta(\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)-(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)]. \tag{D4b}$$

The above estimation for the center of gravity suggests that the position $\boldsymbol{r}_c$ defined by Eq. (42) may be regarded as the center of gravity even in the case with a perturbative modulation. In the derivation of the effective Lagrangian, we need to estimate the inner product between the wave packet and its time derivative. Finally we present the detail for the calculation of this product, by regarding Eq. (42) as the definition of the center of wave packet,

$$\langle W|i\frac{d}{dt}|W\rangle$$
$$=i\int_{\mathrm{BZ}} d\boldsymbol{k} w^*(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)\frac{d}{dt}w(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)$$
$$+i(z_c|\dot{z}_c)+i\int_{\mathrm{BZ}} d\boldsymbol{k}d\boldsymbol{k}' w^*(\boldsymbol{k},\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)w(\boldsymbol{k}',\boldsymbol{k}_c,\boldsymbol{r}_c,z_c,t)$$
$$\times\langle 0|a_{nz_c\boldsymbol{k};\boldsymbol{r}_c}[\dot{\boldsymbol{r}}_c\cdot\boldsymbol{\nabla}_{\boldsymbol{r}_c}a^\dagger_{nz_c\boldsymbol{k}';\boldsymbol{r}_c}]|0\rangle$$
$$=\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\frac{d}{dt}\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)+i(z_c|\dot{z}_c)$$
$$=-\dot{\boldsymbol{k}}_c\cdot\int_{\mathrm{BZ}} d\boldsymbol{k}[\boldsymbol{\nabla}_{\boldsymbol{k}_c}w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)]\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)$$
$$+i(z_c|\dot{z}_c)+\frac{d}{dt}\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)$$
$$=\dot{\boldsymbol{k}}_c\cdot\int_{\mathrm{BZ}} d\boldsymbol{k}[\boldsymbol{\nabla}_{\boldsymbol{k}}w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)]\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)+i(z_c|\dot{z}_c)$$
$$+\frac{d}{dt}\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)$$
$$=-\dot{\boldsymbol{k}}_c\cdot\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)[\boldsymbol{\nabla}_{\boldsymbol{k}}\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)]$$
$$+i(z_c|\dot{z}_c)+\frac{d}{dt}\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)$$
$$=-\dot{\boldsymbol{k}}_c\cdot\left[\boldsymbol{r}_c+\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)\right]$$
$$+i(z_c|\dot{z}_c)+\frac{d}{dt}\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)$$
$$\cong\boldsymbol{k}_c\cdot\dot{\boldsymbol{r}}_c-\dot{\boldsymbol{k}}_c\cdot(z_c|\boldsymbol{\Lambda}_{n\boldsymbol{k}_c}|z_c)+i(z_c|\dot{z}_c)$$
$$+\frac{d}{dt}\left[\int_{\mathrm{BZ}} d\boldsymbol{k} w_r^2(\boldsymbol{k}-\boldsymbol{k}_c)\vartheta(\boldsymbol{k},\boldsymbol{r}_c,z_c,t)-\boldsymbol{k}_c\cdot\boldsymbol{r}_c\right], \tag{D5}$$

where Eq. (D1b) and $a_{nz_c\boldsymbol{k};\boldsymbol{r}_c}|0\rangle=0$ are used in the transformation from the first expression to the second expression.

APPENDIX E: BERRY CURVATURE AND INTERNAL ROTATION

In a system with generic periodic structure, it is tough work to analytically calculate the Berry curvature and the internal rotation. However, it is easy to obtain them numerically by rewriting inner products of Bloch functions and their $\boldsymbol{k}$ derivatives to conventional matrix elements. Here we present some formulas which are convenient for numerical calculations.

For later convenience, we separate the Berry curvature as

$$\boldsymbol{\Omega}_{n\boldsymbol{k}}=\frac{1}{2}[\boldsymbol{\Omega}^E_{n\boldsymbol{k}}+\boldsymbol{\Omega}^H_{n\boldsymbol{k}}]-i\boldsymbol{\Delta}_{n\boldsymbol{k}}\times\boldsymbol{\Delta}_{n\boldsymbol{k}}, \tag{E1a}$$

$$\boldsymbol{\Omega}^F_{n\boldsymbol{k}}=\boldsymbol{\nabla}_{\boldsymbol{k}}\times\boldsymbol{\Lambda}^F_{n\boldsymbol{k}}+i\boldsymbol{\Lambda}^F_{n\boldsymbol{k}}\times\boldsymbol{\Lambda}^F_{n\boldsymbol{k}}, \tag{E1b}$$

where $F=E$ or H. In the following, we rewrite the $\boldsymbol{k}$ derivative in the above expression in terms of the Feynman-Hellman relation. However, even if $|U^{E,H}_{n\lambda\boldsymbol{k}}\rangle$ is a Bloch function of a physical state, its derivative may have an unphysical component proportional to $|K\rangle$. In other words, Bloch functions and their derivatives should be expanded by nonorthogonal bases as follows:

$$|V_{\boldsymbol{k}}\rangle = \sum_{n,\lambda} |U^E_{n\lambda\boldsymbol{k}}\rangle\langle U^E_{n\lambda\boldsymbol{k}}|\vec{\vec{\epsilon}}|V_{\boldsymbol{k}}\rangle + \sum_{\boldsymbol{G},\boldsymbol{G}'} |K\rangle[\Gamma^E_{\boldsymbol{k}}]^{-1}(\boldsymbol{G},\boldsymbol{G}')\langle K'|\vec{\vec{\epsilon}}|V_{\boldsymbol{k}}\rangle \tag{E2a}$$

$$= \sum_{n,\lambda} |U^H_{n\lambda\boldsymbol{k}}\rangle\langle U^H_{n\lambda\boldsymbol{k}}|\vec{\vec{\mu}}|V_{\boldsymbol{k}}\rangle + \sum_{\boldsymbol{G},\boldsymbol{G}'} |K\rangle[\Gamma^H_{\boldsymbol{k}}]^{-1}(\boldsymbol{G},\boldsymbol{G}')\langle K'|\vec{\vec{\mu}}|V_{\boldsymbol{k}}\rangle, \tag{E2b}$$

where $\Gamma^E_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}')=\boldsymbol{K}\vec{\vec{\epsilon}}(\boldsymbol{G},\boldsymbol{G}')\boldsymbol{K}'$ and $\Gamma^H_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}')=\boldsymbol{K}\vec{\vec{\mu}}(\boldsymbol{G},\boldsymbol{G}')\boldsymbol{K}'$. By using the above expansion and the Feynman-Hellman relation derived from Eqs. (31a) and (31b), we can rewrite the Berry curvature as

$$[\boldsymbol{\Omega}^E_{n\boldsymbol{k}}]_{\lambda\lambda'} = -i \sum_{m\neq n,\lambda''} \frac{\langle U^E_{n\lambda\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^E_{\boldsymbol{k}}]|U^E_{m\lambda''\boldsymbol{k}}\rangle \times \langle U^E_{m\lambda''\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^E_{\boldsymbol{k}}]|U^E_{n\lambda'\boldsymbol{k}}\rangle}{(E^2_{n\boldsymbol{k}} - E^2_{m\boldsymbol{k}})^2} + \langle U^E_{n\lambda\boldsymbol{k}}|\vec{\vec{\epsilon}}[\Gamma^E_{\boldsymbol{k}}]^{-1}\boldsymbol{S}\vec{\vec{\epsilon}}|U^E_{n\lambda'\boldsymbol{k}}\rangle, \tag{E3a}$$

$$[\boldsymbol{\Omega}^H_{n\boldsymbol{k}}]_{\lambda\lambda'} = -i \sum_{m\neq n,\lambda''} \frac{\langle U^H_{n\lambda\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^H_{\boldsymbol{k}}]|U^H_{m\lambda''\boldsymbol{k}}\rangle \times \langle U^H_{m\lambda''\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^H_{\boldsymbol{k}}]|U^H_{n\lambda'\boldsymbol{k}}\rangle}{(E^2_{n\boldsymbol{k}} - E^2_{m\boldsymbol{k}})^2} + \langle U^H_{n\lambda\boldsymbol{k}}|\vec{\vec{\mu}}[\Gamma^H_{\boldsymbol{k}}]^{-1}\boldsymbol{S}\vec{\vec{\mu}}|U^H_{n\lambda'\boldsymbol{k}}\rangle, \tag{E3b}$$

$$[\boldsymbol{\Delta}_{n\boldsymbol{k}}]_{\lambda\lambda'} = \frac{1}{4E_{n\boldsymbol{k}}}[\langle U^E_{n\lambda\boldsymbol{k}}|\boldsymbol{S}|U^H_{n\lambda'\boldsymbol{k}}\rangle + \langle U^H_{n\lambda\boldsymbol{k}}|\boldsymbol{S}|U^E_{n\lambda'\boldsymbol{k}}\rangle]. \tag{E3c}$$

Thus $\boldsymbol{\Omega}^F_{n\boldsymbol{k}}$ $(F=E,H)$ is enhanced when the band comes close to other bands in energy, with the enhancement being inversely proportional to the square of energy difference. In contrast, $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ does not have such an enhancement. Though Eq. (E3c) seems to diverge at $E_{n\boldsymbol{k}}\to 0$ $(\boldsymbol{k}\to 0)$, it is not the case, as shown in Appendix 12 for a specific case. In the long-wavelength limit $\boldsymbol{k}\to 0$, a propagating light becomes insensitive to a spatial modulation of $\epsilon(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$, and a medium is regarded as uniform. Because $\boldsymbol{\Delta}_{n\boldsymbol{k}}=0$ for a uniform isotropic medium, $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ in a generic periodic medium should behave as $\boldsymbol{\Delta}_{n\boldsymbol{k}}\to 0$ in the long-wavelength limit.

In the same manner, the internal rotation is also rewritten as follows:

$$[\boldsymbol{\mathcal{S}}^E_{n\boldsymbol{k}}]_{\lambda\lambda'} = \frac{1}{2}\Bigg[-i \sum_{m\neq n,\lambda''} \frac{\langle U^E_{n\lambda\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^E_{\boldsymbol{k}}]|U^E_{m\lambda''\boldsymbol{k}}\rangle \times \langle U^E_{m\lambda''\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^E_{\boldsymbol{k}}]|U^E_{n\lambda'\boldsymbol{k}}\rangle}{E^2_{n\boldsymbol{k}} - E^2_{m\boldsymbol{k}}}$$
$$+ E^2_{n\boldsymbol{k}}\langle U^E_{n\lambda\boldsymbol{k}}|\vec{\vec{\epsilon}}[\Gamma^E_{\boldsymbol{k}}]^{-1}\boldsymbol{S}\vec{\vec{\epsilon}}|U^E_{n\lambda'\boldsymbol{k}}\rangle - i\langle U^E_{n\lambda\boldsymbol{k}}|\boldsymbol{S}\times\vec{\vec{\mu}}^{-1}\boldsymbol{S}|U^E_{n\lambda'\boldsymbol{k}}\rangle\Bigg], \tag{E4a}$$

$$[\boldsymbol{\mathcal{S}}^H_{n\boldsymbol{k}}]_{\lambda\lambda'} = \frac{1}{2}\Bigg[-i \sum_{m\neq n,\lambda''} \frac{\langle U^H_{n\lambda\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^H_{\boldsymbol{k}}]|U^H_{m\lambda''\boldsymbol{k}}\rangle \times \langle U^H_{m\lambda''\boldsymbol{k}}|[\boldsymbol{\nabla}_{\boldsymbol{k}}\Xi^H_{\boldsymbol{k}}]|U^H_{n\lambda'\boldsymbol{k}}\rangle}{E^2_{n\boldsymbol{k}} - E^2_{m\boldsymbol{k}}}$$
$$+ E^2_{n\boldsymbol{k}}\langle U^H_{n\lambda\boldsymbol{k}}|\vec{\vec{\mu}}[\Gamma^H_{\boldsymbol{k}}]^{-1}\boldsymbol{S}\vec{\vec{\mu}}|U^H_{n\lambda'\boldsymbol{k}}\rangle - i\langle U^H_{n\lambda\boldsymbol{k}}|\boldsymbol{S}\times\vec{\vec{\epsilon}}^{-1}\boldsymbol{S}|U^H_{n\lambda'\boldsymbol{k}}\rangle\Bigg]. \tag{E4b}$$

It should be noted that $\boldsymbol{\Omega}^E_{n\boldsymbol{k}}$ and $\boldsymbol{\Omega}^H_{n\boldsymbol{k}}$ have very similar expressions to $\boldsymbol{\mathcal{S}}^E_{n\boldsymbol{k}}$ and $\boldsymbol{\mathcal{S}}^H_{n\boldsymbol{k}}$, respectively. This suggests that there is always some kind of rotation when the Berry curvatures are nonzero. In this sense, we have generalized the argument for the quantum Hall system in Ref. [36] to a photonic system. In the quantum Hall system, the internal rotation is the internal orbital rotation originated by the cyclotron motion under an external magnetic field. On the other hand, in the present case, the internal rotation is the combination of the polarization and the internal orbital rotation originated from a periodic structure. When a system is isotopic and homogeneous, Eqs. (E3a) and (E4b) are reduced to the Berry curvature, $\frac{\boldsymbol{k}}{k^3}\sigma_3$, and the spin divided by $\epsilon\mu$, $\frac{1}{\epsilon\mu}\cdot\frac{\boldsymbol{k}}{k}\sigma_3$. These contributions come only from the terms including the spin operator $\boldsymbol{S}$, and $\boldsymbol{\Delta}_{n\boldsymbol{k}}=0$, i.e., nonzero $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ is originated by the anisotropy or the periodic structure of $\vec{\vec{\epsilon}}$ and $\vec{\vec{\mu}}$. Even in generic cases, $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ has the dimension of length, and its magnitude is a lattice constant at most.

APPENDIX F: TRANSVERSE SHIFT IN CLASSICAL ELECTRODYNAMICS

Here we prove the consistency between our result for the transverse shift [Eq. (63)], which is consistent with the TAM conservation for individual photons [Eq. (65)], and the result by Fedoseev [26,27], which is based on classical electrody-

namics. In Refs. [26,27], each wave packet is constructed as a superposition of plane waves with wave vectors $\boldsymbol{k}=\boldsymbol{k}_c+\boldsymbol{\kappa}$, where $\boldsymbol{\kappa}$ are distributed around zero vector. (In the notation of Refs. [26,27], $\boldsymbol{k}_c$ is represented by $\boldsymbol{K}$.) The polarization vector of each constituent plane wave is defined by Eq. (23) in Ref. [26] with Eq. (7) in Ref. [27],

$$\boldsymbol{e}^{(j)}(\boldsymbol{\kappa}) = z_s^{(j)}(\boldsymbol{\kappa})\boldsymbol{s}^{(j)}(\boldsymbol{\kappa}) + z_p^{(j)}(\boldsymbol{\kappa})\boldsymbol{p}^{(j)}(\boldsymbol{\kappa}), \tag{F1}$$

where $j=i,\rho,\tau$ for incident, reflected, and transmitted beams, respectively, $z_s^{(j)}(\boldsymbol{\kappa})$ and $z_p^{(j)}(\boldsymbol{\kappa})$ $[|z_s^{(j)}(\boldsymbol{\kappa})|^2+|z_p^{(j)}(\boldsymbol{\kappa})|^2=1]$ represent the polarization state of each plane wave, $\boldsymbol{s}^{(j)}(\boldsymbol{\kappa})$ and $\boldsymbol{p}^{(j)}(\boldsymbol{\kappa})$ are the s- and p-polarization vectors defined by

$$\boldsymbol{s}^{(j)}(\boldsymbol{\kappa}) = \frac{\boldsymbol{n}\times\boldsymbol{k}}{|\boldsymbol{n}\times\boldsymbol{k}|}, \quad \boldsymbol{p}^{(j)}(\boldsymbol{\kappa}) = \boldsymbol{s}(\boldsymbol{\kappa})\times\frac{\boldsymbol{k}}{|\boldsymbol{k}|}, \tag{F2}$$

where $\boldsymbol{n}=(0,0,1)$ is normal to the interface, and we consider the same configuration of the interface and beams as those in Sec. III A. The relation between the present notation and that in Ref. [27] is represented as $z_s^{(j)}(\boldsymbol{\kappa})\leftrightarrow A^{(j)}(\boldsymbol{\kappa})/\sqrt{|A^{(j)}(\boldsymbol{\kappa})|^2+|B^{(j)}(\boldsymbol{\kappa})|^2}$ and $z_p^{(j)}(\boldsymbol{\kappa})\leftrightarrow B^{(j)}(\boldsymbol{\kappa})/\sqrt{|A^{(j)}(\boldsymbol{\kappa})|^2+|B^{(j)}(\boldsymbol{\kappa})|^2}$, $\boldsymbol{n}\leftrightarrow\boldsymbol{N}$, and $\boldsymbol{k}/|\boldsymbol{k}|\leftrightarrow\boldsymbol{m}(\boldsymbol{\kappa})$.

By the Maxwell equations, $z_s^{(\rho,\tau)}(\boldsymbol{\kappa})$ and $z_p^{(\rho,\tau)}(\boldsymbol{\kappa})$ are exactly given by

$$z_s^{(j)}(\boldsymbol{\kappa}) = \frac{t_s^{(j)}(\boldsymbol{\kappa})z_s^{(i)}(\boldsymbol{\kappa})}{\sqrt{|t_s^{(j)}(\boldsymbol{\kappa})z_s^{(i)}(\boldsymbol{\kappa})|^2+|t_p^{(j)}(\boldsymbol{\kappa})z_p^{(i)}(\boldsymbol{\kappa})|^2}}, \tag{F3a}$$

$$z_p^{(j)}(\boldsymbol{\kappa}) = \frac{t_p^{(j)}(\boldsymbol{\kappa})z_p^{(i)}(\boldsymbol{\kappa})}{\sqrt{|t_s^{(j)}(\boldsymbol{\kappa})z_s^{(i)}(\boldsymbol{\kappa})|^2+|t_p^{(j)}(\boldsymbol{\kappa})z_p^{(i)}(\boldsymbol{\kappa})|^2}}, \tag{F3b}$$

where $j=\rho$ or τ, $t_s^{(\rho)}(\boldsymbol{\kappa})$ and $t_p^{(\rho)}(\boldsymbol{\kappa})$ are the amplitude reflection coefficients for the s- and p-polarized plane waves, $t_s^{(\tau)}(\boldsymbol{\kappa})$ and $t_p^{(\tau)}(\boldsymbol{\kappa})$ are the amplitude transmission coefficients for the s- and p-polarized plane waves, i.e., $t_s^{(\rho)}\leftrightarrow R_s$, $t_p^{(\rho)}\leftrightarrow R_p$, $t_s^{(\tau)}\leftrightarrow T_s$, and $t_p^{(\tau)}\leftrightarrow T_p$ in our notation in Sec. III A.

In our constitution method for an incident wave packet, the polarization state of each constituent plane wave, i.e., the set of $z_s^{(i)}$ and $z_p^{(i)}$, is independent of $\boldsymbol{\kappa}$. Otherwise, the concept of "an elliptically polarized incident wave packet" gets fuzzy (see Sec. III A 3). Thus this is a natural definition for an elliptically polarized incident wave packet. Its polarization vector is represented also in the following form,

$$\boldsymbol{e}^{(i)}(\boldsymbol{\kappa}) = \frac{\boldsymbol{p}^{(i)}(\boldsymbol{\kappa}) + m\boldsymbol{s}^{(i)}(\boldsymbol{\kappa})}{1+|m|^2}, \tag{F4}$$

where m is a complex constant, representing the polarization state. This m is identical with m defined by Bliokh *et al.* [50], and related with our $|z^I)$ in Sec. III A by

$$|z^I) = \frac{1}{\sqrt{2(1+|m|^2)}}\begin{pmatrix}1-im\\1+im\end{pmatrix}. \tag{F5}$$

It yields

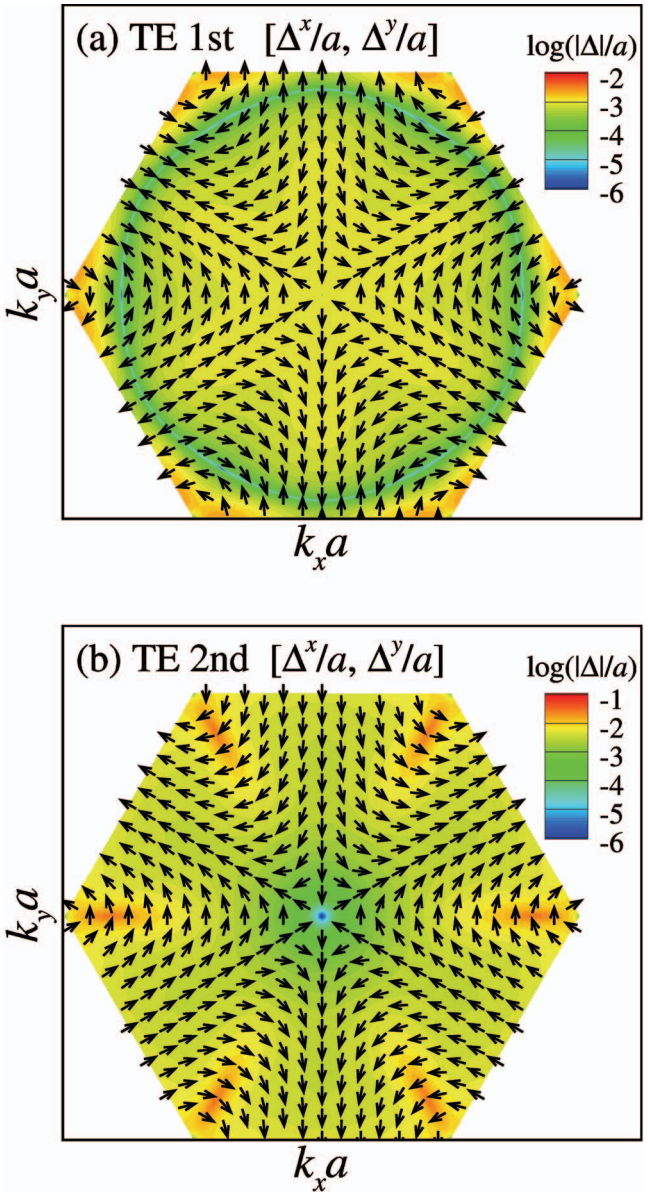

FIG. 6. (Color) Difference between the electric and magnetic parts of the Berry connection for each of (a) the TE first band and (b) the TE second band, i.e., $\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}}=(\boldsymbol{\Lambda}^E_{\mathrm{TE}n\boldsymbol{k}}-\boldsymbol{\Lambda}^H_{\mathrm{TE}n\boldsymbol{k}})/2$. The base of a logarithm is 10.

$$(z^I|\boldsymbol{\sigma}|z^I) = \frac{1}{1+|m|^2}[1-|m|^2, 2\,\mathrm{Re}(m), 2\,\mathrm{Im}(m)] \tag{F6}$$

which is used for comparison between the results here and those based on our theory of the TAM conservation for individual photons.

We now calculate the transverse shift from Eqs. (15)–(17) in Ref. [27]. The result is a sum of two terms

$$\delta y^{(j)} = h^{(j1)} + h^{(j2)}, \tag{F7}$$

where $j=\rho$ or τ, and $\delta y^{(\rho)}\leftrightarrow\delta y^R$ and $\delta y^{(\tau)}\leftrightarrow\delta y^T$ in our notation in Sec. III A. From Eqs. (13a), (13b), (17), and (18) in Ref. [27], the second term of right-hand side, $h^{(j2)}$, is proportional to the κ_y derivative of $\mathrm{Im}[\ln z_s^{(j)}(\boldsymbol{\kappa})-\ln z_p^{(j)}(\boldsymbol{\kappa})]$

at $\boldsymbol{\kappa}=0$. The amplitude reflection and refraction coefficients depend only on the polar angle, and thus their derivatives by κ_y at $\boldsymbol{\kappa}=0$ are zero, because the y component of $\boldsymbol{k}_c$ is zero in the present configuration. As was mentioned previously, $z_s^{(i)}$ and $z_p^{(i)}$ are independent of $\boldsymbol{\kappa}$. Therefore from Eqs. (F3a) and (F3b), $h^{(j2)}=0$ $(j=\rho,\tau)$, and we have

$$\delta y^{(j)} = h^{(j1)} = -i\frac{\boldsymbol{n}\cdot[\boldsymbol{e}^{(j)}(0)\times\boldsymbol{e}^{(j)*}(0)]}{|\boldsymbol{n}\times\boldsymbol{k}^{(j)}|} + i\frac{\boldsymbol{n}\cdot[\boldsymbol{e}^{(i)}(0)\times\boldsymbol{e}^{(i)*}(0)]}{|\boldsymbol{n}\times\boldsymbol{k}^{(i)}|}, \tag{F8}$$

where $\boldsymbol{k}^{(j)}$ $(j=\rho$ or $\tau)$ are mean wave vectors for reflected (ρ) and transmitted (τ) wave packets. Note that the correspondence between these wave vectors and those in Sec. III A are $\boldsymbol{k}^{(i)}\leftrightarrow\boldsymbol{k}^I$, $\boldsymbol{k}^{(\rho)}\leftrightarrow\boldsymbol{k}^R$, and $\boldsymbol{k}^{(\tau)}\leftrightarrow\boldsymbol{k}^T$. Equation (F8) is identical with Eq. (63), showing an equivalence between Fedoseev's theory based on classical electrodynamics and ours.

Finally, we rewrite Eq. (F8) in terms of our notation in Sec. III A. For partial reflection, A_p and A_s are real, and we get

$$\delta y^A = \frac{2\,\mathrm{Im}(m)}{k^I\sin\theta_I}\left[\frac{(A_s/A_p)\cos\theta_A}{1+(A_s/A_p)^2|m|^2} - \frac{\cos\theta_I}{1+|m|^2}\right], \tag{F9}$$

where $A=T$ or R. By rewriting Eq. (F9) in terms of $|z^I)$, the shift is equal to our result in Eq. (68) but not to Eq. (5) in Ref. [50]. For total reflection, R_p and R_s are complex numbers with $|R_p|=|R_s|=1$, and we get

$$\delta y^A = \frac{-2\cos\theta_I}{k^I\sin\theta_I}\frac{\mathrm{Im}(m)[\mathrm{Re}(R_p^*R_s)+1]+\mathrm{Re}(m)\mathrm{Im}(R_p^*R_s)}{1+|m|^2}. \tag{F10}$$

This is exactly the same as ours in Eq. (69). To summarize, for every case, the calculation based on classical electrodynamics gives the identical result with ours based on our quantum-mechanical formalism, and this result is consistent with the TAM conservation for individual photons.

APPENDIX G: BERRY CURVATURE IN A TWO-DIMENSIONAL PHOTONIC CRYSTAL

In order to discuss the TM and TE modes, it is convenient to introduce the following unit vectors:

$$\boldsymbol{e}_K = \frac{\boldsymbol{K}}{K}, \quad \boldsymbol{e}_1 = \frac{\boldsymbol{e}_z\times\boldsymbol{e}_K}{|\boldsymbol{e}_z\times\boldsymbol{e}_K|}, \tag{G1}$$

and the Bloch functions are represented by

$$\epsilon|U^E_{\mathrm{TE}m\boldsymbol{k}}\rangle = \boldsymbol{e}_1\otimes|U^D_{\mathrm{TE}m\boldsymbol{k}}\rangle, \tag{G2a}$$

$$\mu|U^H_{\mathrm{TE}m\boldsymbol{k}}\rangle = \boldsymbol{e}_z\otimes|U^B_{\mathrm{TE}m\boldsymbol{k}}\rangle, \tag{G2b}$$

for the TE modes and

$$\epsilon|U^E_{\mathrm{TM}m\boldsymbol{k}}\rangle = \boldsymbol{e}_z\otimes|U^D_{\mathrm{TM}m\boldsymbol{k}}\rangle, \tag{G3a}$$

$$\mu|U^H_{\mathrm{TM}m\boldsymbol{k}}\rangle = \boldsymbol{e}_1\otimes|U^B_{\mathrm{TM}m\boldsymbol{k}}\rangle, \tag{G3b}$$

for the TM modes. The superscripts, D and B, mean that they correspond to the electric and magnetic flux densities, respectively, satisfying the transversality condition, i.e., being perpendicular to $\boldsymbol{K}$.

The matrices for the eigen equations given in Eqs. (26c) and (26d) are simplified in the case with scalar $\epsilon(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$ as follows:

$$\Xi^E_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}') = \Theta_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}')\mu^{-1}(\boldsymbol{G},\boldsymbol{G}'), \tag{G4a}$$

$$\Xi^H_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}') = \Theta_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}')\epsilon^{-1}(\boldsymbol{G},\boldsymbol{G}'), \tag{G4b}$$

where $\Theta_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}')=(\boldsymbol{K}\cdot\boldsymbol{K}'I-\boldsymbol{K}'\otimes\boldsymbol{K})$, and their derivatives are represented by

$$\nabla_{k_x}\Theta_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}') = \begin{pmatrix} 0 & -K_y & -k_z \\ -K'_y & K_x+K'_x & 0 \\ -k_z & 0 & K_x+K'_x \end{pmatrix}, \tag{G5a}$$

$$\nabla_{k_y}\Theta_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}') = \begin{pmatrix} K_y+K'_y & -K'_x & 0 \\ -K_x & 0 & -k_z \\ 0 & -k_z & K_y+K'_y \end{pmatrix}, \tag{G5b}$$

$$\nabla_{k_z}\Theta_{\boldsymbol{k}}(\boldsymbol{G},\boldsymbol{G}') = \begin{pmatrix} 2k_z & 0 & -K'_x \\ 0 & 2k_z & -K'_y \\ -K_x & -K_y & 0 \end{pmatrix}. \tag{G5c}$$

because $\boldsymbol{G}$ and $\boldsymbol{G}'$ have no z components.

From the above formula together with Eqs. (E3a) and (E3b) and by setting $k_z=0$, we can easily show that the Berry curvature of a nondegenerate TM (TE) mode has only a z component,

$$\Omega^{E,z}_{\mathrm{TM}n\boldsymbol{k}} = 2\sum_{m\neq n}\frac{\mathrm{Im}[\langle U^E_{\mathrm{TM}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TM}m\boldsymbol{k}}\rangle\langle U^E_{\mathrm{TM}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TM}n\boldsymbol{k}}\rangle]}{(E^2_{\mathrm{TM}n\boldsymbol{k}}-E^2_{\mathrm{TM}m\boldsymbol{k}})^2}, \tag{G6a}$$

$$\Omega^{H,z}_{\mathrm{TM}n\boldsymbol{k}} = 2\sum_{m\neq n}\frac{\mathrm{Im}[\langle U^H_{\mathrm{TM}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TM}m\boldsymbol{k}}\rangle\langle U^H_{\mathrm{TM}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TM}n\boldsymbol{k}}\rangle]}{(E^2_{\mathrm{TM}n\boldsymbol{k}}-E^2_{\mathrm{TM}m\boldsymbol{k}})^2} + \langle U^H_{\mathrm{TM}n\boldsymbol{k}}|\mu[\Gamma^H_{\boldsymbol{k}}]^{-1}S^z\mu|U^H_{\mathrm{TM}n\boldsymbol{k}}\rangle, \tag{G6b}$$

$$\Omega^{E,z}_{\mathrm{TE}n\boldsymbol{k}} = 2\sum_{m\neq n}\frac{\mathrm{Im}[\langle U^E_{\mathrm{TE}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TE}m\boldsymbol{k}}\rangle\langle U^E_{\mathrm{TE}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TE}n\boldsymbol{k}}\rangle]}{(E^2_{\mathrm{TE}n\boldsymbol{k}}-E^2_{\mathrm{TE}m\boldsymbol{k}})^2} + \langle U^E_{\mathrm{TE}n\boldsymbol{k}}|\epsilon[\Gamma^E_{\boldsymbol{k}}]^{-1}S^z\epsilon|U^E_{\mathrm{TE}n\boldsymbol{k}}\rangle, \tag{G6c}$$

$$\Omega^{H,z}_{\mathrm{TE}n\boldsymbol{k}} = 2\sum_{m\neq n}\frac{\mathrm{Im}[\langle U^H_{\mathrm{TE}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TE}m\boldsymbol{k}}\rangle\langle U^H_{\mathrm{TE}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TE}n\boldsymbol{k}}\rangle]}{(E^2_{\mathrm{TE}n\boldsymbol{k}}-E^2_{\mathrm{TE}m\boldsymbol{k}})^2}. \tag{G6d}$$

Note that the above contributions for the Berry curvature of the TM (TE) nth band do not necessarily decrease when the energy $E_{\mathrm{TM(TE)}n\boldsymbol{k}}$ increases in contrast to the case without a periodic structure, because nearly degenerate points due to a band structure enhance the magnitude of the Berry curvature.

Following the same argument as that about the Berry curvature, the internal rotation also has only the z component for nondegenerate bands,

$$\mathcal{S}^{E,z}_{\mathrm{TM}n\boldsymbol{k}} = \sum_{m\neq n}\frac{\mathrm{Im}[\langle U^E_{\mathrm{TM}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TM}m\boldsymbol{k}}\rangle\langle U^E_{\mathrm{TM}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TM}n\boldsymbol{k}}\rangle]}{E^2_{\mathrm{TM}n\boldsymbol{k}}-E^2_{\mathrm{TM}m\boldsymbol{k}}}, \tag{G7a}$$

$$\mathcal{S}^{H,z}_{\mathrm{TM}n\boldsymbol{k}} = \sum_{m\neq n}\frac{\mathrm{Im}[\langle U^H_{\mathrm{TM}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TM}m\boldsymbol{k}}\rangle\langle U^H_{\mathrm{TM}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TM}n\boldsymbol{k}}\rangle]}{E^2_{\mathrm{TM}n\boldsymbol{k}}-E^2_{\mathrm{TM}m\boldsymbol{k}}} + \frac{1}{2}[E^2_{\mathrm{TM}n\boldsymbol{k}}\langle U^H_{\mathrm{TM}n\boldsymbol{k}}|\mu[\Gamma^H_{\boldsymbol{k}}]^{-1}S^z\mu|U^H_{\mathrm{TM}n\boldsymbol{k}}\rangle + \langle U^H_{\mathrm{TM}n\boldsymbol{k}}|\epsilon^{-1}S^z|U^H_{\mathrm{TM}n\boldsymbol{k}}\rangle], \tag{G7b}$$

$$\mathcal{S}^{E,z}_{\mathrm{TE}n\boldsymbol{k}} = \sum_{m\neq n}\frac{\mathrm{Im}[\langle U^E_{\mathrm{TE}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TE}m\boldsymbol{k}}\rangle\langle U^E_{\mathrm{TE}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^E_{\boldsymbol{k}}]|U^E_{\mathrm{TE}n\boldsymbol{k}}\rangle]}{E^2_{\mathrm{TE}n\boldsymbol{k}}-E^2_{\mathrm{TE}m\boldsymbol{k}}} + \frac{1}{2}[E^2_{\mathrm{TE}n\boldsymbol{k}}\langle U^E_{\mathrm{TE}n\lambda\boldsymbol{k}}|\epsilon[\Gamma^E_{\boldsymbol{k}}]^{-1}S^z\epsilon|U^E_{\mathrm{TE}n\lambda'\boldsymbol{k}}\rangle + \langle U^E_{\mathrm{TE}n\boldsymbol{k}}|\mu^{-1}S^z|U^E_{\mathrm{TE}n\boldsymbol{k}}\rangle], \tag{G7c}$$

$$\mathcal{S}^{H,z}_{\mathrm{TE}n\boldsymbol{k}} = \sum_{m\neq n}\frac{\mathrm{Im}[\langle U^H_{\mathrm{TE}n\boldsymbol{k}}|[\nabla_{k_x}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TE}m\boldsymbol{k}}\rangle\langle U^H_{\mathrm{TE}m\boldsymbol{k}}|[\nabla_{k_y}\Xi^H_{\boldsymbol{k}}]|U^H_{\mathrm{TE}n\boldsymbol{k}}\rangle]}{E^2_{\mathrm{TE}n\boldsymbol{k}}-E^2_{\mathrm{TE}m\boldsymbol{k}}}. \tag{G7d}$$

It is interesting that the internal rotation of a photon can be perpendicular to its propagating direction in a two-dimensional photonic crystal.

We also calculate $\boldsymbol{\Delta}_{n\boldsymbol{k}}$ from Eq. (E3c). For nondegenerate bands, there is no contribution to the Berry curvature $\boldsymbol{\Omega}_{n\boldsymbol{k}}$ from the vector product of $\boldsymbol{\Delta}_{\mathrm{TM(TE)}n\boldsymbol{k}}$ in Eq. (E1a), because $\boldsymbol{\Delta}_{\mathrm{TM(TE)}n\boldsymbol{k}}$ is a simple vector variable, not a set of matrices. Meanwhile, $\boldsymbol{\Delta}_{\mathrm{TM(TE)}n\boldsymbol{k}}$ may modify the energy spectrum when a modulation is applied. $\boldsymbol{\Delta}_{\mathrm{TM(TE)}n\boldsymbol{k}}$ is given as follows:

$$\boldsymbol{\Delta}_{\mathrm{TM}n\boldsymbol{k}} = \frac{1}{2E^2_{\mathrm{TM}n\boldsymbol{k}}}\mathrm{Im}[\langle U^D_{\mathrm{TM}n\boldsymbol{k}}|\epsilon^{-1}\mu^{-1}\boldsymbol{P}_{\boldsymbol{k}}\epsilon^{-1}|U^D_{\mathrm{TM}n\boldsymbol{k}}\rangle], \tag{G8a}$$

$$\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}} = \frac{1}{2E^2_{\mathrm{TE}n\boldsymbol{k}}}\mathrm{Im}[\langle U^B_{\mathrm{TE}n\boldsymbol{k}}|\mu^{-1}\boldsymbol{P}_{\boldsymbol{k}}\epsilon^{-1}\mu^{-1}|U^B_{\mathrm{TE}n\boldsymbol{k}}\rangle]. \tag{G8b}$$

In many cases, we can approximately regard the magnetic permeability μ to be constant. Then $\boldsymbol{\Delta}_{\mathrm{TM}n\boldsymbol{k}}$ vanishes from Eq. (G8b), whereas $\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}}$ does not in general.

APPENDIX H: REMARKS ON $\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}}$

Here we evaluate $\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}}$ for the two-dimensional photonic crystal discussed in Sec. III C and see its effect on the energy dispersion and group velocity for each of the first and second bands of TE mode. From Eq. (48), an additional correction appears in the energy of each TE mode as

$$\frac{\mathcal{E}_{\mathrm{TE}n\boldsymbol{k}_c;\boldsymbol{r}_c}}{E_{\mathrm{TE}n\boldsymbol{k}_c;\boldsymbol{r}_c}} = 1 - [\nabla_{\boldsymbol{r}_c}\ln\gamma_\epsilon(\boldsymbol{r}_c)]\cdot\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}_c}, \tag{H1}$$

where $E_{\mathrm{TE}n\boldsymbol{k}_c;\boldsymbol{r}_c} = \gamma_\epsilon(\boldsymbol{r}_c)E_{\mathrm{TE}n\boldsymbol{k}_c}$. Figure 6 shows $\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}}$ for the first and second bands of TE mode and we can see $\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}} \lesssim 0.1a$. Therefore the correction is at most a few percent as long as the modulation is sufficiently weak, i.e., $|a\nabla_{\boldsymbol{r}_c}\ln\gamma_\epsilon(\boldsymbol{r}_c)| \ll 1$. In order to make the argument complete, we also calculate a correction to the group velocity of a TE mode,

$$\nabla_{\boldsymbol{k}_c}\mathcal{E}_{\mathrm{TE}n\boldsymbol{k}_c;\boldsymbol{r}_c} = \nabla_{\boldsymbol{k}_c}E_{\mathrm{TE}n\boldsymbol{k}_c;\boldsymbol{r}_c} - \nabla_{\boldsymbol{k}_c}[[\nabla_{\boldsymbol{r}_c}\gamma_\epsilon(\boldsymbol{r}_c)]\cdot\boldsymbol{\Delta}_{\mathrm{TE}n\boldsymbol{k}_c}E_{\mathrm{TE}n\boldsymbol{k}_c}] \cong \nabla_{\boldsymbol{k}_c}E_{\mathrm{TE}n\boldsymbol{k}_c;\boldsymbol{r}_c} + \overleftrightarrow{\Pi}_{\mathrm{TE}n\boldsymbol{k}_c}\dot{\boldsymbol{k}}_c, \tag{H2a}$$

$$\overleftrightarrow{\Pi}^{ij}_{\mathrm{TE}n\boldsymbol{k}} = \nabla^i_{\boldsymbol{k}}\boldsymbol{\Delta}^j_{\mathrm{TE}n\boldsymbol{k}} + [\nabla^i_{\boldsymbol{k}}\ln E_{\mathrm{TE}n\boldsymbol{k}}]\boldsymbol{\Delta}^j_{\mathrm{TE}n\boldsymbol{k}}. \tag{H2b}$$

Here we used the relation $\dot{\boldsymbol{k}}_c \cong -[\nabla_{\boldsymbol{r}_c}\gamma_\epsilon(\boldsymbol{r}_c)]E_{\mathrm{TE}n\boldsymbol{k}_c}$ for smooth and weak modulation. By plugging Eq. (H2a) to the equation of motion for $\boldsymbol{r}_c$ in Eq. (53a), $\overleftrightarrow{\Pi}_{\mathrm{TE}n\boldsymbol{k}}$ is a variable to be compared with the Berry curvature. Figure 7 shows that the ef-

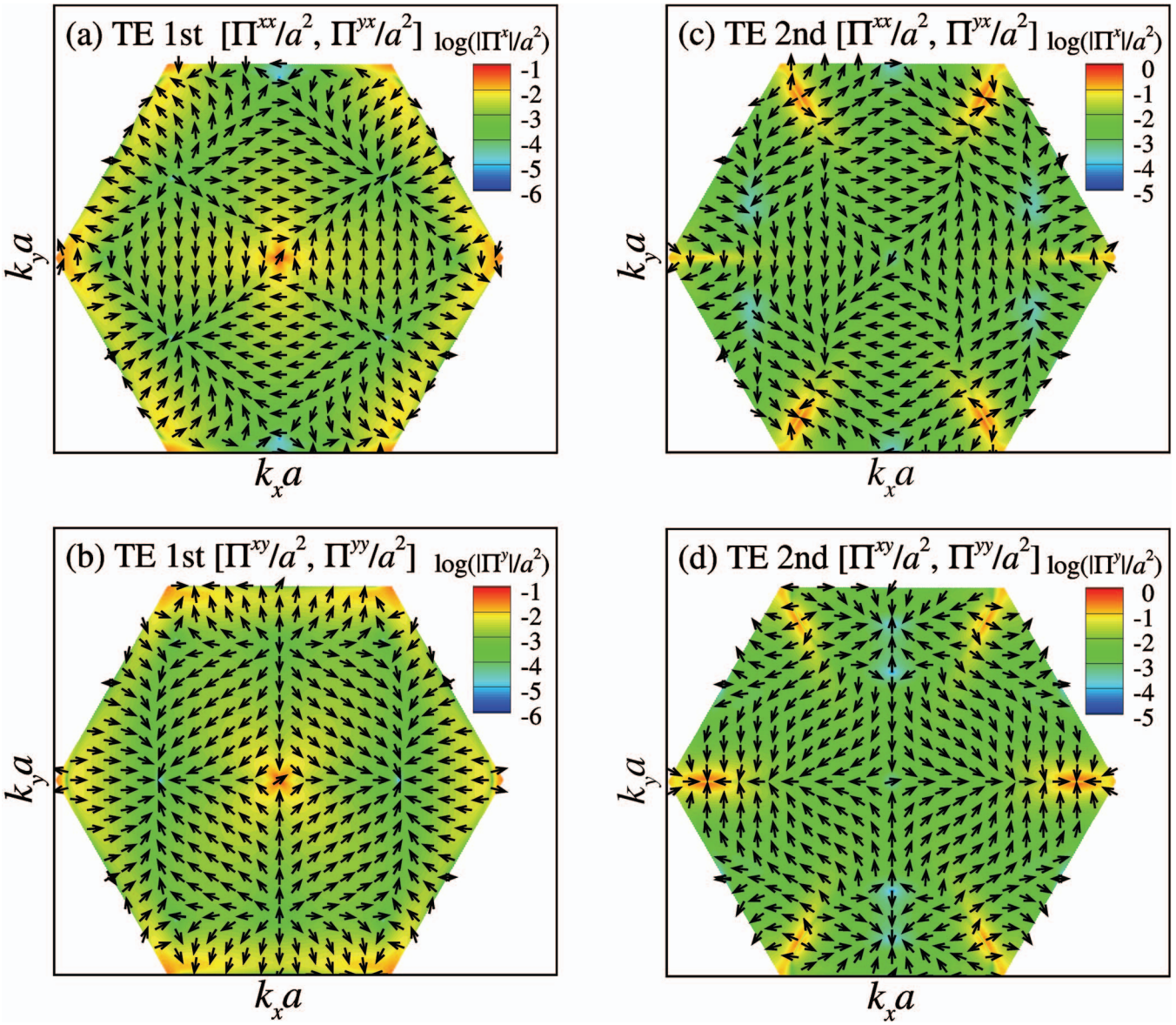

FIG. 7. (Color) $\vec{\vec{\Pi}}_{\mathrm{TE}nk}$, which is related to the correction of group velocity for each of (a) the TE first band with an x-directional modulation, (b) the TE first band with a y-directional modulation, (c) the TE second band with an x-directional modulation, and (d) the TE second band with a y-directional modulation. The base of a logarithm is 10.

fect of $\vec{\vec{\Pi}}_{\mathrm{TE}nk}$ is negligibly small compared to the effect of the Berry curvature in the present case. However, it is noted that, even when a Berry connection is nonzero, the corresponding Berry curvature can vanish. (This is easily understood by the analogy of a Berry connection and a Berry curvature to a vector potential and a magnetic field.) In such a case, $\mathbf{\Delta}_{\mathrm{TE}nk}$ and $\vec{\vec{\Pi}}_{\mathrm{TE}nk}$ are not necessarily minor corrections. These corrections may become measurable enough for a generic modulation superimposed onto a periodic structure, while, in this paper, we mainly consider a slowly varying modulation in order to assure the validity of our argument.

APPENDIX I: DIFFERENCE FROM MAGNETICALLY INDUCED DEFLECTIONS

It was proposed theoretically [39] and observed experimentally [40] that, in a Faraday-active random medium subject to a magnetic field perpendicular to an incident beam, the diffusion flow of light is deflected in a direction perpendicular to both the incident light beam and the externally applied magnetic field. This effect seems to be more similar to the conventional electrical Hall effect than the optical Hall effect is, because the effect is caused by the external magnetic field and the direction of deflection is perpendicular to it. However, it should be noted that, unlike electrons, photons are not charged, and their orbital motions do not directly couple to an external magnetic field. This effect is theoretically interpreted by the magnetically induced off-diagonal components of a diffusion tensor and experimentally proved to be due to the magnetically induced changes in the optical properties of scatterers [39,40]. In this sense, this effect is similar to the anomalous Hall effect due to the skew scattering mechanism, rather than to the conventional Hall effect. On the other hand, the optical Hall effect is originated by the

anomalous velocity of an optical wave packet which appears without external magnetic field or scatterers.

This kind of phenomenon, i.e., magnetically induced deflection, is not restricted to random media. The deflection of light by a magnetic field in a nonscattering homogeneous medium has also been discussed theoretically [41] and observed experimentally [42]. When the effect of absorption in a Faraday-active medium is negligible, the linear effect of external magnetic field $\boldsymbol{B}$ on this medium is described by the dielectric tensor,

$$\vec{\epsilon}_{ij} = n^2(\delta_{ij} + 2i\epsilon_{ijk}\Delta_k), \tag{I1a}$$

$$\boldsymbol{\Delta} = \frac{\gamma}{2n^2}\boldsymbol{B}, \tag{I1b}$$

where n is the refractive index of the medium in the case of $\boldsymbol{B}=0$, $\mathrm{Re}\gamma$ and $\mathrm{Im}\gamma$ represent the strength of the magnetic circular birefringence and that of magnetic circular dichroism respectively, while we set $\mathrm{Im}\gamma=0$. The eigenmodes of the dielectric displacement $\boldsymbol{D}=\vec{\epsilon}\boldsymbol{E}$ in such a medium are explicitly given in Ref. [43], and they are represented in terms of the orthogonal unit vectors $\boldsymbol{e}_k$, $\boldsymbol{e}_\theta$, and $\boldsymbol{e}_\phi$ in the spherical coordinate of the $\boldsymbol{k}$ space as

$$\boldsymbol{D}_+ \propto (\boldsymbol{e}_B\cdot\boldsymbol{e}_k)\boldsymbol{e}_\theta + i(C_B + \Delta|\boldsymbol{e}_B\times\boldsymbol{e}_k|^2)\boldsymbol{e}_\phi, \tag{I2a}$$

$$\boldsymbol{D}_- \propto (C_B + \Delta|\boldsymbol{e}_B\times\boldsymbol{e}_k|^2)\boldsymbol{e}_\theta - i(\boldsymbol{e}_B\cdot\boldsymbol{e}_k)\boldsymbol{e}_\phi, \tag{I2b}$$

$$C_B = \sqrt{(\boldsymbol{e}_B\cdot\boldsymbol{e}_k)^2 + \Delta^2|\boldsymbol{e}_B\times\boldsymbol{e}_k|^4}. \tag{I2c}$$

Here $\boldsymbol{e}_B$ is a unit vector defined by $\boldsymbol{\Delta}=\Delta\boldsymbol{e}_B$ with the condition $\boldsymbol{e}_B\cdot\boldsymbol{e}_k\geqslant 0$, and $\boldsymbol{e}_\phi\|\boldsymbol{e}_B\times\boldsymbol{e}_k$, $\boldsymbol{e}_\theta=\boldsymbol{e}_\phi\times\boldsymbol{e}_k$. (When $\mathrm{Im}\gamma\neq 0$, Δ is a complex-valued parameter.) These eigenmodes have the dispersion relations and the group velocities

$$E_{\pm,\boldsymbol{k}} = \frac{vk}{\sqrt{1-2\Delta^2|\boldsymbol{e}_B\times\boldsymbol{e}_k|^2 \mp 2\Delta C_B}}, \tag{I3a}$$

$$\boldsymbol{v}_{\pm,\boldsymbol{k}} = \frac{E_{\pm,\boldsymbol{k}}}{k}\left[\boldsymbol{e}_k \mp \frac{\Delta}{C_B}|\boldsymbol{e}_B\times\boldsymbol{e}_k|(\boldsymbol{e}_B\cdot\boldsymbol{e}_k)\boldsymbol{e}_\theta\right], \tag{I3b}$$

where $v=1/n$. The direction of the Poynting vector of each mode coincides with $\boldsymbol{v}_{\pm,\boldsymbol{k}}$ as long as $\mathrm{Im}\gamma=0$. It should be noted that the deflection occurs within the plane determined by $\boldsymbol{k}$ and $\boldsymbol{B}$. The angle $\delta\theta$ between the propagating directions of two eigenmodes with the same $\boldsymbol{k}$ is given in Ref. [43] and represented in the present notation as

$$\delta\theta = 2\arctan\frac{\Delta|\boldsymbol{e}_B\times\boldsymbol{e}_k|(\boldsymbol{e}_B\cdot\boldsymbol{e}_k)}{\sqrt{(\boldsymbol{e}_B\cdot\boldsymbol{e}_k)^2 + \Delta^2|\boldsymbol{e}_B\times\boldsymbol{e}_k|^4}}. \tag{I4}$$

For the exact Voigt geometry ($\boldsymbol{e}_B\cdot\boldsymbol{e}_k=0$), there appears no deflection [43,44]. The physics of this phenomenon is intuitively interpreted by considering the first order perturbation with respect to Δ and the situation in which the angles between $\boldsymbol{e}_B$ and $\boldsymbol{e}_k$ are not close to the Voigt geometry, i.e., $\boldsymbol{e}_B\cdot\boldsymbol{e}_k \gg |\Delta||\boldsymbol{e}_B\times\boldsymbol{e}_k|^2$. The approximated eigenvalues and group velocities are represented as follows;

$$E_{\pm,\boldsymbol{k}} \cong v(k \pm \boldsymbol{\Delta}\cdot\boldsymbol{k}), \tag{I5a}$$

$$\boldsymbol{v}_{\pm,\boldsymbol{k}} \cong v(\boldsymbol{e}_k \pm \boldsymbol{\Delta}). \tag{I5b}$$

This effect comes from the magnetically induced change in the dispersion relation of each mode due to the Pitaevskii magnetization, $\pm v\Delta\cdot\boldsymbol{k}$ [41]. On the other hand, the optical Hall effect is caused by the anomalous velocity due to the geometrical property of a wave packet.

In the above perturbative picture, $\boldsymbol{D}_\pm$ are approximately equivalent to right/left circularly polarized modes which have the spin angular momenta, $\pm\boldsymbol{e}_k$. Therefore the above interpretation based on the Pitaevskii magnetization means that an external magnetic field couples to the spin of a photon through a Faraday-active medium. From this consideration, we reasonably expect that an external magnetic field couples not only to the spin but also to a generic internal rotation of photon in the form of dipole coupling. (Consequently, this effect is expected for Laguerre-Gauss beams which have internal orbital angular momenta.) As shown in Sec. III C, there appear eigenmodes with large internal rotations in a two-dimensional photonic crystal without inversion symmetry. Here we take the configuration in which the photonic crystal is periodic in the xy plane and uniform along the z direction. Considering an eigenmode with $k_z=0$, its internal rotation is oriented in the z direction, i.e., perpendicular to its propagating direction. Thus when the photonic crystal is composed of Faraday-active media and subject to an external magnetic field, it is expected that the magnetically induced deflection can be enhanced. In addition, this effect would be observed even in the Voigt geometry. The details of this problem is beyond the scope of the present paper and will be discussed elsewhere. Here we just note that this effect in a Faraday-active photonic crystal is due to the magnetically induced change of dispersion relation as well as that in a homogeneous Faraday-active medium, and is different from the optical Hall effect in a photonic crystal discussed in Sec. III C.

[1] M. V. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984); J. Mod. Opt. **34**, 1401 (1987).

[2] *Geometrical Phases in Physics*, edited by A. Shapere and F. Wilczek (World Scientific, Singapore, 1989).

[3] *The Geometrical Phase in Quantum Systems*, edited by A. Bohm, A. Mostafazadeh, H. Koizumi, Q. Niu, and J. Zwanziger (Springer-Verlag, Berlin, 2003).

[4] R. Karplus and J. M. Luttinger, Phys. Rev. **95**, 1154 (1954).

[5] J. M. Luttinger, Phys. Rev. **112**, 739 (1958).

[6] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).

[7] M. Kohmoto, Ann. Phys. (N.Y.) **160**, 343 (1985).

[8] H. Aoki and T. Ando, Phys. Rev. Lett. **57**, 3093 (1986).

[9] M. Onoda and N. Nagaosa, J. Phys. Soc. Jpn. **71**, 19 (2002).

[10] T. Jungwirth, Q. Niu, and A. H. MacDonald, Phys. Rev. Lett. **88**, 207208 (2002).
[11] Z. Fang, N. Nagaosa, K. S. Takahashi, A. Asamitsu, R. Mathieu, T. Ogasawara, H. Yamada, M. Kawasaki, Y. Tokura, and K. Terakura, Science **302**, 92 (2003).
[12] S. Murakami, N. Nagaosa, and S.-C. Zhang, Science **301**, 1348 (2003).
[13] J. Sinova, D. Culcer, Q. Niu, N. A. Sinitsyn, T. Jungwirth, and A. H. MacDonald, Phys. Rev. Lett. **92**, 126603 (2004).
[14] S. M. Rytov, Dokl. Akad. Nauk SSSR **18**, 263 (1938).
[15] V. V. Vladimirski, Dokl. Akad. Nauk SSSR **31**, 222 (1941).
[16] S. Pancharatnam, *The Proceedings of the Indian Academy of Sciences* (1956) Vol. XLIV, No. 5, Sec. A, p. 247.
[17] R. Y. Chiao and Y. S. Wu, Phys. Rev. Lett. **57**, 933 (1986).
[18] A. Tomita and R. Y. Chiao, Phys. Rev. Lett. **57**, 937 (1986).
[19] M. V. Berry, Nature (London) **326**, 277 (1987).
[20] M. Onoda, S. Murakami, and N. Nagaosa, Phys. Rev. Lett. **93**, 083901 (2004).
[21] F. I. Fedorov, Dokl. Akad. Nauk SSSR **105**, 465 (1955).
[22] C. Imbert, Phys. Rev. D **5**, 787 (1972).
[23] D. G. Boulware, Phys. Rev. D **7**, 2375 (1973).
[24] N. Ashby and S. C. Miller, Jr., Phys. Rev. D **7**, 2383 (1973).
[25] H. Schilling, Ann. Phys. **16**, 122 (1965).
[26] V. G. Fedoseev, Opt. Spectrosc. **71**, 829 (1991); [Opt. Spectrosc. **71**, 483 (1991)].
[27] V. G. Fedoseev, Opt. Spektrosk. **71**, 992 (1991); [Opt. Spectrosc. **71**, 570 (1991)].
[28] F. Pillon, H. Gilles, and S. Girard, Appl. Opt. **43**, 1863 (2004).
[29] A. V. Dooghin, N. D. Kundikova, V. S. Liberman, and B. Ya. Zel'dovich, Phys. Rev. A **45**, 8204 (1992).
[30] V. S. Liberman and B. Ya. Zel'dovich, Phys. Rev. A **46**, 5199 (1992).
[31] K. Yu. Bliokh and Yu. P. Bliokh, Phys. Rev. E **70**, 026605 (2004).
[32] J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals* (Princeton University Press, Princeton, 1995).
[33] K. Sakoda, *Optical Properties of Photonic Crystals* (Springer, Berlin, 2005).
[34] R. Jackiw and A. Kerman, Phys. Lett. **71A**, 158 (1979).
[35] A. K. Pattanayak and W. C. Schieve, Phys. Rev. E **50**, 3601 (1994).
[36] M.-C. Chang and Q. Niu, Phys. Rev. B **53**, 7010 (1996).
[37] G. Sundaram and Q. Niu, Phys. Rev. B **59**, 14915 (1999).
[38] M. Born and E. Wolf, *Principles of Optics*, 7th edition (Cambridge University Press, Cambridge, England, 1999).
[39] B. A. van Tiggelen, Phys. Rev. Lett. **75**, 422 (1995).
[40] G. L. J. A. Rikken and B. A. van Tiggelen, Nature (London) **381**, 54 (1996).
[41] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984).
[42] G. L. J. A. Rikken and B. A. van Tiggelen, Phys. Rev. Lett. **78**, 847 (1997).
[43] G. W. 't Hooft, G. Nienhuis, and J. C. J. Paasschens, Phys. Rev. Lett. **80**, 1114 (1998).
[44] G. L. J. A. Rikken and B. A. van Tiggelen, Phys. Rev. Lett. **80**, 1115 (1998).
[45] Jiang Yimin and M. Liu, Phys. Rev. Lett. **90**, 099401 (2003).
[46] G. L. J. A. Rikken and B. A. van Tiggelen, Phys. Rev. Lett. **90**, 099402 (2003).
[47] F. D. M. Haldane and S. Raghu, e-print cond-mat/0503588.
[48] S. Raghu and F. D. M. Haldane, e-print cond-mat/0602501.
[49] K. Sawada and N. Nagaosa, Phys. Rev. Lett. **95**, 237402 (2005).
[50] K. Yu. Bliokh and Yu. P. Bliokh, Phys. Rev. Lett. **96**, 073903 (2006).
[51] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Yeshiva University, New York, 1964).
[52] M. Kristensen and J. P. Woerdman, Phys. Rev. Lett. **72**, 2171 (1994).
[53] Except for a simple system with a quadratic dispersion, the concept of "force" often becomes ambiguous, while the concept of "acceleration" is still well-defined. This issue is rather crucial for systems with spin-orbit interactions, i.e., systems which are relativistic in nature. Here, for the sake of convenience, we refer to the time derivative of (lattice) momentum as the driving force.
[54] Although the arguments in Refs. [36,37] are based on the first quantized formalism, the reformulation in the second quantized one is straightforward. Then the similar argument can be applicable to photonic systems. In a relativistic boson system like a photonic system, we cannot construct a positive definite provability density. This is why we have formulated our theory in the second quantized formalism.
[55] This form of the Berry curvature, $\mathbf{\Omega}_{\boldsymbol{k}}=\boldsymbol{k}/k^3\sigma_3$, is specific to the massless particle with spin 1. For the relativistic fermion with mass m, and spin-$1/2$, the Berry curvature in the helicity basis is expressed by
$$\mathbf{\Omega}_{\mathbf{k}}=\frac{1}{2E_{\boldsymbol{k}}^2}\left[\frac{m}{E_{\boldsymbol{k}}}(\sigma_1\boldsymbol{e}_\theta+\sigma_2\boldsymbol{e}_\phi)+\boldsymbol{e}_k\sigma_3\right],$$
where $E_{\boldsymbol{k}}=\sqrt{k^2+m^2}$. In the massless limit, this coincides with the Berry curvature of photon except for the overall coefficient due to different magnitude of spin. In the nonrelativistic limit, i.e., increasing m, with fixing k, the Berry curvature decreases as $1/m^2$, because the spin-orbit interaction also scales in the same manner.
[56] F. Goos and M. Hänchen, Ann. Phys. **1**, 333 (1947).
[57] J. D. Jackson, *Classical Electrodynamics*, 3rd edition (John Wiley & Sons, Inc., New York, 1999).
[58] C. Duval, Z. Horváth, and P. A. Horváthy, Phys. Rev. D **74**, 021701(R) (2006).
[59] L. Allen, S. M. Barnett, and M. J. Padgett, *Optical Angular Momentum* (Institute of Physics Publishing, Bristol and Philadelphia, 2003).
[60] V. G. Fedoseyev, Opt. Commun. **193**, 9 (2001).
[61] R. Dasgupta and P. K. Gupta, Opt. Commun. **257**, 91 (2006).
[62] H. Okuda and H. Sasada, Opt. Express **14**, 8393 (2006).
[63] M. S. Kushwaha, P. Halevi, L. Dobrzynski, and B. Djafari-Rouhani, Phys. Rev. Lett. **71**, 2022 (1993).
[64] M. S. Kushwaha, P. Halevi, G. Martínez, L. Dobrzynski, and B. Djafari-Rouhani, Phys. Rev. B **49**, 2313 (1994).