

# Charge Conservation beyond Uniformity: Spatially Inhomogeneous Electromagnetic Response in Periodic Solids

Robert C. McKay<sup>1</sup>,<sup>\*</sup> Fahad Mahmood,<sup>2,3</sup> and Barry Bradlyn<sup>1,\*</sup>

<sup>1</sup>*Department of Physics and Institute for Condensed Matter Theory,  
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801-3080, USA*

<sup>2</sup>*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

<sup>3</sup>*Materials Research Laboratory, University of Illinois at Urbana-Champaign,  
Urbana, Illinois 61801, USA*



(Received 25 September 2023; revised 23 February 2024; accepted 1 March 2024; published 27 March 2024)

Nonlinear electromagnetic response functions have reemerged as a crucial tool for studying quantum materials, due to recently appreciated connections between optical response functions, quantum geometry, and band topology. Most attention has been paid to responses to spatially uniform electric fields, relevant to low-energy optical experiments in conventional solid state materials. However, magnetic and magnetoelectric phenomena are naturally connected by responses to spatially varying electric fields due to Maxwell's equations. Furthermore, in the emerging field of moiré materials, characteristic lattice scales are much longer, allowing spatial variation of optical electric fields to potentially have a measurable effect in experiments. In order to address these issues, we develop a formalism for computing linear and nonlinear responses to spatially inhomogeneous electromagnetic fields. Starting with the continuity equation, we derive an expression for the second-quantized current operator that is manifestly conserved and model independent. Crucially, our formalism makes no assumptions on the form of the microscopic Hamiltonian and so is applicable to model Hamiltonians derived from tight-binding or *ab initio* calculations. We then develop a diagrammatic Kubo formalism for computing the wave vector dependence of linear and nonlinear conductivities, using Ward identities to fix the value of the diamagnetic current order by order in the vector potential. We apply our formula to compute the magnitude of the Kerr effect at oblique incidence for a model of a moiré-Chern insulator and demonstrate the experimental relevance of spatially inhomogeneous fields in these systems. We further show how our formalism allows us to compute the (orbital) magnetic multipole moments and magnetic susceptibilities in insulators. Turning to nonlinear response, we use our formalism to compute the second-order transverse response to spatially varying transverse electric fields in our moiré-Chern insulator model, with an eye toward the next generation of experiments in these systems.

DOI: [10.1103/PhysRevX.14.011058](https://doi.org/10.1103/PhysRevX.14.011058)

Subject Areas: Condensed Matter Physics

## I. INTRODUCTION

The study of optical properties of materials has long been one of the guiding themes in condensed matter physics. From the ability of x-ray crystallography to determine crystal structure to spectroscopic probes of band structure and collective dynamics, optical response experiments have served as a vital tool to learn about the structure of solid state systems. On the theoretical side, sum rules for optical response functions provide some of the few exact and experimentally relevant results in quantum many-body

physics [1–4]. Combined with first-principles calculations, the Kubo formula for linear and nonlinear optical responses has allowed for an understanding of the electronic properties of materials [5–10]. More recently, it has come to be appreciated that optical response functions such as linear and nonlinear Hall conductivity [11–14], the photogalvanic effect [15–17], magnetoelectric polarizability [18–22], and nonlinear magnetoresistance [23–27] all receive large topological contributions and can be used to both diagnose and functionalize topological materials.

The fundamental objects of interest for optical responses are the (nonlinear) conductivity tensors of a material, which relate the measured current density  $\langle \mathbf{j}(\mathbf{r}, t) \rangle$  at position  $\mathbf{r}$  and time  $t$  that flow in a material in response to an electric field  $\mathbf{E}(\mathbf{r}, t)$ . Typically, the electric field  $\mathbf{E}(\mathbf{r}, t)$  is introduced via a voltage source or an optical probe. The typical objects of interest are the experimentally measurable (nonlinear) conductivities

<sup>\*</sup>bbradlyn@illinois.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

$$\sigma_{(n)}^{\mu,\nu_1,\dots,\nu_n}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_n, t, t_1, \dots, t_n), \quad (1)$$

which determine the measured current via

$$\langle j^\mu(\mathbf{r}, t) \rangle = \sum_n \int \prod_i d\mathbf{r}_i dt_i \sigma_{(n)}^{\mu,\nu_1,\dots,\nu_n}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_n, t, t_1, \dots, t_n) E_{\nu_1}(\mathbf{r}_1, t_1) \times \dots \times E_{\nu_n}(\mathbf{r}_n, t_n). \quad (2)$$

Here and throughout, we use Greek indices  $\mu, \nu = 1, 2, 3$  to index spatial components of vectors, and the Einstein summation convention is used for repeated indices. Since we work entirely in flat space time, there is no functional distinction between upper and lower indices; we try to choose the index arrangement that makes expressions easiest to parse. Equation (2) parametrizes the current density order by order in powers of the external electric field. At the order of  $n = 1$ , we recover the familiar generalized Ohm's law with  $\sigma_1^{\mu\nu}(\mathbf{r}, \mathbf{r}_1, t, t_1) = \sigma^{\mu\nu}(\mathbf{r}, \mathbf{r}_1, t - t_1)$  being the linear conductivity tensor. Note that in writing Eq. (2) we have not assumed that our system is translationally invariant.

Because the speed of light is large compared with typical velocity scales in a solid state material and because the typical energy scales of interest are in the microwave, terahertz, and optical regimes, the typical wavelength of the probe electric field (on the order of thousands of angstroms for visible light all the way up to the order of meters for ac fields) is significantly larger than the typical lattice spacing of crystals (on the order of nanometers). This means that, for most experimentally relevant situations, we can ignore the spatial variation of the electric field [28]. In this approximation, we can introduce the uniform optical conductivities  $\sigma_n^{\mu,\nu_1,\dots,\nu_n}(t, t_1, \dots, t_n)$  which depend on time and not space and which determine the spatially averaged current density  $\langle \mathbf{j}(t) \rangle$  via

$$\langle j^\mu(t) \rangle = \sum_n \int \prod_i dt_i \sigma_n^{\mu,\nu_1,\dots,\nu_n}(t, t_1, \dots, t_n) \prod_i E_{\nu_i}(t_i). \quad (3)$$

Equation (3) follows from Eq. (2) after assuming the electric field is spatially uniform.

The uniform optical conductivities can be computed using generalizations of the Kubo formula, in terms of correlation functions of the electronic velocity operator  $\mathbf{v}$ . Recent theoretical work [29–33] has developed streamlined and efficient analytical machinery for computing  $\sigma_n^{\mu,\nu_1,\dots,\nu_n}(t, t_1, \dots, t_n)$  for a variety of interesting materials, for finite systems, continuum crystalline systems, and crystalline systems in the tight-binding approximation. For noninteracting electron systems, recent work has shown how the Kubo formula for the uniform optical conductivities can be interpreted in terms of the quantum geometry of electronic wave functions in the systems [34–36]. This theoretical work has been complemented by advances in nonlinear spectroscopy, which have seen signatures of intriguing

topological and band-geometric effects in quantum materials [37,38].

That said, there are limitations to focusing on the uniform optical conductivity. First, the uniform optical conductivity captures response to spatially uniform electric fields but does not capture response to magnetic fields. To understand this, we can see from Faraday's law

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (4)$$

that a spatially uniform and time-varying magnetic field generates a spatially inhomogeneous (transverse) electric field. Magnetic and magnetoelectric responses are thus encoded in the gradient expansion of  $\sigma_{(n)}^{\mu,\nu_1,\dots,\nu_n}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_n, t, t_1, \dots, t_n)$ . Such a gradient expansion has been carried out to low orders to obtain, for instance, the Streda formula relating the Hall conductivity to the magnetization density [39,40], the Kubo formula for the magnetoelectric polarizability [41], and optical gyrotropy [42]. Similarly, a gradient expansion of the second-order ( $n = 2$ ) longitudinal optical conductivity yielded diagrammatic Kubo formulas for electric quadrupole transitions in periodic crystals [43]. Lastly, the study of electron hydrodynamics has brought renewed attention to the connection between the viscosity tensor of a Galilean-invariant electron fluid and the (linear) optical conductivity expanded to second order in gradients [44–50]. Attempts to generalize this relationship to periodic solids have yielded unintuitive results [51].

Second, the explosion of interest in moiré materials [52–61] gives cause to reevaluate the focus on the uniform optical conductivity. In systems such as twisted multilayer graphene, twisted transition metal dichalcogenides, and lattice-mismatched van der Waals interfaces, the effective moiré lattice constant can be several orders of magnitude larger than the atomic spacing in individual layers. In such a system, the wavelength of optical frequency light can be an appreciable fraction of the moiré lattice spacing. Thus, we can expect that optical properties such as Kerr and Faraday rotation, as well as nonlinear shift current and second-harmonic generation, may obtain measurable corrections from the spatial inhomogeneity of the applied optical electric field. Recent advances in ultrafast spectroscopy have put the study of nonlinear optical effects in moiré and van der Waals materials within reach of modern experiments [50,62,63].

Thus, in order to study both generalized magnetoelectric responses and optical properties of moiré materials, it would be desirable to have a complete theory of spatially inhomogeneous electromagnetic response applicable to periodic systems. From the Kubo formula, we know that  $\sigma_{(n)}^{\mu, \nu_1, \dots, \nu_n}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_n, t, t_1, \dots, t_n)$  can be computed from a retarded correlation function of  $n$  current operators  $\mathbf{j}(\mathbf{r})$ . When the full microscopic Hamiltonian  $H_0$  for a system of interest is known, the operator  $\mathbf{j}(\mathbf{r})$  can be identified via minimal coupling, by replacing the momentum of each particle  $\mathbf{p}_i$  by the covariant combination  $\mathbf{p}_i - \bar{e}\mathbf{A}(\mathbf{x}_i)$ , where  $\mathbf{A}(\mathbf{x}_i)$  is the electromagnetic vector potential as a function of the position operator  $\mathbf{x}_i$  of particle  $i$  and  $\bar{e}$  is the charge of particle  $i$  [64,65]. Denoting by  $H_A$  the Hamiltonian we obtain after minimal coupling, we can define the current operator as

$$\mathbf{j}_A(\mathbf{r}) = -\frac{\delta H_A}{\delta \mathbf{A}(\mathbf{r})}, \quad (5)$$

where  $\delta/\delta \mathbf{A}(\mathbf{r})$  denotes a variational derivative with respect to the vector potential. This allows us to write

$$H_A = H_0 - \int d\mathbf{r} \mathbf{j}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r}), \quad (6)$$

from which the nonlinear conductivities can be extracted via standard response theory upon identifying (in this gauge [64,65]) the electric field  $\mathbf{E} = -\partial \mathbf{A}/\partial t$ . Crucially, the current defined via Eq. (5) manifestly obeys the continuity equation (charge conservation equation)

$$\partial_t \rho(\mathbf{r}) + \nabla \cdot \mathbf{j}_A(\mathbf{r}) = 0, \quad (7)$$

where  $\rho(\mathbf{r})$  is the charge density operator. Equation (7) reflects the conservation of electric charge and is equivalent to gauge invariance.

While this gives an in-principle complete procedure for computing  $\sigma_{(n)}^{\mu, \nu_1, \dots, \nu_n}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_n, t, t_1, \dots, t_n)$ , it has one main drawback: In most situations in solid state physics, we do not have access to the full microscopic Hamiltonian but instead only an effective low-energy model for the degrees of freedom of interest. This can arise due to an approximate treatment of relativistic effects of core electrons in a solid (as in, e.g., a density functional theory calculation) or via the use of a tight-binding model that describes only the low-energy subset of states in the full Hilbert space of a material. When only an effective or approximate Hamiltonian is known, the minimal coupling substitution  $H_0 \rightarrow H_A$  cannot be carried out uniquely.

One common approach [41,42,66–68] to circumvent this problem in nonrelativistic materials is to *assume* that the full and unknown microscopic electronic Hamiltonian has the form

$$H_0 \stackrel{!}{=} \sum_i \frac{|\mathbf{p}_i|^2}{2m} + V(\mathbf{x}_i) - \vec{\lambda}(\mathbf{x}_i) \cdot (\mathbf{p} \times \vec{\sigma}) + \frac{1}{2} \sum_{i \neq j} U(\mathbf{x}_i - \mathbf{x}_j), \quad (8)$$

where  $m$  is the electron mass,  $V$  is the periodic potential of the ions in the system,  $\vec{\lambda}$  is the approximate spin-orbit potential (usually expressed in terms of gradients of  $V$ ),  $\vec{\sigma}$  is a vector of Pauli matrices acting on electron spins, and  $U$  is the Coulomb potential. With this assumption, the minimally coupled current operator can be evaluated exactly. In the absence of the external electromagnetic field, it is given by

$$\mathbf{j}_{\min}(\mathbf{r}) \equiv -\frac{\delta H_A}{\delta \mathbf{A}(\mathbf{r})} \Big|_{\mathbf{A}=0} = \sum_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}} \mathbf{j}_{\mathbf{q}, \min}, \quad (9)$$

where the Fourier components  $\mathbf{j}_{\mathbf{q}, \min}$  can be expressed in two equivalent and often used forms:

$$\mathbf{j}_{\mathbf{q}, \min} = \frac{\bar{e}}{2} \sum_i \{e^{-i\mathbf{q} \cdot \mathbf{x}_i}, \mathbf{v}_i\} = \tilde{\mathbf{j}}_{\mathbf{q}} \quad (10)$$

$$= \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i/2} \mathbf{v}_i e^{-i\mathbf{q} \cdot \mathbf{x}_i/2} = \mathbf{j}_{\text{mid}, \mathbf{q}}, \quad (11)$$

where  $\mathbf{v}_i = i[H, \mathbf{x}_i]$  is the velocity operator for particle  $i$  and  $\bar{e}$  is the charge of the particles (which for convenience we assume to be the same for all particles throughout this work). The customary approach is to then take either Eq. (10) or Eq. (11) as the definition of the current operator in the low-energy effective model and use that to compute optical response functions.

There are two problems with this approach. The first is that, although Eqs. (10) and (11) are equivalent for the microscopic Hamiltonian Eq. (8), they will, in general, give different results when applied to any effective model. This raises the question of which, if either, of them should be used to compute optical response functions. Additionally, neither Eq. (10) nor Eq. (11) coincides with the minimally coupled current  $\mathbf{j}_A(\mathbf{r})$  when relativistic corrections to the kinetic energy are incorporated, which may be important for core electrons in solid state systems. Even more severe, we show below that generally neither Eq. (10) nor Eq. (11) satisfies the continuity equation (7) even when  $\mathbf{A} = 0$ . When used in the context of effective low-energy models for even nonrelativistic systems, this means that Eqs. (10) and (11) represent non-number-conserving approximations to the minimally coupled current. Thus, neither Eq. (10) nor Eq. (11) provides a satisfactory starting point for the computation of optical response functions in low-energy models of quantum materials.

In what follows, we develop a theory of spatially inhomogeneous linear and nonlinear optical response that

manifestly obeys the continuity equation (7) when applied to arbitrary effective models of condensed matter systems. First, in Sec. II, we show how to construct the position-dependent current operator for a generic Hamiltonian that manifestly respects the continuity equation. Crucially, our construction makes no assumptions on the form of the Hamiltonian and is applicable to non-relativistic, semirelativistic, and approximately tight-binding Hamiltonians. We show that our current agrees with Eqs. (10) and (11) to linear order in the wave vector  $\mathbf{q}$ , implying that our theory gives the same linear response to uniform electric and magnetic fields as in Refs. [18,41,42,66]. We also show how to define the diamagnetic current such that the continuity equation is satisfied order by order in the electromagnetic vector potential.

Next, in Sec. III, we set out the Feynman diagrammatic rules for evaluating the spatially inhomogeneous conductivity in Fourier space as a function of frequency and wave vector (where the uniform optical conductivities correspond to the limit of zero wave vector). Focusing on the linear conductivity, we derive a modified  $f$ -sum rule relating the diamagnetic conductivity to the density-density response function. In Sec. V, we explore various observable phenomena that arises from our wave-vector-dependent response theory at the linear response level, applied to noninteracting electron systems in a periodic potential. As a proof of concept, we examine the wave vector dependence of the Hall conductivity in a time-reversal-breaking Weyl semimetal model first. We then show how our formulation of the wave-vector-dependent linear conductivity can be applied to compute the Kerr rotation and ellipticity in a model of a moiré-Chern insulator. We compare our predictions with an analogous calculation using the current operators in Eqs. (10) and (11), showing that our gauge-invariant formulation predicts a measurably different Kerr response on the order of arcseconds to arcminutes for typical moiré length and energy scales. We round off this section by examining the magnetic properties of insulators. We analyze the relationship between magnetic susceptibility and transverse conductivity, present expressions for the magnetic quadrupole moment, and give a new derivation of the Streda formula for the Hall conductivity.

Then, in Sec. VI, we move on from phenomena derived from linear conductivity and evaluate the diagrammatic Kubo formula for the second-order electromagnetic conductivity for noninteracting electrons in a periodic potential. We then focus on the component of the second-order conductivity that is second-harmonic generation in frequency, and self-focusing in wave vector, which is relevant to future ultrafast optical experiments. We show how to compute this component of the conductivity for our model of a moiré-Chern insulator. Finally, we conclude our paper by summarizing the key results and suggesting new directions in Sec. VII.

## II. THE CURRENT OPERATOR

Motivated by recent efforts to obtain a complete spatially inhomogeneous electromagnetic response theory [43,51,66,68], our goal in this section is to derive the Fourier components of the conserved current operator in a model-independent fashion. Measurable phenomena such as conductivities, optical response, and magnetization can be accurately modeled only in terms of a charge-conserving current operator. As we show, the conventionally defined current densities Eqs. (11) and (10) fail to satisfy the continuity equation (7) for generic Hamiltonians.

To find the conserved current density, let us consider a general Hamiltonian

$$H_0 = \sum_i T(\mathbf{p}_i) + V(\mathbf{x}_i, \mathbf{p}_i) + \frac{1}{2} \sum_{i \neq j} U(\mathbf{x}_i - \mathbf{x}_j) \quad (12)$$

for a (possibly interacting) system of electrons. Here,  $i = 1 \dots N$  indexes the particles of the system,  $\mathbf{p}_i$  is the momentum operator for particle  $i$ , and  $\mathbf{x}_i$  is the position operator for particle  $i$ . Also,  $T(\mathbf{p}_i)$  is the kinetic energy operator for the  $i$ th particle. Although this kinetic operator is usually quadratic in momenta, it can extend to semi-relativistic systems. For example, the kinetic energy could take the form

$$T(\mathbf{p}_i) \approx \frac{|\mathbf{p}_i|^2}{2m} + \frac{|\mathbf{p}_i|^4}{8m^3 c^2} + \dots, \quad (13)$$

which is relevant for all-electron *ab initio* modeling of spin-orbit coupled materials [69]. Our results hold for arbitrary  $T(\mathbf{p}_i)$ . The single-particle potential  $V(\mathbf{x}_i, \mathbf{p}_i)$  includes both the momentum-independent external potential as well as momentum-dependent external potential terms such as the spin-orbit potential. Finally,  $U(\mathbf{x}_i - \mathbf{x}_j)$  is the interaction potential, which we take to be momentum independent for simplicity (we expect this is a good assumption for most models of interest; however, for long-range interacting models of Mott insulators that have recently attracted attention, this assumption must be checked [70]). We assume that the potential  $V$  has discrete translation symmetry such that

$$\begin{aligned} T_{\mathbf{R}}^\dagger H T_{\mathbf{R}} &= e^{i\mathbf{R} \cdot \sum_i \mathbf{p}_i} H e^{-i\mathbf{R} \cdot \sum_i \mathbf{p}_i} \\ &= \sum_i T(\mathbf{p}_i) + V(\mathbf{x}_i + \mathbf{R}, \mathbf{p}_i) + \frac{1}{2} \sum_{i \neq j} U(\mathbf{x}_i - \mathbf{x}_j) \\ &= H, \end{aligned} \quad (14)$$

where  $\mathbf{R}$  is a Bravais lattice vector.

The global  $U(1)$  symmetry of  $H_0$  implies that the density operator

$$\rho(\mathbf{r}) = \bar{e} \sum_i \delta(\mathbf{r} - \mathbf{x}_i) \quad (15)$$



satisfies the continuity equation (7), where all particles have charge  $\bar{e}$ . We can rewrite the continuity equation in the Heisenberg picture as

$$\partial_t \rho(\mathbf{r}) = i[H_0, \rho(\mathbf{r})] = -\nabla \cdot \mathbf{j}_{\min}(\mathbf{r}), \quad (16)$$

where we work in natural units where  $\hbar = c = 1$  unless stated otherwise. In the presence of an external electromagnetic field with scalar potential  $A_0$  and vector potential  $\mathbf{A}$ , we have that the minimally coupled Hamiltonian becomes

$$H_0 \rightarrow H_A = \sum_i T[\mathbf{p}_i - \bar{e}\mathbf{A}(\mathbf{x}_i)] + V[\mathbf{x}_i, \mathbf{p}_i - \bar{e}\mathbf{A}(\mathbf{x}_i)] - \bar{e}A_0(\mathbf{x}_i) + \frac{1}{2} \sum_{i \neq j} U(\mathbf{x}_i - \mathbf{x}_j). \quad (17)$$

We can take the functional derivative with respect to the scalar potential to arrive at

$$\rho_A(\mathbf{r}) = -\frac{\delta H}{\delta A_0(\mathbf{r})} = \rho(\mathbf{r}), \quad (18)$$

where  $\rho_A(\mathbf{r})$  indicates the density operator in the presence of the electromagnetic field,  $\rho(\mathbf{r})$  is defined in Eq. (15), and we have made use of the identity

$$\frac{\delta A_0(\mathbf{x}_i)}{\delta A_0(\mathbf{r})} = \delta(\mathbf{r} - \mathbf{x}_i). \quad (19)$$

Furthermore, we have that

$$\partial_t \rho_A(\mathbf{r}) = i[H_A, \rho_A(\mathbf{r})] = -\nabla \cdot \mathbf{j}_A(\mathbf{r}), \quad (20)$$

where by evaluating the commutator we find

$$j_A^\mu(\mathbf{r}) = -\frac{\delta H_A}{\delta A_\mu(\mathbf{r})}. \quad (21)$$

Equation (21) allows us to define the current in terms of the Hamiltonian  $H_A$  minimally coupled to the external electromagnetic field. In particular, we can find the unperturbed current operator  $\mathbf{j}_{\min}(\mathbf{r})$  from Eq. (16) as

$$j_{\min}^\mu(\mathbf{r}) = -\frac{\delta H_A}{\delta A_\mu(\mathbf{r})} \Big|_{\mathbf{A}=0}. \quad (22)$$

However, in order to utilize Eq. (22), we need to know the explicit microscopic form of the Hamiltonian in Eq. (12). In many cases of interest, we do not have this information available. For instance, in modeling electronic systems we often have access to only Wannier-based tight-binding models, or other effective Hamiltonians, that are projected into a subset of bands of interest. For an effective Hamiltonian, it is not possible to implement the minimal coupling of Eq. (17) in a model-independent fashion. One possible way to circumvent this complication is to make

use of a Peierls substitution to couple a Wannierized effective model to the vector potential, provided the Wannier orbitals are sufficiently well localized and the vector potential varies sufficiently slowly [4]. However, defining the current via a Peierls substitution requires making choices for how to evaluate path integrals of the vector potential between Wannier centers that are not necessarily natural. It also requires assuming that the vector potential varies slowly on the length scale of the Wannier function localization length, which may not be appropriate for all models (especially for models of moiré or topologically nontrivial systems). Finally, the Peierls substitution cannot be used to define the current when a Wannierized tight-binding model has not been explicitly constructed.

In order to circumvent these issues, we derive an alternative expression for the current  $\mathbf{j}(\mathbf{r})$  that is manifestly conserved and that does not require detailed knowledge of the microscopic Hamiltonian. To do so, let us first recast the continuity equation [Eq. (16)] in momentum space. We introduce the Fourier-transformed density

$$\rho_{\mathbf{q}} = \int d\mathbf{r} e^{-i\mathbf{q} \cdot \mathbf{r}} \rho(\mathbf{r}) = \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i}. \quad (23)$$

In Fourier space, the continuity equation reads

$$\partial_t \rho_{\mathbf{q}} = i[H_0, \rho_{\mathbf{q}}] = -i\mathbf{q} \cdot \mathbf{j}_{\mathbf{q}}, \quad (24)$$

where

$$\mathbf{j}_{\mathbf{q}} = \int d\mathbf{r} e^{-i\mathbf{q} \cdot \mathbf{r}} \mathbf{j}(\mathbf{r}). \quad (25)$$

Rather than attempt to evaluate the commutator in Eq. (24) directly, we take a more general approach. In the Heisenberg picture, we see that  $\rho_{\mathbf{q}}$  in Eq. (23) depends on time implicitly through the operators  $\mathbf{x}_i$ . Defining the single-particle velocity operators

$$\mathbf{v}_i \equiv \partial_t \mathbf{x}_i = i[H_0, \mathbf{x}_i], \quad (26)$$

we can write

$$\rho_{\mathbf{q}}(t + \delta t) = \bar{e} e^{-i\mathbf{q} \cdot \mathbf{x}_i(t + \delta t)} \approx \bar{e} e^{-i\mathbf{q} \cdot (\mathbf{x}_i(t) + \delta t \mathbf{v}_i(t))}. \quad (27)$$

Thus, using the definition of the derivative, the rate of change of the density operator is

$$\partial_t \rho_{\mathbf{q}} = \bar{e} \lim_{\delta t \rightarrow 0} \frac{e^{-i\mathbf{q} \cdot (\mathbf{x}_i + \delta t \mathbf{v}_i)} - e^{-i\mathbf{q} \cdot \mathbf{x}_i}}{\delta t}. \quad (28)$$

To simplify this further, we make use of a general result of Karplus and Schwinger [71], who showed that, for any operators  $A$  and  $B$ ,

$$e^{A + \delta t B} = e^A + \delta t \int_0^1 d\lambda e^{(1-\lambda)A} B e^{\lambda A} + \mathcal{O}[(\delta t)^2]. \quad (29)$$

We prove this result in Appendix A. Inserting Eq. (29) with  $A = -i\mathbf{q} \cdot \mathbf{x}_i$  and  $B = -i\mathbf{q} \cdot \mathbf{v}_i$  into Eq. (28), we find that

$$\partial_i \rho_{\mathbf{q}} = \bar{e} \sum_i \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q} \cdot \mathbf{x}_i} (-i\mathbf{q} \cdot \mathbf{v}_i) e^{-i\lambda\mathbf{q} \cdot \mathbf{x}_i}. \quad (30)$$

Comparing with the continuity equation in Eq. (24), we can identify

$$\mathbf{j}_{\mathbf{q}} = \bar{e} \sum_i \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q} \cdot \mathbf{x}_i} \mathbf{v}_i e^{-i\lambda\mathbf{q} \cdot \mathbf{x}_i}. \quad (31)$$

Equation (31) has several desirable features. Most importantly, it is manifestly conserved. Second, for nonrelativistic Hamiltonians [where  $\mathbf{v}(\mathbf{p}, \mathbf{x})$  is a linear function of  $\mathbf{p}$ ], a short calculation (see Appendix B) shows that  $\mathbf{j}_{\mathbf{q}}$  coincides with the minimally coupled current. Note that, since the Hamiltonian  $H_0$  commutes with the projection operator onto a set of low-energy bands, the projected low-energy density operator satisfies Eq. (16) with the conserved current given by the low-energy projection of Eq. (31). Thus, for nonrelativistic systems with microscopic Hamiltonians of the form of Eq. (8), Eq. (31) is the approximation to the minimally coupled current in Eq. (5) that is manifestly conserved in any low-energy approximation to the Hamiltonian. However, we emphasize that a choice is made in going from Eq. (30) to Eq. (31): We can add any operator  $\delta\mathbf{j}_{\mathbf{q}}$  that is purely transverse ( $\mathbf{q} \cdot \delta\mathbf{j} = 0$ ) to Eq. (31) without changing the continuity equation (16). Different choices of  $\delta\mathbf{j}$  correspond to different definitions for the (orbital) *magnetization current*, which do not contribute to transport and are not constrained by gauge invariance. As such, it is, in principle, not possible to fix  $\delta\mathbf{j}$  unambiguously from effective Hamiltonians alone. Even in the context of the microscopic Hamiltonian nonrelativistic approximations to the Dirac equation, nonminimal coupling to the electromagnetic field can modify the definition of the magnetization current. Faced with these challenges, we take the natural choice  $\delta\mathbf{j} = 0$ —corresponding to our definition Eq. (31) for the current operator—as our working, model-independent definition of the current.

As we now show, Eq. (31) resolves many of the problems of the nonconserved choices Eqs. (10) and (11) for the current. In Sec. II A, we show that both the longitudinal and transverse components of Eqs. (10) and (11) differ from those of Eq. (31) at second order in wave vector. Then, in Sec. II B, we compute the matrix elements of the conserved current Eq. (31) in the basis of Bloch eigenstates, allowing us to apply our formalism to compute electromagnetic responses. Finally, in Sec. II C, we show that the formula of Karplus and Schwinger can be used to define the generalized wave-vector-dependent diamagnetic contributions to the current operator in the presence of a nonzero electromagnetic field.

### A. The nonconserved current operators

Equation (31) is our first main result. Let us compare it quantitatively with the (nonconserved) current operators Eqs. (10) and (11), which have been used to compute spatially inhomogeneous electromagnetic response functions to low order in the wave vector [51,66]. Recall that Eq. (10) proposes the definition

$$\tilde{\mathbf{j}}_{\mathbf{q}} = \frac{\bar{e}}{2} \sum_i \{e^{-i\mathbf{q} \cdot \mathbf{x}_i}, \mathbf{v}_i\} \quad (32)$$

for the current operator. We refer to this as the “trapezoidal” current, since it is the trapezoidal approximation the integration over  $\lambda$  in our conserved current Eq. (31).

Similarly, Eq. (11) proposes the commonly used [51,66,68] definition

$$\mathbf{j}_{\text{mid},\mathbf{q}} = \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i/2} \mathbf{v}_i e^{-i\mathbf{q} \cdot \mathbf{x}_i/2}. \quad (33)$$

We refer to this as the “midpoint” current, since it is the midpoint approximation the integration over  $\lambda$  in our conserved current Eq. (31).

To see how the definitions of current operator from Eqs. (32) and (33) compare with our Eq. (31), it is helpful to write  $\mathbf{v}_i = \mathbf{v}_i(\mathbf{p}_i, \mathbf{x}_i)$  to make explicit the  $\mathbf{p}_i$  and  $\mathbf{x}_i$  dependence of  $\mathbf{v}_i$ . Inserting this into Eq. (31) and using the fact that

$$e^{i\alpha \cdot \mathbf{x}_i} f(\mathbf{p}_i, \mathbf{x}_i) = f(\mathbf{p}_i - \alpha, \mathbf{x}_i) e^{i\alpha \cdot \mathbf{x}_i}, \quad (34)$$

we find

$$\begin{aligned} \mathbf{j}_{\mathbf{q}} &= \bar{e} \sum_i \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q} \cdot \mathbf{x}_i} \mathbf{v}_i(\mathbf{p}_i, \mathbf{x}_i) e^{-i\lambda\mathbf{q} \cdot \mathbf{x}_i} \\ &= \bar{e} \sum_i \int_0^1 d\lambda e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{v}_i(\mathbf{p}_i - \lambda\mathbf{q}, \mathbf{x}_i). \end{aligned} \quad (35)$$

Similarly, using Eq. (34) to simplify Eq. (32) yields

$$\tilde{\mathbf{j}}_{\mathbf{q}} = \bar{e} \frac{1}{2} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} (\mathbf{v}_i(\mathbf{p}_i, \mathbf{x}_i) + \mathbf{v}_i(\mathbf{p}_i - \mathbf{q}, \mathbf{x}_i)). \quad (36)$$

Lastly, using Eq. (34) to simplify Eq. (33) yields

$$\mathbf{j}_{\text{mid},\mathbf{q}} = \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{v}_i(\mathbf{p}_i - \mathbf{q}/2, \mathbf{x}_i). \quad (37)$$

Thus, we generically have that  $\mathbf{j}_{\mathbf{q}} \neq \tilde{\mathbf{j}}_{\mathbf{q}} \neq \mathbf{j}_{\text{mid},\mathbf{q}}$ . Note that if  $\mathbf{v}_i$  is a linear function of momentum (which occurs when the kinetic energy includes no relativistic corrections and the spin-orbit potential is linear in momentum), then we can carry out the integral in Eq. (35) to find that  $\mathbf{j}_{\mathbf{q}} = \tilde{\mathbf{j}}_{\mathbf{q}} = \mathbf{j}_{\text{mid},\mathbf{q}}$ . However, for Hamiltonians with relativistically corrected kinetic energy and complicated

spin-orbit potentials, this will not be the case. This is relevant not just for *ab initio* calculations of charge transport in heavy elements, but also for effective models where integrating out high-energy degrees of freedom can lead to a renormalization of the effective Hamiltonian away from the standard nonrelativistic form.

We can compute the difference  $\mathbf{j}_q - \tilde{\mathbf{j}}_q$  and  $\mathbf{j}_q - \mathbf{j}_{\text{mid},q}$  order by order in  $\mathbf{q}$ . We find that

$$\mathbf{j}_q - \tilde{\mathbf{j}}_q = -\frac{\bar{e}}{12} \sum_i q_\mu q_\nu \frac{\partial^2 \mathbf{v}_i(\mathbf{p}_i, \mathbf{x}_i)}{\partial p_i^\mu \partial p_i^\nu} + \mathcal{O}(|\mathbf{q}|^3), \quad (38)$$

$$\mathbf{j}_q - \mathbf{j}_{\text{mid},q} = \frac{\bar{e}}{24} \sum_i q_\mu q_\nu \frac{\partial^2 \mathbf{v}_i(\mathbf{p}_i, \mathbf{x}_i)}{\partial p_i^\mu \partial p_i^\nu} + \mathcal{O}(|\mathbf{q}|^3). \quad (39)$$

We see from Eqs. (38) and (39) that  $\mathbf{j}_q$ ,  $\tilde{\mathbf{j}}_q$ , and  $\mathbf{j}_{\text{mid},q}$  all coincide to linear order in  $\mathbf{q}$ , with the first discrepancy at the order of  $|\mathbf{q}|^2$ . This suggests that, while calculations of uniform magnetoelectric response functions (which requires knowing  $\mathbf{j}_q$  to linear order in  $\mathbf{q}$ , as in Refs. [41,42,66]) can use  $\tilde{\mathbf{j}}_q$  or  $\mathbf{j}_{\text{mid},q}$  in place of the conserved Eq. (31), the calculations of the  $\mathcal{O}(|\mathbf{q}|^2)$  corrections to the conductivity in Refs. [51,68] should be reexamined. We show below that, while  $\mathbf{j}_q$ ,  $\tilde{\mathbf{j}}_q$ , and  $\mathbf{j}_{\text{mid},q}$  all yield the same Hall conductivity to the order of  $\mathcal{O}(|\mathbf{q}|^2)$ , they yield different predictions for the longitudinal conductivity at the same order. Even further, we see from Eq. (38) that  $\mathbf{q} \cdot (\mathbf{j}_q - \tilde{\mathbf{j}}_q) \neq 0$  and  $\mathbf{q} \cdot (\mathbf{j}_q - \mathbf{j}_{\text{mid},q}) \neq 0$ , so that *only*  $\mathbf{j}_q$  and *not*  $\tilde{\mathbf{j}}_q$  *nor*  $\mathbf{j}_{\text{mid},q}$  satisfies the continuity equation (16). Generically, we have

$$\mathbf{q} \cdot \mathbf{j}_q - \mathbf{q} \cdot \mathbf{j}_{\text{mid},q} = 0, \quad (40)$$

$$\mathbf{q} \cdot \tilde{\mathbf{j}}_q - \mathbf{q} \cdot \mathbf{j}_{\text{mid},q} \neq 0, \quad (41)$$

$$\mathbf{q} \cdot \mathbf{j}_{\text{mid},q} - \mathbf{q} \cdot \mathbf{j}_{\text{min},q} \neq 0. \quad (42)$$

In Appendix B, we give a direct analysis of how the conserved  $\mathbf{j}_q$  differs from nonconserved  $\tilde{\mathbf{j}}_q$  and  $\mathbf{j}_{\text{mid},q}$  for a concrete model of a semirelativistic free-electron system. We show explicitly that, once the free electron is energetic enough for quartic or higher-order corrections to the kinetic energy to become appreciable, the conserved  $\mathbf{j}_q$  and nonconserved  $\tilde{\mathbf{j}}_q$  and  $\mathbf{j}_{\text{mid},q}$  disagree at the order of  $|\mathbf{q}|^2$ .

Although we have shown how the conserved  $\mathbf{j}_q$  of Eq. (31) differs from the nonconserved  $\tilde{\mathbf{j}}_q$  and  $\mathbf{j}_{\text{mid},q}$ , in the current form our results apply to only continuum first-quantized Hamiltonians of the form of Eq. (12). To apply our same approach to formulate the current density operator for a wider variety of condensed matter problems, we in Sec. II B calculate the second-quantized form of the current

operator Eq. (31) by examining its matrix elements in a basis of single-particle Bloch eigenstates.

## B. Matrix elements of the current operator in the Bloch basis with second quantization

In this section, we evaluate the matrix elements of the conserved current operator Eq. (31)—as well as the non-conserved current operators Eqs. (32) and (33)—in a basis of single-particle Bloch eigenstates. Consider a noninteracting Hamiltonian of the form of Eq. (12) with  $U = 0$  with discrete translation symmetry as in Eq. (14). From Bloch's theorem, we can introduce single-particle eigenstates

$$\psi_{n\mathbf{k}}(r) = \frac{1}{\sqrt{N}} e^{i\mathbf{k} \cdot \mathbf{r}} u_{n\mathbf{k}}(r), \quad (43)$$

where  $N$  is the nominally infinite number of unit cells in the system. We denote the eigenstates with kets  $|\psi_{n\mathbf{k}}\rangle$  with inner product

$$\langle \psi_{n\mathbf{k}} | \psi_{m\mathbf{k}'} \rangle = \int d\mathbf{r} \psi_{n\mathbf{k}}^*(\mathbf{r}) \psi_{m\mathbf{k}'}(\mathbf{r}) = \delta_{\mathbf{k},\mathbf{k}'} \delta_{nm}. \quad (44)$$

It is also convenient to introduce kets  $|u_{n\mathbf{k}}\rangle$  to represent the cell-periodic functions  $u_{n\mathbf{k}}(\mathbf{r})$ . The inner product of cell-periodic kets is given by

$$\langle u_{n\mathbf{k}} | u_{m\mathbf{k}'} \rangle = \int_{\text{cell}} d\mathbf{r} u_{n\mathbf{k}}^*(\mathbf{r}) u_{m\mathbf{k}'}(\mathbf{r}), \quad (45)$$

where  $\int_{\text{cell}}$  denotes an integration over a single unit cell. We work in the periodic gauge such that

$$|\psi_{n\mathbf{k}+\mathbf{G}}\rangle = |\psi_{n\mathbf{k}}\rangle \quad (46)$$

for any reciprocal lattice vector  $\mathbf{G}$ . Note that the periodic gauge always exists, even for topologically nontrivial systems, as we discuss further at the end of this subsection. The periodic gauge constraint implies [72]

$$|u_{n\mathbf{k}+\mathbf{G}}\rangle = e^{-i\mathbf{G} \cdot \mathbf{x}} |u_{n\mathbf{k}}\rangle. \quad (47)$$

Let us begin by computing the matrix elements of the density operator  $\rho_q = \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i}$  between two Bloch states. We have

$$\begin{aligned} \langle \psi_{n\mathbf{k}} | \rho_q | \psi_{m\mathbf{k}'} \rangle &= \bar{e} \int d\mathbf{r} \psi_{n\mathbf{k}}^*(\mathbf{r}) e^{-i\mathbf{q} \cdot \mathbf{r}} \psi_{m\mathbf{k}'}(\mathbf{r}) \\ &= \bar{e} \frac{1}{N} \sum_{\mathbf{R}} e^{i\mathbf{R} \cdot (\mathbf{k}' - \mathbf{k} - \mathbf{q})} \\ &\quad \times \int_{\text{cell}} d\mathbf{r} u_{n\mathbf{k}}^*(\mathbf{r}) e^{i\mathbf{r} \cdot (\mathbf{k}' - \mathbf{k} - \mathbf{q})} u_{m\mathbf{k}'}(\mathbf{r}) \\ &= \delta_{\mathbf{k}', \mathbf{k} + \mathbf{q}} \bar{e} \langle u_{n\mathbf{k}} | u_{m\mathbf{k} + \mathbf{q}} \rangle. \end{aligned} \quad (48)$$

Introducing creation and annihilation operators  $c_{n\mathbf{k}}^\dagger$  and  $c_{n\mathbf{k}}$  that create or annihilate electrons in the state  $|\psi_{n\mathbf{k}}\rangle$ , Eq. (48) implies that we can write the density operator in second-quantized notation as

$$\begin{aligned}\rho_{\mathbf{q}} &= \sum_{\mathbf{k}\mathbf{k}'nm} \langle \psi_{n\mathbf{k}} | \rho_{\mathbf{q}} | \psi_{m\mathbf{k}'} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k}'} \\ &= \bar{e} \sum_{\mathbf{k}nm} \langle u_{n\mathbf{k}} | u_{m\mathbf{k}+\mathbf{q}} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k}+\mathbf{q}}.\end{aligned}\quad (49)$$

Next, we can examine the second-quantized form of our conserved current  $\mathbf{j}_{\mathbf{q}}$  from Eq. (31). Note first that

$$\begin{aligned}T_{\mathbf{R}} \mathbf{j}_{\mathbf{q}} |\psi_{n\mathbf{k}}\rangle &= \bar{e} T_{\mathbf{R}} \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} \mathbf{v} e^{-i\lambda\mathbf{q}\cdot\mathbf{x}} |\psi_{n\mathbf{k}}\rangle \\ &= \bar{e} \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot(\mathbf{x}-\mathbf{R})} \mathbf{v} e^{-i\lambda\mathbf{q}\cdot(\mathbf{x}-\mathbf{R})} T_{\mathbf{R}} |\psi_{n\mathbf{k}}\rangle \\ &= e^{-i(\mathbf{k}-\mathbf{q})\cdot\mathbf{R}} \mathbf{j}_{\mathbf{q}} |\psi_{n\mathbf{k}}\rangle.\end{aligned}\quad (50)$$

Equation (50) indicates that  $\mathbf{j}_{\mathbf{q}} |\psi_{n\mathbf{k}}\rangle$  has crystal momentum  $\mathbf{k} - \mathbf{q}$ . Since Bloch states with different crystal momentum are orthogonal, we deduce that the only nonvanishing single-particle matrix elements of  $\mathbf{j}_{\mathbf{q}}$  are

$$\begin{aligned}\langle \psi_{n\mathbf{k}} | \mathbf{j}_{\mathbf{q}} | \psi_{m\mathbf{k}+\mathbf{q}} \rangle &= \frac{\bar{e}}{N} \int_0^1 d\lambda \int d\mathbf{r} u_{n\mathbf{k}}^*(\mathbf{r}) e^{-i[(1-\lambda)\mathbf{q}+\mathbf{k}]\cdot\mathbf{r}} \\ &\quad \times \mathbf{v} e^{i[\mathbf{k}+(1-\lambda)\mathbf{q}]\cdot\mathbf{r}} u_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}) \\ &= \bar{e} \int_0^1 d\lambda \int_{\text{cell}} d\mathbf{r} u_{n\mathbf{k}}^*(\mathbf{r}) \mathbf{v}_{\mathbf{k}+(1-\lambda)\mathbf{q}} u_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}) \\ &= \bar{e} \int_0^1 d\lambda \langle u_{n\mathbf{k}} | \mathbf{v}_{\mathbf{k}+(1-\lambda)\mathbf{q}} | u_{m\mathbf{k}+\mathbf{q}} \rangle,\end{aligned}\quad (51)$$

where we have defined

$$\mathbf{v}_{\mathbf{k}} = e^{-i\mathbf{k}\cdot\mathbf{x}} \mathbf{v} e^{i\mathbf{k}\cdot\mathbf{x}} = \partial_{\mathbf{k}} (e^{-i\mathbf{k}\cdot\mathbf{x}} H e^{i\mathbf{k}\cdot\mathbf{x}}) \equiv \partial_{\mathbf{k}} H_{\mathbf{k}}. \quad (52)$$

Using Eq. (51), we conclude that the current  $\mathbf{j}_{\mathbf{q}}$  has the second-quantized representation

$$\mathbf{j}_{\mathbf{q}} = \bar{e} \sum_{\mathbf{k}nm} \int_0^1 d\lambda \langle u_{n\mathbf{k}} | \mathbf{v}_{\mathbf{k}+(1-\lambda)\mathbf{q}} | u_{m\mathbf{k}+\mathbf{q}} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k}+\mathbf{q}}. \quad (53)$$

By the same logic, we find the second-quantized representation of the nonconserved trapezoid current  $\tilde{\mathbf{j}}_{\mathbf{q}}$  from Eq. (32) is

$$\tilde{\mathbf{j}}_{\mathbf{q}} = \bar{e} \sum_{\mathbf{k}nm} \frac{1}{2} \langle u_{n\mathbf{k}} | \mathbf{v}_{\mathbf{k}+\mathbf{q}} + \mathbf{v}_{\mathbf{k}} | u_{m\mathbf{k}+\mathbf{q}} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k}+\mathbf{q}}. \quad (54)$$

Finally, for the nonconserved midpoint current  $\mathbf{j}_{\text{mid},\mathbf{q}}$  in Eq. (33), we find the second-quantized representation

$$\mathbf{j}_{\text{mid},\mathbf{q}} = \bar{e} \sum_{\mathbf{k}nm} \langle u_{n\mathbf{k}} | \mathbf{v}_{\mathbf{k}+\mathbf{q}/2} | u_{m\mathbf{k}+\mathbf{q}} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k}+\mathbf{q}}. \quad (55)$$

Equation (53) is the main result of this section and allows us to reformulate our conserved current operator Eq. (31) in terms of second-quantized Bloch orbitals. Equation (53) can be used to compute the matrix elements of the conserved current for effective models that include only a subset of the energy bands of interest in a solid, in which the effective kinetic energy need not be a quadratic function of momentum. Additionally, Eq. (53) manifestly satisfies the continuity equation and so gives a proper starting point for the evaluation of linear and nonlinear electromagnetic response coefficients.

Although we supposed that the wave functions  $|u_{n\mathbf{k}}\rangle$  were the (cell periodic parts of the) energy eigenstates for the noninteracting part of a Hamiltonian, we can follow the logic leading to Eq. (53) for any orthonormal Bloch-like basis set. In Appendix C, for example, we derive the second-quantized representation of the current Eq. (31) in a basis of ultralocalized tight-binding orbitals which are useful for calculations in model systems. Additionally, provided that the interaction energy depends only on (unprojected) density operators, Eq. (53) still gives the second-quantized current operator.

Note that, for systems with nonvanishing Chern numbers, the periodic gauge constraint for  $|u_{n\mathbf{k}}\rangle$  in Eq. (46) requires that the phase of  $|u_{n\mathbf{k}}\rangle$  be nonsmooth in the Brillouin zone [73,74]. This presents no difficulty for our formalism, which does not involve derivatives or Taylor expansions of the wave functions themselves with respect to  $\mathbf{k}$  but only derivatives of the manifestly smooth operator  $H_{\mathbf{k}}$ . Additionally, when we work in a basis of tight-binding orbitals as in Appendix C, we can include a sufficiently large number of bands to guarantee that the basis Wannier orbitals have smooth Fourier transforms. Finally, we note that, in any numerical computation, the wave functions  $|u\rangle_{n\mathbf{k}}$  are typically evaluated on a discrete grid of points in the Brillouin zone; we can then choose a periodic gauge such that any nonremovable singularities due to nonvanishing Chern number occur at points that are not included in the grid.

Having expressed the current operator Eq. (31) in the second-quantized Bloch basis, we can now turn our attention to how the current density is modified in the presence of a background electromagnetic field. This allows us to generalize the usual diamagnetic current to nonzero wave vector for systems with generic Hamiltonians. We see that these generalized diamagnetic current operators—which are essential to maintain gauge invariance of response functions—appear naturally in our formalism for  $\mathbf{j}_{\mathbf{q}}$  defined in Eq. (31).

### C. Diamagnetic current operators

In order to compute electromagnetic response functions, we need to know the current  $\mathbf{j}_{\mathbf{A}}(\mathbf{r})$  order by order in the



vector potential  $\mathbf{A}$ . Although we primarily focus on  $\mathbf{j}_{\mathbf{q},\min}$  in Eq. (22), for a general system, we can write

$$j_A^\mu(\mathbf{r}) = j_{\min}^\mu(\mathbf{r}) + \int d\mathbf{r}' j^{\mu\nu}(\mathbf{r}, \mathbf{r}') A_\nu(\mathbf{r}', t) + \dots \quad (56)$$

or in Fourier space

$$j_{A,\mathbf{q}}^\mu = j_{\min,\mathbf{q}}^\mu + \sum_{\mathbf{q}'} j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu} A_{\nu,\mathbf{q}'} + \dots \quad (57)$$

While knowledge of  $j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu}$  requires knowledge of the minimally coupled Hamiltonian in Eq. (17), the continuity equation at Eq. (16) places constraints on the longitudinal components of  $j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu}$ . In particular, note first that

$$\begin{aligned} j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu} &= -\frac{\delta^2 H_A}{\delta A_{\mu,-\mathbf{q}} \delta A_{\nu,\mathbf{q}'}} \Big|_{\mathbf{A}=0} \\ &= \frac{\delta j_{\mathbf{q},\mathbf{A}}^\mu}{\delta A_{\nu,\mathbf{q}'}} \Big|_{\mathbf{A}=0}. \end{aligned} \quad (58)$$

Inserting Eq. (58) into the continuity equation (16), we find that

$$\begin{aligned} i q_\mu j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu} &= \frac{\delta}{\delta A_{\nu,\mathbf{q}'}} (i q_\mu j_{\mathbf{q},\mathbf{A}}^\mu) \Big|_{\mathbf{A}=0} \\ &= -i \frac{\delta}{\delta A_{\nu,\mathbf{q}'}} [H_A, \rho_{\mathbf{q}}] \Big|_{\mathbf{A}=0} \\ &= i [j_{-\mathbf{q}',\min}^\nu, \rho_{\mathbf{q}}], \end{aligned} \quad (59)$$

where we have used from Eq. (23) that  $\rho_{\mathbf{q}}$  is independent of the vector potential. Equation (59) shows that the longitudinal component of the diamagnetic current  $j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu}$  can be expressed entirely in terms of the unperturbed current  $j_{-\mathbf{q}',\min}^\nu$ . Furthermore, since the longitudinal component of  $j_{-\mathbf{q}',\min}^\nu$  is equal to the current in Eq. (31), we find that the longitudinal component  $q_\mu q'_\nu j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu}$  of the diamagnetic

current is entirely expressible in terms of the velocity operator.

Equation (59) is equivalent to the Ward identity from quantum field theory [75]. We can iterate the logic leading to Eq. (59) in order to obtain the longitudinal part of all higher-order current vertices  $q_{\mu_1} q'_{\mu_2} q''_{\mu_3} \dots j_{\mathbf{q},\mathbf{q}',\mathbf{q}''}^{\mu_1\mu_2\mu_3\dots}$ . Since the operators  $j_{\mathbf{q},\mathbf{q}',\mathbf{q}''}^{\mu_1\mu_2\mu_3\dots}$  appear as vertex functions in calculations of higher-order electromagnetic responses [29,76], we see that our expression for the current in Eq. (31) in terms of the velocity operator can be used to generate the longitudinal components of these higher-order vertices.

Using Eqs. (58) and (59), we can recast this second-order operator in terms of an integration over an auxiliary variable  $\lambda$  via the Karplus-Schwinger relation. Namely, we show in Appendix D that, for any single-particle operator  $O$ ,

$$i[O, e^{-i\mathbf{q}\cdot\mathbf{x}}] = \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} (\mathbf{q} \cdot [O, \mathbf{x}]) e^{-i\lambda\mathbf{q}\cdot\mathbf{x}}. \quad (60)$$

Applying Eq. (60) to our Ward identity Eq. (59), we find

$$\begin{aligned} i q_\mu q'_\nu j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu} &= i \left[ q'_\nu j_{-\mathbf{q}',\min}^\nu, \bar{e} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \right] \\ &= \bar{e} \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} (q_\mu q'_\nu [j_{-\mathbf{q}',\min}^\nu, x^\mu]) e^{-i\lambda\mathbf{q}\cdot\mathbf{x}} \\ &= \bar{e} \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} (q_\mu q'_\nu [j_{-\mathbf{q}',\min}^\nu, x^\mu]) e^{-i\lambda\mathbf{q}\cdot\mathbf{x}}, \end{aligned} \quad (61)$$

where we have used the fact that both  $\mathbf{j}_{\min,\mathbf{q}}$  and  $\mathbf{j}_{\mathbf{q}}$  in Eq. (31) are conserved. The advantage to using the Karplus-Schwinger relation is being able to strip off the  $\mathbf{q}$  and  $\mathbf{q}'$ , thus allowing us to define  $j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu}$  that satisfies the Ward identity for generic systems.

We can rewrite Eq. (61) in terms of the second-quantized creation and annihilation operators for Bloch eigenstates following the formalism of Sec. II B in order to define

$$j_{\mathbf{q},-\mathbf{q}'}^{\mu\nu} = -(\bar{e})^2 \int_0^1 \int_0^1 d\lambda d\lambda' \langle u_{n\mathbf{k}} | \left[ \frac{\partial^2 H_{\mathbf{k}}}{\partial k^\mu \partial k^\nu} \right]_{\mathbf{k} \rightarrow \mathbf{k} - (1-\lambda')\mathbf{q}' + (1-\lambda)\mathbf{q}} | u_{m\mathbf{k}-\mathbf{q}'+\mathbf{q}} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k}-\mathbf{q}'+\mathbf{q}}. \quad (62)$$

In a similar manner, we can iterate this procedure to derive expressions for the  $N$ th-order current operator

$$j_{\mathbf{q},-\mathbf{q}_1,-\mathbf{q}_2,\dots,-\mathbf{q}_N}^{\mu\nu_1\nu_2\dots\nu_N} = -\frac{\delta^{N+1} H_A}{\delta A_{\mu,-\mathbf{q}} \delta A_{\nu_1,\mathbf{q}_1} \delta A_{\nu_2,\mathbf{q}_2} \dots \delta A_{\nu_N,\mathbf{q}_N}} \Big|_{\mathbf{A}=0}, \quad (63)$$

which satisfies the generalized Ward identity

$$(i q_\mu) (-i q_{1,\nu_1}) (-i q_{2,\nu_2}) \dots (-i q_{N-1,\nu_{N-1}}) j_{\mathbf{q},-\mathbf{q}_1,-\mathbf{q}_2,\dots,-\mathbf{q}_N}^{\mu\nu_1\nu_2\dots\nu_N} = i \left[ i \left[ j_{-\mathbf{q}_N,\min}^{\nu_N}, \rho_{-\mathbf{q}_{N-1}} \right], \rho_{-\mathbf{q}_{N-2}} \right], \dots, \rho_{\mathbf{q}}. \quad (64)$$

Applying the Karplus-Schwinger relation Eq. (60) iteratively and using our expressions Eqs. (31) and (53), we find that we can write the  $n$ th-order diamagnetic current operator in second-quantized notation as

$$j_{\mathbf{q}, -\mathbf{q}_1, \dots, -\mathbf{q}_N}^{\mu\nu_1 \dots \nu_N} = -(-\bar{e})^{N+1} \int_0^1 d\lambda \int_0^1 d\lambda_1 \dots \int_0^1 d\lambda_N \times \langle u_{n\mathbf{k}} | [\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^{\nu_1}} \dots \partial_{\mathbf{k}^{\nu_N}} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + (1-\lambda)\mathbf{q} - (1-\lambda_1)\mathbf{q}_1 \dots - (1-\lambda_N)\mathbf{q}_N} | u_{m\mathbf{k} + \mathbf{q} - \mathbf{q}_1 \dots - \mathbf{q}_N} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k} + \mathbf{q} - \mathbf{q}_1 \dots - \mathbf{q}_N}. \quad (65)$$

Note, importantly, that Eq. (65) is explicitly symmetric under the exchange of any pair of indices, since partial derivatives commute.

For later convenience, we define the  $N$ th-order velocity vertex  $v_{(N)}^{\nu_N}(\mathbf{k}, \mathbf{q}, -\mathbf{q}_1, \dots, -\mathbf{q}_N)$  as the operator appearing in the matrix element of Eq. (65), i.e.,

$$v_{(N)}^{\nu_N}(\mathbf{k}, \mathbf{q}, -\mathbf{q}_1, \dots, -\mathbf{q}_N) \equiv \int_0^1 d\lambda \int_0^1 d\lambda_1 \dots \int_0^1 d\lambda_N [\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^{\nu_1}} \dots \partial_{\mathbf{k}^{\nu_N}} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + (1-\lambda)\mathbf{q} - (1-\lambda_1)\mathbf{q}_1 \dots - (1-\lambda_N)\mathbf{q}_N}. \quad (66)$$

Therefore, applying Eq. (66) to Eq. (65), we can simply write

$$j_{\mathbf{q}, -\mathbf{q}_1, \dots, -\mathbf{q}_N}^{\mu\nu_1 \dots \nu_N} = -(-\bar{e})^{N+1} \sum_{\mathbf{k}} \langle u_{n\mathbf{k}} | v_{(N)}^{\nu_N}(\mathbf{k}, \mathbf{q}, -\mathbf{q}_1, \dots, -\mathbf{q}_N) | u_{m\mathbf{k} + \mathbf{q} - \mathbf{q}_1 \dots - \mathbf{q}_N} \rangle c_{n\mathbf{k}}^\dagger c_{m\mathbf{k} + \mathbf{q} - \mathbf{q}_1 \dots - \mathbf{q}_N}. \quad (67)$$

Note that, since the trapezoid and midpoint currents [Eqs. (10) and (11)] do not satisfy the continuity equation, there is no general procedure for determining their corresponding diamagnetic currents. Nevertheless, for the sake of comparison, we define trapezoid and midpoint diamagnetic current operators by approximating each auxiliary integral in Eq. (65) with the trapezoid or midpoint approximation, respectively.

Given Eq. (65) for the generalized diamagnetic current operators as a function of wave vector(s), we now develop a diagrammatic formalism to compute linear and nonlinear conductivities as a function of frequency and wave vector. Our construction of Eqs. (65) and (53) from generalized Ward identities ensures that these conductivities give a current that respects the continuity equation.

### III. FEYNMAN DIAGRAMMATIC METHOD AND SUM RULES

Using the conserved current operator, Eq. (31) we now develop the formalism for computing the spatially inhomogeneous linear and nonlinear conductivities defined in Eq. (2). We build off the work of Ref. [29], which presented a simple yet powerful framework for calculating the spatially uniform ( $\mathbf{q} = 0$ ) nonlinear conductivities. Continuing off of this framework, we look to expand this framework to accommodate spatially inhomogeneous fields and currents. This framework utilizes diagrammatic perturbation theory in terms of noninteracting Matsubara Green's functions for electrons, with interactions with the external electromagnetic field given by the (noninteracting) vertex functions governed by Eq. (66). We generalize this diagrammatic method in Refs. [29,33] to allow for electromagnetic fields and currents with nonzero wave vectors.

As a consequence, we are able to capture not just response to electric fields, but—via Faraday's law Eq. (4)—response to magnetic fields as well. We begin in Sec. III A by establishing the foundation for diagrammatic evaluation of the Kubo formula. Then, in Sec. III B, we give the rules for setting up and evaluating our Feynman diagrams. Finally, in Secs. IV–VI, we apply our formalism to analyze the linear and second-order response as a function of wave vector for 3D Weyl and 2D moiré systems.

#### A. Feynman diagram setup

To write the Feynman diagram rules for evaluating the average current, we must first relate the conductivities to the generating function for correlations. Since the Feynman diagrams are a direct interpretation of perturbative expansions of the generating function, this allows us to write the conductivities as a sum of diagrams, which may then be translated to a mathematical statement involving Green's functions and interaction vertices [29,33,77–83].

First, consider the generic Hamiltonian that has been coupled to a vector potential, as posed in Eq. (17). Notice that the minimally coupled Hamiltonian can be expanded in powers of the vector potential as

$$H_A = H_0 - \left( A_{\alpha_1, -\mathbf{q}_1} j_{\mathbf{q}_1}^{\alpha_1} + \frac{1}{2} A_{\alpha_1, -\mathbf{q}_1} A_{\alpha_2, \mathbf{q}_2} j_{\mathbf{q}_1, -\mathbf{q}_2}^{\alpha_1 \alpha_2} + \frac{1}{6} A_{\alpha_1, -\mathbf{q}_1} A_{\alpha_2, \mathbf{q}_2} A_{\alpha_3, \mathbf{q}_3} j_{\mathbf{q}_1, -\mathbf{q}_2, -\mathbf{q}_3}^{\alpha_1 \alpha_2 \alpha_3} + \dots \right). \quad (68)$$

In the last term in Eq. (68), we have replaced the functional derivatives of the coupled Hamiltonian with the generalized current term described in Eq. (63). We now specify to studying systems of electrons, where we take  $\bar{e} = -e$ , with  $e$  being the (positive) elementary charge. With this

definition of the minimally coupled Hamiltonian, we can define the expectation values of the current operator to all orders in the vector potential [29]:

$$\langle j^\mu \rangle = \frac{1}{Z} \text{Tr} \left[ T_t e^{j_A^\mu(\mathbf{r})} e^{-i \int dt' H_A(t', \mathbf{r})} \right], \quad (69)$$

where  $j_A^\mu(\mathbf{r})$  is defined via Eq. (56) in terms of our conserved current [Eq. (53)] and the diamagnetic currents [Eq. (65)]. Also,  $T_t$  is the time-ordering operator.

In Eq. (69), two levels of expansion may be implemented in orders of the vector potential: an expansion in  $j_A^\mu(\mathbf{r})$  and an expansion in  $H_A$ . We assume the vector potential can be written as a superposition of plane waves:

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\omega, \mathbf{q}} \mathbf{A}_{\mathbf{q}}(\omega) e^{i\mathbf{r} \cdot \mathbf{q} - i\omega t}. \quad (70)$$

To obtain the average current, we can expand the left-hand side of Eq. (69) in terms of a Fourier-transformed response coefficient and the corresponding fields, as described in Eq. (2), taking the Fourier transform of every time and position argument.

Note that, for systems of electrons in a periodic potential, umklapp processes allow for the generation of currents that oscillate in space with wave vectors that differ from the wave vectors of the applied field by a reciprocal lattice vector. Concretely, expanding Eq. (69) in Fourier space [or, equivalently, taking the Fourier transform of Eq. (2)] yields

$$\langle j_{\mathbf{q}}^\mu(t) \rangle = \sum_n \sum_{\mathbf{G}} \int \prod d\mathbf{q}_i dt_i \sigma_{(n)}^{\mu, \nu_1, \dots, \nu_n}(\mathbf{q}, \mathbf{q}_1, \dots, \mathbf{q}_n, t, t_1, \dots, t_n) \delta\left(\mathbf{q} - \sum_i \mathbf{q}_i - \mathbf{G}\right) \prod_i E_{\nu_i, \mathbf{q}_i}(t_i), \quad (71)$$

where  $\mathbf{G}$  are the reciprocal lattice vectors. Since the spatial scale of variation of a current with wave vector  $\mathbf{G} \neq 0$  is on the order of (fractions of) an angstrom, these umklapp terms are typically not experimentally interesting. Therefore, we focus the remaining work on computing the  $\mathbf{G} = 0$  component of the electromagnetic response. We emphasize, however, that our diagrammatic calculation can be easily generalized to compute the  $\mathbf{G} \neq 0$  umklapp components of the current as well.

By comparing Eq. (69) with Eq. (71), we may extract the  $n$ th-order conductivity by using the following process: We first expand the average current to  $n$ th order in the vector potential. Next, we Fourier transform with respect to time and make use of the definition of the electric field

$$\mathbf{E}_{\mathbf{q}}(\omega) = i\omega \mathbf{A}_{\mathbf{q}}(\omega) \quad (72)$$

in the  $A_0 = 0$  gauge. Note that this produces a time-ordered, rather than a causal, response function. To recover the usual causal response function that can be measured in experiment, we can analytically continue all frequencies into the complex plane,  $\omega \rightarrow \omega + i\eta$  for a small infinitesimal  $\eta$ , following the procedure outlined in Refs. [29,65,80]. In addition to preserving causality, a small positive finite  $\eta$  can be used to approximate the electron self-energy due to impurity scattering [29,65].

Practically speaking, we carry out the perturbative calculation using the imaginary-time Matsubara formalism, using Matsubara frequencies  $i\nu_n = [(2n+1)i\pi/\beta]$  for fermions and  $i\omega_n = (2n\pi/\beta)$  for bosons, where  $\beta = (k_B T)^{-1}$  [78]. We take the  $T \rightarrow 0$  limit at the end of all computations, which has the effect of turning Fermi-Dirac distributions  $n_F$  into Heaviside step functions. Moreover, the  $T \rightarrow 0$  limit still gives a good approximation for low-temperature transport coefficients in insulators. Since

the Matsubara frequencies become continuous in this limit, we suppress the subscript  $n$  [29,78,81,82].

With this as our starting point, we can now outline the diagrammatic rules that allow us to express Eq. (71) in terms of Matsubara Green's functions and the vertex functions of Eq. (65).

## B. List of Feynman diagram rules

In this section, we describe the rules for writing down and evaluating the Feynman diagrams for the wave-vector- and frequency-dependent conductivity tensors. These diagrams rely on the generalized current operator derived in Sec. II C. Our formalism closely follows Ref. [29], contrasting only in the inclusion of the wave vector dependence of the external fields and the current operator.

First, let us define the symbology that composes our diagrams.

- (1) The free fermion propagator (Matsubara Green's function)  $G(k)$  is denoted by  $\longrightarrow$ .
- (2) The perturbing vector potential field is denoted by  $\sim\sim\sim\sim$  and carries with it a (Matsubara) frequency and a wave vector.
- (3) Where the fermion and photon propagator meet, we place a vertex in the  $\alpha$  direction. When corresponding to the output current operator it is denoted by an outgoing photon vertex  $\sim\sim$ , and when corresponding to a perturbing field it is denoted by an incoming vertex  $\sim\sim\bullet$ .

We introduce a four-vector notation for Matsubara frequencies, with  $k = (i\nu, \mathbf{k})$  representing the fermionic Matsubara frequencies and momenta and  $q = (i\omega, \mathbf{q})$  representing the bosonic Matsubara frequencies and momenta. We also use the shorthand  $dk \equiv d\mathbf{k} d\nu_1$  to denote the integration measure and  $q_1 + q_2 \equiv q_{12}$  to denote the componentwise sum of two four-vectors.

With these components defined, we may write diagrams describing the contributions to  $\langle j_{\mathbf{q}}^{\mu} \rangle$  order by order in perturbation theory. For concreteness, we specialize at this point to a free-electron system. As such, the Feynman rules below require us to consider only diagrams with a single fermion loop. We note, however, that our diagrammatic rules can be extended in the presence of interactions, such as from phonons [80], by including additional interaction vertices. To simplify the bookkeeping arising from the negative signs in Eq. (65) and from the negative sign of the electron charge, we formulate our diagrammatic rules in terms of the velocity vertex functions in Eq. (66).

The diagrammatic rules are as follows.

- (1) Every loop has an output vertex, and most have input vertices. Diagrams contributing the  $n$ th-order response have  $n + 1$  photon lines. All photon lines are incoming at input vertices. At output vertices, exactly one photon line is outgoing.
- (2) Every loop denotes an integral over both  $\mathbf{k}$  space as well as a Matsubara sum (which can be made into an integral) in frequency  $\nu$ .
- (3) Every closed loop conserves momentum and energy [84]. By convention, incoming photon lines carry momentum out of the loop, and outgoing photon lines carry momentum into the loop.
- (4) Each incoming field vector also contributes a factor of  $ei/\hbar i\omega_{\alpha}$ , where  $\omega_{\alpha}$  is the Matsubara frequency for the  $\alpha$ th photon line.
- (5) To avoid double counting, only topologically unique diagrams should be considered. In particular, if the exchange of two or more four-vector and index labels on photon lines (not including the photon corresponding to the output current) does not change a diagram, then the diagram should be divided by the appropriate multiplicity factor [29]. In practice, this means that  $N$ th-order input diamagnetic current vertices are accompanied by a factor of  $1/N!$ , and  $N$ th-order output diamagnetic vertices are accompanied by a factor of  $1/(N-1)!$ . We see an application of this rule in Sec. VI.
- (6) A vertex with  $N$  photon lines corresponds a factor of (a matrix elements of)  $v_{(N)}^{\nu_N}(\mathbf{k}, \mathbf{q}_1, \dots, \mathbf{q}_N)$  defined in Eq. (65), where  $\mathbf{k}$  is the fermion momenta going into the vertex.

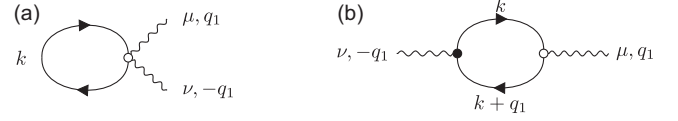


FIG. 1. Linear-order response diagrams.

After evaluating a particular set of diagrams, we can finally obtain a causal response function by analytically continuing each Matsubara frequency back to a real frequency,  $i\omega \rightarrow \omega + i\eta$ , for  $\eta$  a positive infinitesimal.

#### IV. LINEAR RESPONSE

In this section, we apply the Feynman diagrammatic rules from Sec. III B to calculate the linear conductivity as a function of frequency and wave vector for a general noninteracting periodic solid. First, in Sec. IV A, we use our diagrammatic method to recover the Kubo formula for the conductivity, where we comment on the relationship between our approach and that in Refs. [51,66,68]. Then, in Sec. IV B, we show how our formalism relates to the generalized  $f$ -sum rule [4,85].

##### A. Full linear response from diagrams

The linear conductivity is given by the sum of the two diagrams in Fig. 1. Using our Feynman rules, this becomes

$$\begin{aligned} \sigma^{\mu\nu}(i\omega, \mathbf{q}) = & \frac{ie^2}{i\omega} \int dk \left[ G_{n_1}(k) v_{(2),n_1 n_1}^{\mu\nu}(\mathbf{k}, -\mathbf{q}, \mathbf{q}) \right. \\ & + G_{n_2}(k + \mathbf{q}) v_{(1)n_2 n_1}^{\nu}(\mathbf{k} + \mathbf{q}, -\mathbf{q}) \\ & \left. \times G_{n_1}(k) v_{(1)n_1 n_2}^{\mu}(\mathbf{k}, \mathbf{q}) \right], \end{aligned} \quad (73)$$

where  $G_n(k) = (\nu - \epsilon_{n\mathbf{k}})^{-1}$  is the Matsubara Green's function in the energy eigenbasis  $|\psi_{n\mathbf{k}}\rangle$  of Eq. (43),  $\epsilon_{n\mathbf{k}}$  is the corresponding energy, and we introduce the notation  $v_{(N+1),n_1 n_2}^{\mu\nu_1 \dots \nu_N}$  to denote the matrix elements of the velocity vertex Eq. (66) in the basis of Bloch eigenstates  $|u_{n\mathbf{k}}\rangle$ .

The first term in Eq. (73), corresponding to the diagram in Fig. 1(a), is the diamagnetic conductivity. We can evaluate Eq. (73) and analytically continue to real frequency to find

$$\begin{aligned} \sigma_{\text{dia}}^{\mu\nu}(\omega) = & \frac{ie^2}{\omega^+} \int dk \sum_{n_1} G_{n_1}(k) v_{(2),n_1 n_1}^{\mu\nu}(\mathbf{k}, -\mathbf{q}, \mathbf{q}) \\ = & \frac{ie^2}{\omega^+} \int d\nu d\mathbf{k} d\lambda d\lambda_1 \sum_{n_1} \frac{\langle u_{n_1 \mathbf{k}} | (\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^\nu} H_{\mathbf{k}}) |_{\mathbf{k} \rightarrow \mathbf{k} - (1-\lambda_1)\mathbf{q} + (1-\lambda)\mathbf{q}} | u_{n_1 \mathbf{k}} \rangle}{i\nu - \epsilon_{n_1 \mathbf{k}}} \\ = & \frac{ie^2}{\omega^+} \int d\mathbf{k} d\lambda d\lambda_1 \sum_{n_1} n_F(\epsilon_{n_1}(\mathbf{k})) \langle u_{n_1 \mathbf{k}} | (\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^\nu} H_{\mathbf{k}}) |_{\mathbf{k} \rightarrow \mathbf{k} + (\lambda_1 - \lambda)\mathbf{q}} | u_{n_1 \mathbf{k}} \rangle, \end{aligned} \quad (74)$$



where  $n_F$  is the Fermi-Dirac distribution and  $\omega^+ = \omega + i\eta$ . In the limit  $\mathbf{q} \rightarrow 0$ , this reproduces the generalized diamagnetic conductivity of Ref. [29]. Here, however, we see that, when  $H_{\mathbf{k}}$  contains nonquadratic momentum dependence—as it will for any semirelativistic system or for any system modeled by a low-energy effective Hamiltonian—the diamagnetic conductivity acquires a nontrivial  $\mathbf{q}$  dependence. The diamagnetic conductivity serves to regularize the conductivity in the limit  $\omega \rightarrow 0$ . We can gain some intuition for this by noting that Fig. 1(a) gives the diamagnetic conductivity in terms of the average

$$\sigma_{\text{dia}}^{\mu\nu}(\omega) = \frac{i}{\omega^+} \langle j_{-\mathbf{q},\mathbf{q}}^{\mu\nu} \rangle \quad (75)$$

of the diamagnetic current Eq. (62). The generalized Ward identity from Eq. (59) constrains the diamagnetic conductivity.

The bubble diagram in Fig. 1(b) gives the paramagnetic conductivity, corresponding to the second term in Eq. (73). By evaluating the Matsubara integrals in Eq. (73) and analytically continuing back to real frequencies, we find that the paramagnetic conductivity is given by

$$\sigma_{\text{para}}^{\alpha\beta}(\omega, \mathbf{q}) = \frac{-ie^2}{\omega^+} Q^{\alpha\beta}(\omega, \mathbf{q}), \quad (76)$$

where  $Q^{\alpha\beta}(t - t', \mathbf{q}) \equiv i \langle [j_{\mathbf{q}}^{\alpha}(t), j_{-\mathbf{q}}^{\beta}(t')] \rangle \theta(t - t')$  is the retarded current-current correlation function in the time domain. In the frequency domain, we can use our expression Eq. (53) for the current operator to find

$$\begin{aligned} Q^{\alpha\beta}(\omega, \mathbf{q}) &= \sum_{\mathbf{k}} v_{ab}^{\alpha}(\mathbf{k}, \mathbf{q}) v_{ba}^{\beta}(\mathbf{k} + \mathbf{q}, -\mathbf{q}) \frac{n_F(\epsilon_{a\mathbf{k}}) - n_F(\epsilon_{b\mathbf{k}+\mathbf{q}})}{\epsilon_{a\mathbf{k}} - \epsilon_{b\mathbf{k}+\mathbf{q}} + \omega^+} \\ &\equiv \sum_{\mathbf{k}} F_{\alpha\beta}^{ab}(\mathbf{k}, \mathbf{q}) \frac{n_F(\epsilon_{a\mathbf{k}}) - n_F(\epsilon_{b\mathbf{k}+\mathbf{q}})}{\epsilon_{a\mathbf{k}} - \epsilon_{b\mathbf{k}+\mathbf{q}} + \omega^+}, \end{aligned} \quad (77)$$

where  $v_{ab}^{\alpha}(\mathbf{k}, \mathbf{q})$  is the charge-conserving velocity, given by

$$v_{ab}^{\alpha}(\mathbf{k}, \mathbf{q}) = \int_0^1 d\lambda \langle u_{a\mathbf{k}} | (\partial_{\mathbf{k}^{\alpha}} H_{\mathbf{k}}) |_{\mathbf{k} \rightarrow \mathbf{k} + (1+\lambda)\mathbf{q}} | u_{b\mathbf{k}+\mathbf{q}} \rangle \quad (78)$$

as indicated in Eq. (65).

Our improved definition for the conserved current operator enters into the matrix elements  $F_{\alpha\beta}^{ab}(\mathbf{k}, \mathbf{q})$ , while the ratio of Fermi functions to energy differences is universal and arises from the evaluation of the Green's function integral in Eq. (73). We can compare our result for the paramagnetic conductivity to that in Refs. [51,68], which uses the midpoint current from Eq. (33) instead of the conserved current. This leads to the appearance of a modified matrix element

$$F_{\text{mid};\alpha\beta}^{ab}(\mathbf{k}, \mathbf{q}) = \left\langle u_{a\mathbf{k}} \left| \hat{v}_{\mathbf{k}+\frac{\mathbf{q}}{2}}^{\alpha} \right| u_{b\mathbf{k}+\mathbf{q}} \right\rangle \left\langle u_{b\mathbf{k}+\mathbf{q}} \left| \hat{v}_{\mathbf{k}+\frac{\mathbf{q}}{2}}^{\beta} \right| u_{a\mathbf{k}} \right\rangle, \quad (79)$$

in place of  $F_{\alpha\beta}^{ab}(\mathbf{k}, \mathbf{q})$ , with  $\hat{v}_{\mathbf{k}}^{\alpha} \equiv (\partial \hat{H}_{\mathbf{k}} / \partial k_{\alpha})$ . The subscript “mid” is assigned to this quantity to remind the reader that it is derived using the midpoint definition of the current operator.

Similarly, if the (nonconserved) trapezoid current  $\tilde{\mathbf{j}}$  from Eq. (32) were used in place of  $\mathbf{j}$  to derive the response, then we would find for the matrix element

$$\begin{aligned} \tilde{F}_{\alpha\beta}^{ab}(\mathbf{k}, \mathbf{q}) &= \left\langle u_{a\mathbf{k}} \left| \frac{\hat{v}_{\mathbf{k}+\mathbf{q}}^{\alpha} + \hat{v}_{\mathbf{k}}^{\alpha}}{2} \right| u_{b\mathbf{k}+\mathbf{q}} \right\rangle \\ &\times \left\langle u_{b\mathbf{k}+\mathbf{q}} \left| \frac{\hat{v}_{\mathbf{k}+\mathbf{q}}^{\beta} + \hat{v}_{\mathbf{k}}^{\beta}}{2} \right| u_{a\mathbf{k}} \right\rangle. \end{aligned} \quad (80)$$

We emphasize that, since Eq. (77) is computed using the conserved current in Eq. (53), it gives the physically meaningful conductivity. We use Eqs. (79) and (80) in Sec. V to show how the trapezoid and midpoint currents give quantitatively different predictions for the conductivity as compared to Eq. (77).

In particular, Ref. [51] uses the (nonconserved) Eq. (79) to evaluate the  $\mathbf{q}$ -dependent Hall response for a two-band tight-binding model, where we expect the use of the conserved current Eq. (53) to be important to obtain reliable results. We now argue that, since the current operators Eqs. (31)–(33) agree to linear order in  $\mathbf{q}$ , they yield identical calculations of the paramagnetic conductivity Eq. (76) to the order of  $|\mathbf{q}|^2$ . To see this, let us rewrite the current-current correlator in Eq. (76) as

$$\begin{aligned} \langle [j_{\mathbf{q}}^{\alpha}(t), j_{-\mathbf{q}}^{\beta}(t')] \rangle &= \langle [j_{\mathbf{0}}^{\alpha}(t), j_{-\mathbf{q}}^{\beta}(t')] \rangle + \langle [j_{\mathbf{q}}^{\alpha}(t), j_{\mathbf{0}}^{\beta}(t')] \rangle \\ &+ \langle [\delta j_{\mathbf{q}}^{\alpha}(t), \delta j_{-\mathbf{q}}^{\beta}(t')] \rangle \\ &- \langle [j_{\mathbf{0}}^{\alpha}(t), j_{\mathbf{0}}^{\beta}(t')] \rangle, \end{aligned} \quad (81)$$

where we have introduced

$$\delta j_{\mathbf{q}}^{\alpha} = j_{\mathbf{q}}^{\alpha} - j_{\mathbf{0}}^{\alpha}. \quad (82)$$

To proceed, note that, for crystalline systems, we know from conservation of (crystal) momentum that the ground state average of any operator with nonzero momentum (modulo reciprocal lattice vectors) must vanish. We can apply this observation to conclude that the first two terms on the right-hand side of Eq. (81) vanish unless  $\mathbf{q} = 0$ . Thus, we have

$$\begin{aligned} \left\langle \left[ j_{\mathbf{q}}^{\alpha}(t), j_{-\mathbf{q}}^{\beta}(t') \right] \right\rangle &= \left\langle \left[ \delta j_{\mathbf{q}}^{\alpha}(t), \delta j_{-\mathbf{q}}^{\beta}(t') \right] \right\rangle \\ &- \left\langle \left[ j_{\mathbf{0}}^{\alpha}(t), j_{\mathbf{0}}^{\beta}(t') \right] \right\rangle. \end{aligned} \quad (83)$$

Finally, Taylor expanding Eq. (83) for small  $\mathbf{q}$  shows that the  $\mathcal{O}(|\mathbf{q}|^2)$  contribution to the paramagnetic conductivity is determined by the  $\mathcal{O}(|\mathbf{q}|)$  term in  $\delta j_{\mathbf{q}}^{\alpha}$ . Since Eqs. (31)–(33) agree to linear order in  $\mathbf{q}$ , they yield the same prediction for the paramagnetic (and, hence, the Hall) conductivity.

Note, however, that the diamagnetic vertex Eq. (75) derived from the conserved current in Eq. (31) differs from the naive  $\mathbf{q}$ -independent diamagnetic current derived from the nonrelativistic Hamiltonian in Eq. (8) and used in Refs. [66–68] to calculate the symmetric part of the conductivity from tight-binding models. In particular, since the conserved two-photon vertex  $j_{-\mathbf{q},\mathbf{q}}^{\mu\nu}$ , defined in Eq. (62), is an even and symmetric function of  $\mathbf{q}$ , there will be  $\mathcal{O}(|\mathbf{q}|^2)$  contributions to the (symmetric) diamagnetic conductivity for tight-binding models that are necessary to ensure charge conservation.

### B. The $f$ -sum rule and diamagnetic conductivity

In this section, we show how our definition for the conserved current  $\mathbf{j}_{\mathbf{q}}$  allows us to derive a generalized  $f$ -sum rule valid for any effective model. While the derivations in this section do not make use of the explicit form of the current operator, they highlight constraints on the conductivity imposed by charge conservation that are satisfied only when the conserved current operator is used. We begin with a short review of density-density response and the derivation of the  $f$ -sum rule [86]. Recall that the  $f$ -sum rule constrains the spectral weight of the density-density response function

$$\chi(\omega, \mathbf{q}) = -i \int_0^{\infty} dt e^{i\omega^+ t} \langle [\rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle. \quad (84)$$

We start by defining the spectral density

$$\chi''(\omega, \mathbf{q}) = \frac{1}{2} \int_{-\infty}^{\infty} dt e^{i\omega t} \langle [\rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \quad (85)$$

which satisfies the Kramers-Kronig relation

$$\chi(\omega, \mathbf{q}) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega' \frac{\chi''(\omega', \mathbf{q})}{\omega^+ - \omega'}. \quad (86)$$

The spectral density  $\chi''(\omega, \mathbf{q})$  can also be expressed as the imaginary part of  $\chi(\omega, \mathbf{q})$  and can be directly measured through absorption spectroscopy.

From Eq. (86), we can deduce an expansion for the large- $\omega$  asymptotics of  $\chi(\omega, \mathbf{q})$  in terms of moments of  $\chi''(\omega, \mathbf{q})$ . By Taylor expanding the denominator in the integral in Eq. (86), we obtain the asymptotic expansion

$$\chi(\omega \rightarrow \infty, \mathbf{q}) \sim \sum_{n=0}^{\infty} \frac{\chi_n''(\mathbf{q})}{(\omega)^{n+1}}, \quad (87)$$

where

$$\chi_n''(\mathbf{q}) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega (\omega)^n \chi''(\omega, \mathbf{q}) \quad (88)$$

is  $(1/\pi)$  times the  $n$ th frequency moment of  $\chi''(\omega, \mathbf{q})$ . Note that, since the (unprojected) density operators commute, i.e.,  $[\rho_{\mathbf{q}}, \rho_{\mathbf{q}'}] = 0$ , we have

$$\chi_0''(\mathbf{q}) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \chi''(\omega, \mathbf{q}) \quad (89)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dt e^{i\omega t} \langle [\rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \quad (90)$$

$$= \int_{-\infty}^{\infty} dt \delta(t) \langle [\rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \quad (91)$$

$$= \langle [\rho_{\mathbf{q}}(0), \rho_{-\mathbf{q}}(0)] \rangle \quad (92)$$

$$= 0, \quad (93)$$

where in going from Eq. (90) to Eq. (91) we exchanged the order of the  $\omega$  and  $t$  integrals and used

$$\frac{1}{2\pi} \int d\omega e^{i\omega t} = \delta(t). \quad (94)$$

This means that the  $n = 0$  term in the asymptotic expansion Eq. (87) vanishes. Using Eqs. (87) and (93), we have, asymptotically to leading order as  $\omega \rightarrow \infty$ ,

$$\lim_{\omega \rightarrow \infty} \omega^2 \chi(\omega, \mathbf{q}) = \chi_1''(\mathbf{q}) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega (\omega) \chi''(\omega, \mathbf{q}). \quad (95)$$

We can go further and evaluate  $\chi_1''(\mathbf{q})$  using the continuity equation to arrive at a general form of the  $f$ -sum rule. In particular, inserting the definition Eq. (85) into the definition Eq. (88) of  $\chi_1''(\mathbf{q})$  and integrating by parts, we find

$$\begin{aligned} \chi_1''(\mathbf{q}) &= \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega (\omega) \chi''(\omega, \mathbf{q}) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dt e^{i\omega t} \omega \langle [\rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \\ &= \frac{-i}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dt \frac{d}{dt} (e^{i\omega t}) \langle [\rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dt e^{i\omega t} \langle [i\partial_t \rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \\ &= \int_{-\infty}^{\infty} dt \delta(t) \langle [i\partial_t \rho_{\mathbf{q}}(t), \rho_{-\mathbf{q}}(0)] \rangle \\ &= \langle [i\partial_t \rho_{\mathbf{q}}, \rho_{-\mathbf{q}}] \rangle, \end{aligned} \quad (96)$$

where we suppress the time arguments in the equal-time commutator in the last line. Using the continuity equation (24), along with the Karplus-Schwinger relationship to simplify commutators of the density operator, we find

$$\begin{aligned}\chi_1''(\mathbf{q}) &= \langle [i\partial_t \rho_{\mathbf{q}}, \rho_{-\mathbf{q}}] \rangle \\ &= q_\mu \langle [j_{\mathbf{q}}^\mu, \rho_{-\mathbf{q}}] \rangle \\ &= q_\mu q_\nu \langle j_{-\mathbf{q},\mathbf{q}}^{\mu\nu} \rangle, \\ &= -i \lim_{\omega \rightarrow 0} \omega q_\mu q_\nu \sigma_{\text{dia}}^{\mu\nu}(\mathbf{q}, \omega),\end{aligned}\quad (97)$$

where we have used Eq. (75) to express the sum rule in terms of the average of the two-photon velocity vertex, which is the coefficient of the diamagnetic conductivity  $\sigma_{\text{dia}}^{\mu\nu}(\mathbf{q}, \omega)$ . For Hamiltonians with quadratic momentum dependence, the diamagnetic conductivity is independent of  $\mathbf{q}$ , and this reproduces the usual  $f$ -sum rule. Note that, since Eq. (96) involves an integral over all frequencies, the sum rule of Eq. (97) in the form we have written it relates to the full (unprojected) density operator and diamagnetic conductivity. We defer the exploration of restricted sum rules over a limited frequency range to future work (though see the recent Ref. [87] for progress along these lines).

Our derivation shows how the  $f$ -sum rule generalizes to nonzero wave vector for general systems. In particular, we show in Sec. II C that the two-photon velocity vertex  $j_{\mathbf{q}',-\mathbf{q}}^{\mu\nu}$ , which uses the conserved current  $\mathbf{j}_{\mathbf{q}}$ , is generally given by Eq. (61). We also show how the correlation-correlation operator through the spectral density function relates back to conductivity in Eq. (97), and, in Appendix F, we explore this connection even further through the plasmon dispersion.

For general Hamiltonians, such as those arising in tight-binding approximations to solid state systems, the  $\mathbf{q}$  dependence of  $j_{\mathbf{q}',-\mathbf{q}}^{\mu\nu}$  leads to a nontrivial  $\mathbf{q}$ -dependent diamagnetic conductivity and, hence, a modification to the  $f$ -sum rule according to Eq. (97). In Ref. [4], the deviation between the generalized  $f$ -sum rule [in the form of Eq. (96)] and the electron filling was taken as a measure of goodness of fit for tight-binding parameters. Here, we take a different point of view: The deviation between the generalized  $f$ -sum rule and the electron filling is a consequence of gauge invariance and quantifies a combination of semirelativistic effects as well as information about the truncation of the Hilbert space in an effective low-energy model. Since the spectral density at nonzero  $\mathbf{q}$  is measurable via absorption spectroscopy, Eq. (97) gives an experimental probe of the two-photon velocity vertex and, hence, of the conserved current  $\mathbf{j}_{\mathbf{q}}$  [85].

## V. APPLICATIONS OF THE LINEAR RESPONSE

Having derived the complete expression for the  $\mathbf{q}$ -dependent linear conductivity in Eq. (73), we now use it to analyze electromagnetic response in insulators and

semimetals. We begin in Sec. VA by computing the frequency- and wave-vector-dependent Hall conductivity in a model of a Weyl semimetal. We see that for large wave vectors the use of the conserved current Eq. (53) in the Kubo formula yields quantitatively different predictions for the Hall response as compared to the trapezoid or midpoint approximations prevalent in the literature. This analysis is applicable to studying the response of the system to spatially modulated ac electromagnetic fields, such as can be applied using standing waves or gate potentials. Next, in Sec. VB, we turn our attention to moiré materials, where the wave vector dependence of  $\sigma^{\mu\nu}(\omega, \mathbf{q})$  has implications even for optical response due to the large effective lattice constant. We compute the Kerr angle and ellipticity at oblique incidence as a function of frequency for a model of a moiré-Chern insulator in two dimensions, showing that spatial inhomogeneous electric fields can lead to experimentally relevant modifications to the Kerr effect. Finally, in Sec. VC, we use our formalism for the conserved current to analyze the magnetic moment and magnetic susceptibility of insulators.

### A. Linear Hall effect in a Weyl semimetal

As a proof of principle, we apply our definition for the conserved current Eq. (53) to compute the wave-vector-dependent Hall conductivity

$$\sigma_H^{xy}(\omega, \mathbf{q}) = \frac{1}{2}(\sigma^{xy}(\omega, \mathbf{q}) - \sigma^{yx}(\omega, \mathbf{q})) \quad (98)$$

for a toy model of a time-reversal symmetry-breaking, inversion symmetric Weyl semimetal with two Weyl points first presented in Ref. [88]. We also compare our predictions with analogous calculations of the Hall conductivity using the (nonconserved) midpoint and trapezoid currents via Eqs. (79) and (80). In the Weyl semimetal, Hall conductivity arises due to the Berry curvature texture of the occupied bands; each Weyl point is a source of Berry curvature of charge  $|C| = 1$ . For two Weyl points of opposite charge separated by momentum  $2k_0\hat{z}$ , the Berry curvature, when the chemical potential is exactly at the two Weyl nodes, leads to a spatially uniform anomalous Hall conductivity [36,89–91]:

$$\sigma_H^{xy}(\omega \rightarrow 0, \mathbf{q} \rightarrow 0) = 2e^2 k_0. \quad (99)$$

For future convenience, we define the wave-vector-dependent anomalous Hall conductivity  $\sigma_{\text{anom}}^{xy}(\mathbf{q})$  as

$$\sigma_H^{xy}(\omega \rightarrow 0, \mathbf{q}) \equiv \sigma_{\text{anom}}^{xy}(\mathbf{q}). \quad (100)$$

We use our formalism to compute how this topological Hall response is modified in the presence of inhomogeneous electric fields. At this stage, we do not assume that our fields are optical—that is, we do not require that

$\omega = c|\mathbf{q}|$ . As such, our analysis is primarily applicable to the response of the Weyl semimetal to spatially dispersive, slowly varying ac electromagnetic fields. Note that the diamagnetic conductivity Eq. (74) is explicitly symmetric, owing to the symmetry of the diamagnetic current vertex in Eq. (59). Therefore, when we go to calculate the Hall conductivity (which is antisymmetric), it will not contribute. Thus, we can focus on the paramagnetic conductivity Eqs. (76) and (77).

### 1. The Weyl semimetal model

We consider a model for a time-reversal symmetry-breaking, inversion symmetric Weyl semimetal with tight-binding Hamiltonian given by [92–97]

$$\begin{aligned}
 H &= \sum_{i=0}^4 d_i(\mathbf{k}) \sigma^i, \\
 d_0(\mathbf{k}) &= \gamma[\cos(k_z) - \cos(k_0)] - \mu, \\
 d_1(\mathbf{k}) &= -2t_x \sin(k_x), \\
 d_2(\mathbf{k}) &= -2t_y \sin(k_y), \\
 d_3(\mathbf{k}) &= -(2t_z(\cos(k_z) - \cos(k_0)) \\
 &\quad + m(2 - \cos(k_x) - \cos(k_y)) \\
 &\quad + \gamma_z(\cos(3k_z) - \cos(3k_0))), \quad (101)
 \end{aligned}$$

where  $\sigma^i$  are the  $2 \times 2$  Pauli matrices with  $\sigma^0$  defined to be the identity matrix. All values of  $\mathbf{k}$  are measured in reduced coordinates, i.e.,  $\mathbf{k}_i \in [-\pi, \pi]$  for  $i = x, y, z$ . Here,  $k_0$  quantifies the separation of the Weyl nodes along the  $\hat{z}$  direction. When  $|m| > 2t_z$ , this model has a gapped spectrum everywhere in the Brillouin zone except at two Weyl points with (reduced) coordinates  $(0, 0, \pm k_0)$ ; we focus on this regime for our analysis. The energy parameters  $t_x, t_y$ , and  $t_z$  determine the velocities close to the Weyl node in each principal direction. The quantities  $\gamma$  and  $\gamma_z$  are tilting parameters that allow for type-II Weyl semimetals to occur [92]. A plot of the spectrum of Eq. (101) with  $k_x = k_y = 0$  is shown in Fig. 2.

Using this model along with our Kubo formula Eq. (77), we examine how the Hall conductivity  $\sigma_H^{xy}(\omega, \mathbf{q})$  deviates from its topological value as a function of both  $\omega$  and  $\mathbf{q}$ .

### 2. Hall response when $\mathbf{q} = (q_0, 0, 0)$

Using Eq. (77), we first compute the Hall response in this Weyl semimetal model when the electric field is parallel to the  $\hat{y}$  direction and the wave vector is oriented in the  $\hat{x}$  direction with a magnitude  $q_0$ . The Hall current then flows along the  $\hat{x}$  direction, parallel to the wave vector. Note that, in this case, the electric field is transverse, while the measured current is longitudinal. We examine this response because it illustrates the differences between the Hall conductivity calculated with the conserved current and

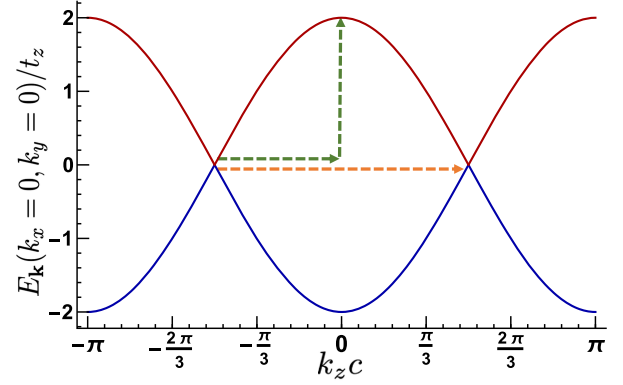


FIG. 2. Energy spectrum of the Weyl semimetal Hamiltonian Eq. (101) along the  $k_x = k_y = 0$  line in the Brillouin zone, with parameter values  $m = 4t_z$ ,  $t_x = t_y = t_z$ ,  $\gamma = 0$ ,  $\gamma_z = 0$ , and  $\mu = 0$ . The dotted lines indicate two types of indirect transitions that can be driven by a time- and space-dependent electric field: The green dotted line shows a jump of  $|\mathbf{q}| = \pi/2$  in momentum and  $\omega = 2t_z$  in energy, and the dotted yellow line shows a jump of  $|\mathbf{q}| = \pi$  in momentum space with zero energy transfer  $\omega = 0$ .

the Hall conductivity calculated with the nonconserved currents commonly used in the literature.

We first illustrate the difference in the three definitions as outlined in Eqs. (77), (79), and (80): the conserved current, the (nonconserved) midpoint definition, and the (nonconserved) trapezoidal definition, respectively.

Each of these three different definitions, yielding different predictions for the Hall conductivity, is shown in Fig. 3. Increasing the wave vector generally leads to a more pronounced difference among the three predictions, which makes sense, since the midpoint and trapezoid currents are approximations to the  $\lambda$  integration in the definition Eq. (53) of the conserved current; these integral approximations become less exact as the wave vector increases. This is further demonstrated in Fig. 4(c), where we can see that the discrepancies between the anomalous Hall conductivity  $\sigma_{\text{anom}}^{xy}(\mathbf{q})$  [defined in Eq. (100)] computed using each of the three methods increase as  $\mathbf{q}$  increases. We note also that, due to the integration over  $\lambda$  in the definition of the conserved current Eq. (53), the anomalous Hall conductivity computed using the conserved current decays as  $1/|\mathbf{q}|^2$  for asymptotically large  $\mathbf{q}$ . Hence, the continuity equation has an influence on the Hall response that becomes more apparent with increasing wave vector. On the contrary, the nonconserved trapezoidal and midpoint definitions of the current fail to capture this feature and instead predict an anomalous Hall conductivity that is periodic in  $q_x$  with period  $2\pi$  and  $4\pi$ , respectively. This point is illustrated in Fig. 4(c).

We also find, as expected, that the three current operators predict the same Hall conductivity in the limit  $\mathbf{q} \rightarrow 0$  for all frequencies. Furthermore, from Fig. 4(c), we see that, for small  $|\mathbf{q}|$ , deviations between the conserved and nonconserved predictions for the anomalous Hall conductivity



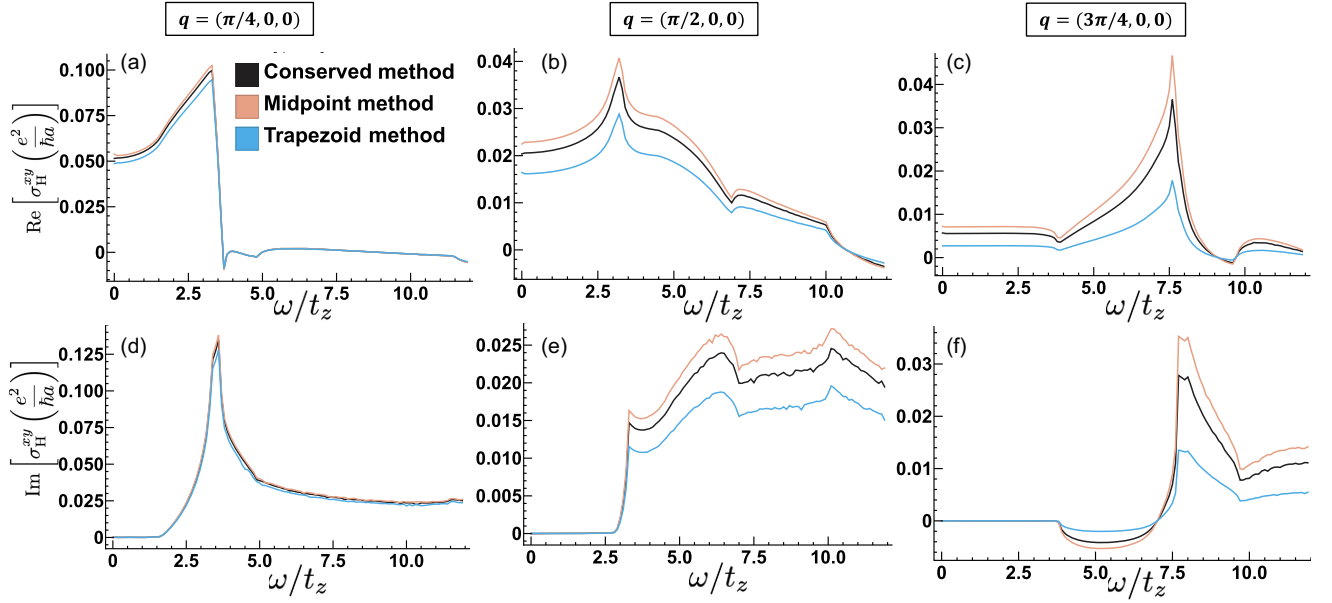


FIG. 3. The Hall conductivity versus the frequency of the perturbing electric field at different values of the wave vector ( $|q| = \pi/4, \pi/2, 3\pi/4$ ) in the  $\hat{x}$  direction. (a)–(c) show the real parts of the Hall conductivity for each wave vector, while (d)–(f) show the imaginary part. The parameters  $m = 4t_z$ ,  $t_x = t_y = t_z$ ,  $\gamma = 0$ ,  $\gamma_z = 0$ , and  $\mu = 0$  are used for these calculations. Black curves correspond to the conductivity computed with the conserved current  $\mathbf{j}_q$  using Eq. (77). For comparison, we also show the Hall conductivity computed with the nonconserved trapezoid current  $\tilde{\mathbf{j}}_q$  (blue) and the nonconserved midpoint current (pink). We see that, at large wave vectors, the nonconserved currents give quantitatively different predictions for the Hall conductivity as compared to the conserved current. All plots are generated with natural units (i.e., all energy parameters in terms of  $t_z$  and  $e = 1$ ).

vanish faster than  $|\mathbf{q}|^2$ . We see that, for small  $\mathbf{q}$ , the anomalous Hall conductivity behaves as

$$\sigma_{\text{anom}}^{xy}(q_x \hat{x}) \sim \sigma_{\text{anom}}^{xy}(0) - |\alpha| q_x^2. \quad (102)$$

This is consistent with the results in Ref. [51], which found that to quadratic order the nonconserved midpoint approximation Eq. (79) predicts a negative  $\mathcal{O}(q^2)$  correction to the anomalous Hall conductivity arising from band-geometric effects (recall that to quadratic order the midpoint current and the conserved current yield the same Hall conductivity).

Focusing now on the physically meaningful conserved current, we show in Figs. 4(a) and 4(b) the real and imaginary parts of the Hall conductivity as a function of frequency for five different wave vectors in the  $\hat{x}$  direction. Strikingly, we see that the Hall conductivity at  $\mathbf{q} = (\pi, 0, 0)$  is identically zero, i.e.,  $\sigma_H^{xy}(\omega, \pi \hat{x}) = 0$ . This is true for computations based on both the conserved and nonconserved currents, arises due to the simplicity of the model, and can be most clearly understood in terms of the trapezoid method: Since  $v_k^x$  depends only on  $k_x$  and includes only nearest-neighbor hopping, we have  $\tilde{F}_{xy}(\mathbf{k}, \pi \hat{x}) = 0$  identically. In a more complicated model with longer-range hopping, we would expect

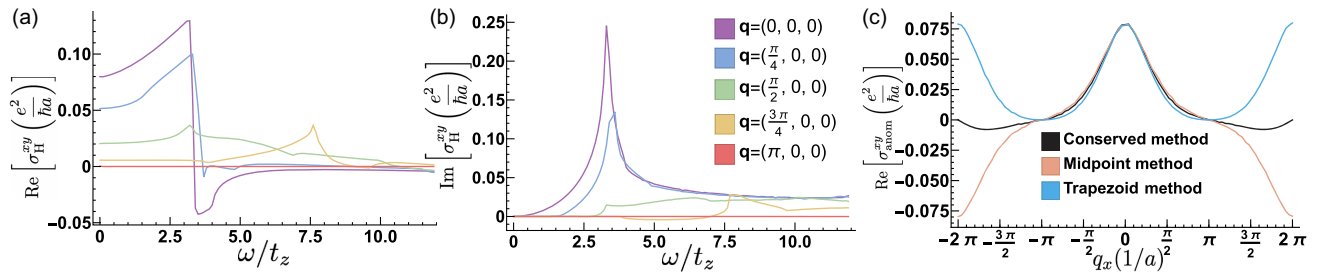


FIG. 4. The real [(a)] and imaginary [(b)] parts of the Hall conductivity as a function of frequency for several values of the wave vector in the  $\hat{x}$  direction. The conductivities in (a) and (b) are computed using the conserved current Eq. (53) and the Kubo formula Eq. (77). (c) shows the real part of the anomalous Hall conductivity [as defined in Eq. (100)] as a function of the wave vector changing in the  $\hat{x}$  direction. The parameters  $m = 4t_z$ ,  $t_x = t_y = t_z$ ,  $\gamma = 0$ ,  $\gamma_z = 0$ , and  $\mu = 0$  are used for these calculations.

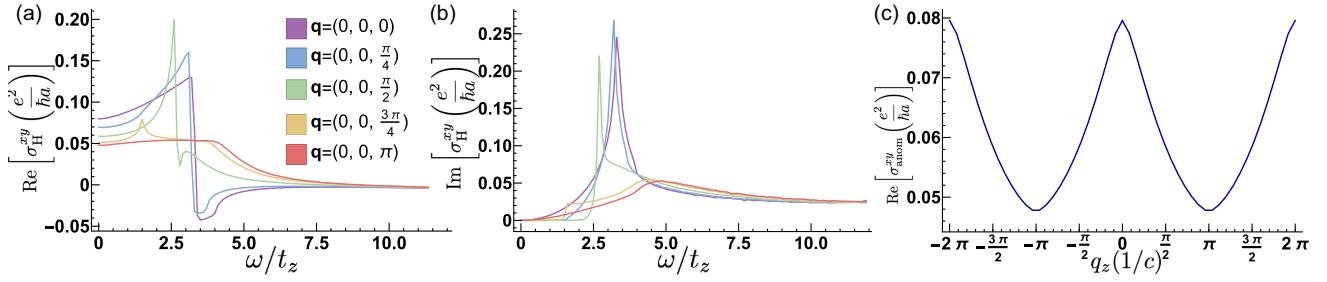


FIG. 5. The real [(a)] and imaginary [(b)] parts of the Hall conductivity as a function of frequency for several values of the wave vector in the  $\hat{z}$  direction. (c) shows the real part of the anomalous Hall conductivity [as defined in Eq. (100)] as a function of the wave vector changing in the  $\hat{z}$  direction. The parameters  $m = 4t_z$ ,  $t_x = t_y = t_z$ ,  $\gamma = 0$ ,  $\gamma_z = 0$ , and  $\mu = 0$  are used for these calculations.

$\sigma_H^{xy}(\omega, \pi\hat{x}) \neq 0$  generically. Examining Figs. 4(a) and 4(b), we see that, as the magnitude of  $\mathbf{q} = q\hat{x}$  is increased, the magnitude of the Hall conductivity decreases. We also see peaks in the Hall response at frequencies corresponding to indirect (nonzero momentum transfer) transitions between occupied and unoccupied states induced by the external electromagnetic field.

### 3. Hall response when $\mathbf{q} = (0,0,q_0)$

Now we take the orientation of the incoming field to be in the  $\hat{z}$  direction, which is parallel to the Weyl node separation. Unlike when  $\mathbf{q}$  was in the  $\hat{x}$  orientation in Sec. VA 2, the three different definitions of the current operators yield identical predictions for the Hall conductivity in this case. This is because, for our toy model Eq. (101), only the matrix elements Eqs. (77), (79), and (80) change in the Kubo formula for the conductivity. Since our model has no diagonal hopping and since the Hall conductivity depends only on the  $x$  and  $y$  components of the current, we have that  $F_{xy}^{ab}(\mathbf{k}, \mathbf{q}) = F_{\text{mid};xy}^{ab}(\mathbf{k}, \mathbf{q}) = \tilde{F}_{xy}^{ab}(\mathbf{k}, \mathbf{q}) = F_{xy}^{ab}(\mathbf{k}, 0)$ . To change this, we could include diagonal hoppings of the form  $\sin(k_i) \rightarrow \sin(k_i)f(k_z)$  or  $\cos(k_i) \rightarrow \cos(k_i)f(k_z)$  for  $i \in \{x, y\}$  and  $f(k_z)$  being a general periodic function. Mixing the  $k_x$  or  $k_y$  dependence with  $k_z$  gives rise to the possibility of a Hall conductance where each of the three mentioned methods would then yield different results.

In Figs. 5(a) and 5(b), we show the real and imaginary parts of the Hall conductivity  $\sigma_H^{xy}(\mathbf{q} = q_z\hat{z}, \omega)$  for five values of  $q_z$ . We point out two relevant wave vector values:  $q_z = \pi/2$  (the momentum transfer of the transition from one Weyl node to the highest-energy eigenstate) and  $q_z = \pi$  (momentum transfer of the transition between Weyl nodes at zero frequency). The  $q_z = \pi/2$  transition is shown in Fig. 2 with the green dotted arrow, and the  $q_z = \pi$  is illustrated in that same figure with a yellow dotted line. At  $q_z = \pi/2$ , we would expect a large Hall response, since the density of states at the band maximum is large. To examine the Hall effect in this orientation, we inspect Figs. 5(a) and 5(b), which plot the Hall coefficient as a function of frequency for several wave vectors. At the value  $q_z = \pi/2$ ,

we notice the real part of the Hall effect achieves its maximal value at  $\omega \sim 2.63t_z$ .

The other interesting value of  $q_z$  is at  $\pi$ . At  $q_z = \pi$ , a zero-frequency electric field can excite transitions between the two Weyl nodes. This is, an electron can directly hop from one Weyl node to the other without having to absorb energy. This type of excitation at  $\mathbf{q} = \pi\hat{z}$  puts the electron in the strongest differential of Berry curvature possible, hopping from one topologically charged Fermi point to the Fermi point with opposite charge. We find that, at low frequencies, this leads to a Hall conductivity  $\sigma_H^{xy}(\mathbf{q} = \pi\hat{z}, \omega)$  with a real part that is nearly constant over a wide frequency range, before decaying to zero at high frequencies. We see this in Fig. 5(a). In Fig. 5(b), see that the imaginary part of  $\sigma_H^{xy}(\mathbf{q} = \pi\hat{z}, \omega)$  is nearly linear at lower frequencies and then decays to zero at high frequencies.

We next consider the anomalous Hall conductivity  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = q_z\hat{z})$  as a function of the wave vector. We plot the anomalous Hall conductivity as a function of the  $\mathbf{q} = q_z\hat{z}$ , in Fig. 5(c). We see that the anomalous Hall conductivity for this model is periodic as a function of  $q_z$  due to the absence of diagonal hopping. We also see that the Hall conductivity decreases quadratically at small  $q_z$ , consistent with previous investigations into  $\mathcal{O}(\mathbf{q}^2)$  corrections to the Hall conductivity [44,45,98].

Lastly, to study the effect of a finite Fermi surface on the Hall conductivity, we study  $\sigma_H^{xy}(\omega, \mathbf{q} = q_z\hat{z})$  for our Weyl semimetal as a function of chemical potential. When the Fermi surface volume is finite, we expect a diverse range of intra- and inter-Fermi surface indirect transitions to contribute to the  $\mathbf{q}$  dependence of the Hall conductivity.

First, we focus on the anomalous Hall conductivity as a function of wave vector for several different values of the chemical potential. For  $|\mu| < |2t_z|$ , we see from Fig. 2 that the Fermi surface consists of two disjoint pockets centered on each Weyl node. At  $|\mu| = |2t_z|$ , there is a Lifshitz transition, where the two Fermi pockets meet; for  $|\mu| > |2t_z|$ , there is a single Fermi pocket. Since the density of states and topological charge of the Fermi surface change drastically at the Lifshitz transition, we expect to see a drastic change in the Hall conductivity near  $\mu = 2t_z$  for all  $\mathbf{q}$ .

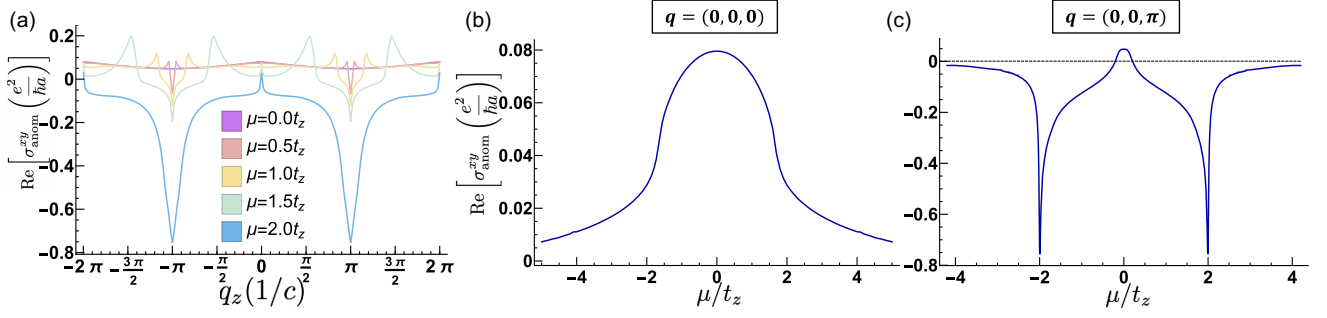


FIG. 6. Chemical potential dependence of the anomalous Hall conductivity [as defined in Eq. (100)] for the Weyl semimetal model. (a) shows the real part of the anomalous Hall response as a function of the wave vector in the  $\hat{z}$  direction, for several different values of  $\mu$  up to the Lifshitz transition at  $\mu = 2t_z$ . (b) shows the real part of the anomalous Hall response as a function of chemical potential at  $q_z = 0$ , and (c) shows the real part of the anomalous Hall response as a function of chemical potential at  $q_z = \pi$ . The parameters  $m = 4t_z$ ,  $t_x = t_y = t_z$ ,  $\gamma = 0$ , and  $\gamma_z = 0$  are used in these calculations.

We show this in Fig. 6(a), where we plot the anomalous Hall conductivity  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = q_z \hat{z})$  as a function of  $q_z$  for five different values of the chemical potential. We see that, when  $\mu = 2t_z$ , the anomalous Hall conductivity is positive at  $q_z = 0$  but decreases dramatically as a function of  $q_z$ . For larger  $q_z$  the anomalous Hall conductivity reverses sign and achieves a large negative value at  $q_z = \pi$ ; the Hall response at  $q_z = \pi$  and  $\mu = 2t_z$  is almost an order of magnitude larger than the topological anomalous Hall response  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = 0)$ .

To investigate this further, we compute the anomalous Hall conductivity  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = q_z \hat{z})$  at  $\omega = 0$  and at fixed  $q_z$  as a function of  $\mu$ , shown in Fig. 6(b) for  $q_z = 0$ , and we see that  $\sigma_{\text{anom}}^{xy}(q_z = 0)$  decreases quadratically as  $\mu \rightarrow 0$ , consistent with previous results [99–102]. Importantly, the anomalous Hall response at  $q_z = 0$  is proportional to the Berry curvature at the Fermi surface leading to this quadratic dependence on the chemical potential [99–102]. Furthermore, we see in Fig. 6(b) that  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = 0)$  is always positive for any value of  $\mu$ . Therefore, since we know from previous literature [36,93] that this contribution is proportional to the Berry curvature, we can establish both the positivity and quadratic behavior as indicators of a Berry curvature-dominated influence at the Fermi surface.

However, as noted previously,  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = \pi \hat{z})$  demonstrates very different behavior as a function of  $\mu$ , as shown in Fig. 6(c). We see that the anomalous response is positive for  $\mu = 0$  and then, as  $|\mu|$  grows, eventually switches sign and achieves a large negative value at the Lifshitz transition. The behavior of  $\sigma_{\text{anom}}^{xy}(\mathbf{q} = \pi \hat{z})$  near the  $\mu = 2t_z$  arises due to the divergence of the joint density of states [essentially, the velocity-independent universal factor in our Kubo formula Eq. (77)] at the Lifshitz transition.

## B. Kerr rotation in moiré materials

While the wave vector dependence of the optical conductivity can usually be ignored in most solids, the

large length scales present in moiré lattice systems mean that spatial inhomogeneities in optical electromagnetic fields may have an appreciable effect on transport. In this section, we focus on applying our formalism to compute the wave vector dependence of the Kerr effect in 2D systems. Recall that the Kerr effect describes the change in relative angle and ellipticity from the light reflected from a material [22,103] and is a direct probe of time-reversal symmetry breaking [104]. As such, the polar Kerr effect has been used as a direct probe of time-reversal symmetry breaking in unconventional superconductors and charge-ordered systems [105–110]. Additionally, reflectivity studies on 2D materials allow us to probe the frequency and wave vector dependence of response functions off the light cone, since, for a fixed frequency, the magnitude of the in-plane wave vector can be varied by changing the incidence angle. Because of the intrinsic time-reversal symmetry breaking of states in twisted graphene and transition-metal dichalcogenides materials, a study of the Kerr effect in these systems is a natural avenue for future experiments [62].

To this end, we compute the magnitude of the Kerr effect in a model of a moiré-Chern insulator. First, in Sec. V B 1, we introduce a toy model for a Chern insulator in a moiré system. Then, in Sec. V B 2, we compute the Kerr angle and ellipticity for the model, using Eq. (73). Unlike in our analysis of the Hall effect in Sec. V A, here both the paramagnetic and diamagnetic conductivities play a role. We compare our results with approximate calculations using the nonconserved trapezoid [Eq. (32)] and midpoint [Eq. (33)] definitions of the current prevalent in the literature, showing that they yield quantitatively distinct predictions for the Kerr angle and ellipticity that could be distinguished in experiment.

### 1. Moiré Haldane Chern insulator model

For our toy model of a moiré-Chern insulator, we start with the spinless Haldane model on a honeycomb lattice [111–114]. The Bravais lattice vectors connecting

next-nearest-neighbor honeycomb lattice sites (in the same sublattice) are, in Cartesian coordinates,

$$\mathbf{b}_1 = \frac{a}{2} \begin{bmatrix} -\sqrt{3} \\ 3 \end{bmatrix}, \quad (103)$$

$$\mathbf{b}_2 = \frac{a}{2} \begin{bmatrix} -\sqrt{3} \\ -3 \end{bmatrix}, \quad (104)$$

$$\mathbf{b}_3 = a \begin{bmatrix} \sqrt{3} \\ 0 \end{bmatrix}, \quad (105)$$

where  $a$  is the moiré lattice constant. The nearest-neighbor vectors connecting honeycomb lattice sites in opposite sublattices can similarly be written as

$$\mathbf{a}_1 = \frac{a}{2} \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}, \quad (106)$$

$$\mathbf{a}_2 = \frac{a}{2} \begin{bmatrix} -\sqrt{3} \\ 1 \end{bmatrix}, \quad (107)$$

$$\mathbf{a}_3 = a \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \quad (108)$$

We construct a tight-binding model with nearest and next-nearest-neighbor hoppings as well as a staggered on-site potential and an orbital magnetic flux. We can write the Hamiltonian as  $H = \sum_{i=0}^3 h_i(\mathbf{k})\sigma_i$ , with

$$h_0(\mathbf{k}) = 2t_2 \sum_{j=1}^3 \cos(\mathbf{k} \cdot \mathbf{b}_j) - \mu + 3t_2 \cos(\phi), \quad (109)$$

$$h_1(\mathbf{k}) = t \sum_{j=1}^3 \cos(\mathbf{k} \cdot \mathbf{a}_j), \quad (110)$$

$$h_2(\mathbf{k}) = t \sum_{j=1}^3 \sin(\mathbf{k} \cdot \mathbf{a}_j), \quad (111)$$

$$h_3(\mathbf{k}) = M - 2t_2 \sin(\phi) \sum_{j=1}^3 \sin(\mathbf{k} \cdot \mathbf{b}_j). \quad (112)$$

Here,  $M$  is the staggered on-site potential,  $t$  is the nearest-neighbor hopping amplitude, and  $t_2$  is the next-nearest-neighbor hopping amplitude. The parameter  $\mu$  is the chemical potential. Also, the last term in Eq. (109) is modified from Haldane's original derivation by the extra term  $3t_2 \cos(\phi)$ . We include this extra term for the convenience of putting the zero of energy in the band gap. Finally,  $\phi$  is the time-reversal symmetry-breaking magnetic flux per plaquette.

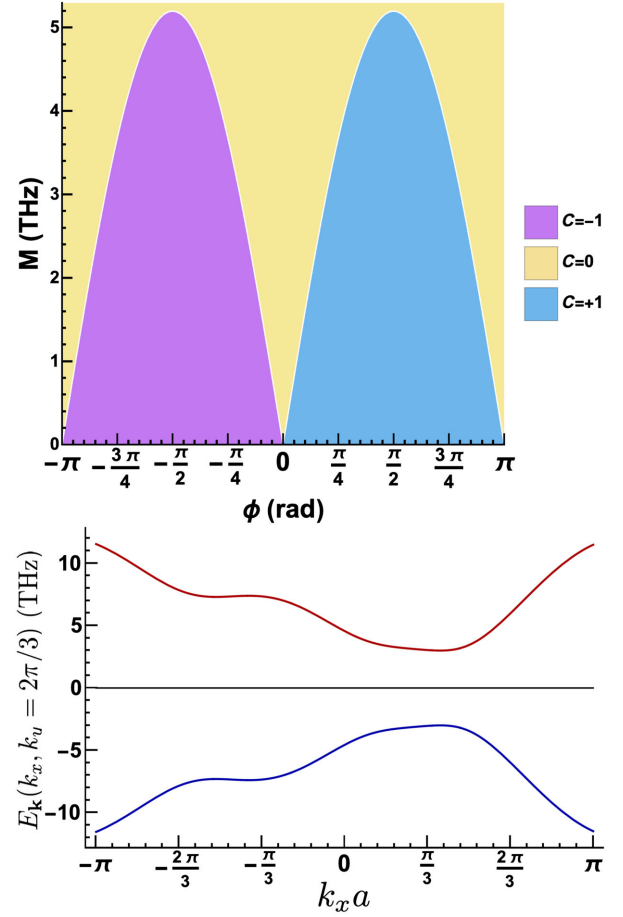


FIG. 7. The top is the topological phase diagram of the Haldane Chern insulator as described by Eqs. (109)–(112). We set  $t = 4t_2$  and  $t_2 = 1$  THz to adhere to the nonoverlap condition of  $|t_2/t| < 1/3$ . The Chern number, as illustrated in this figure, is calculated by integrating the Berry curvature across the Brillouin zone. The bottom is the spectrum for the Haldane Chern insulator, capturing both  $K$  and  $K'$  points when  $k_y = 2\pi/3$ , which is at  $k_x = \pm 2\pi/3\sqrt{3}$ . At these points, the separation in energies follows as  $2|M \pm 3\sqrt{3}t_2 \sin(\phi)|$ , which take on separation values of 6 and 14.8 THz, respectively.

In this model, we take  $|t_2/t| < 1/3$ , which forces the bands to not overlap. Additionally, to be in a topologically nontrivial state with Chern number  $C = \pm 1$ , our choice of parameters must satisfy  $|M/t_2| < 3\sqrt{3}|\sin(\phi)|$ . The phase diagram for this model is shown in Fig. 7.

Furthermore, we note this model contains two gapped Dirac points at points  $K$  and  $K'$ , with coordinates  $(2\pi/3)[(1/\sqrt{3}), 1]$  and  $(2\pi/3)[-(1/\sqrt{3}), 1]$  in units of the inverse moiré lattice constant, respectively. The  $K$  and  $K'$  points are shown in the Brillouin zone in Fig. 8. By adjusting  $M$ ,  $t_2$ , and  $\phi$ , the Dirac nodes can be gapped out. A gap closes at a single Dirac node at each phase boundary as illustrated in Fig. 7. At the  $K$  and  $K'$  points, the band gaps are  $2|M - 3\sqrt{3}t_2 \sin(\phi)|$  and



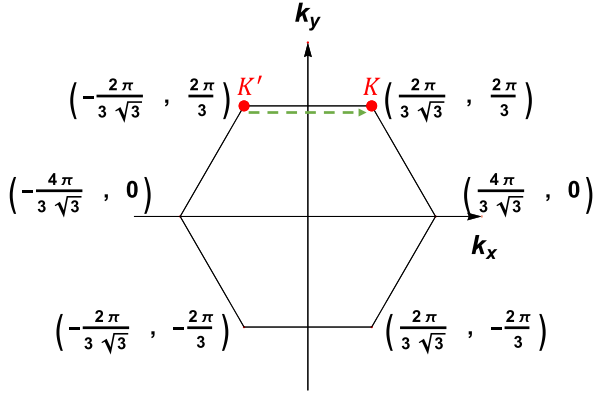


FIG. 8. Brillouin zone for the Chern insulator model in Eqs. (109)–(112) measured in natural units of inverse lattice constants. The dotted green line denotes the momentum transfer between the  $K'$  and  $K$  points. An excitation between these two points can be excited by adjusting the wave vector of the incoming electromagnetic field.

$2|M + 3\sqrt{3}t_2 \sin(\phi)|$ , respectively; we refer to these as the topological gaps. We show the band structure of the model for a representative set of parameters  $t_2 = 1$  THz,  $t = 4$  THz,  $M = 2.2$  THz, and  $\phi = \pi/2$ .

## 2. The momentum-dependent Kerr effect of the Haldane Chern insulator

We now consider the influence of an optical electric field on the moiré-Chern insulator, focusing on the Kerr effect. When the moiré lattice constant is large, the in-plane wave vector  $|\mathbf{q}| \sim \omega/c$  may be an appreciable fraction of the inverse lattice spacing  $1/a$ , at frequencies near the topological gaps. To model this, we choose the moiré length and energy scales of our model to be comparable to those of twisted bilayer graphene reported in the literature [62,115–118]. In particular, we take the lattice constant  $a = 200$  Å, which is achievable at small twist angles in graphene. Additionally, we take  $t_2 = 1$  THz,  $t = 4$  THz, and  $M \approx 4.45$  THz, such that the relevant energy scale for optical excitations is on the order of 10 THz. Additionally, by differentiating the Hamiltonian  $H$  defined in Eqs. (109)–(112) with respect to  $\mathbf{k}$ , we find that the matrix elements of the velocity operator are all on the order of  $ta \sim 2.6 \times 10^{-4}c$  or smaller, where  $c$  is the speed of light; this is comparable to the Fermi velocity in a typical metal.

To compute the Kerr response, we consider the experimental geometry shown in Fig. 9. We consider an infinite sample oriented in the  $xy$  plane, encapsulated within a dielectric substrate with large index of refraction  $n_R = \sqrt{\epsilon_R} = 38$  [119]. We illuminate the system with linearly polarized light at oblique incidence  $\theta_i = 7\pi/16$ ,  $\mathbf{E}_I = E_0 \hat{x} e^{-i\omega t + i\mathbf{q} \cdot \mathbf{r}}$ . Since the in-plane component of the electric field is continuous across the interface, this means

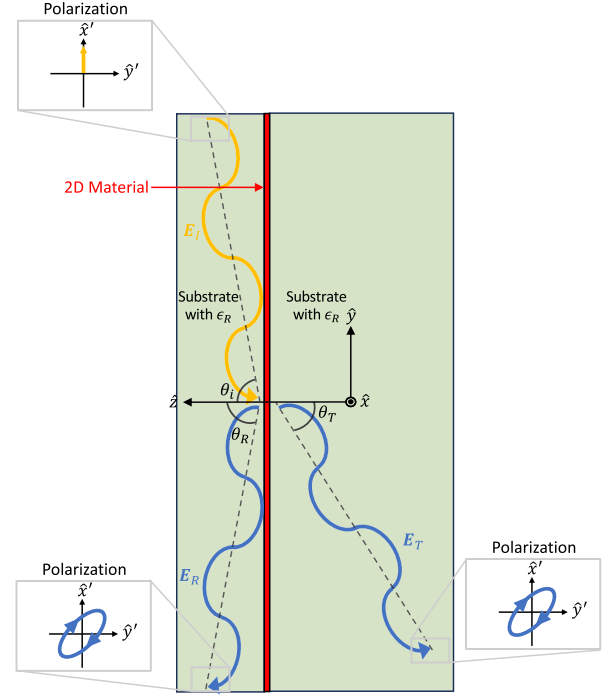


FIG. 9. Diagram of the polarization and configuration of the Kerr effect setup. The 2D sample (red) is encapsulated in a high- $\epsilon_R$  dielectric (green). The yellow curve is the incident light with incidence angle  $\theta_i$  and incident field  $\mathbf{E}_I$ . The left blue curve is the reflected wave with reflected angle  $\theta_R$  and reflected electric field  $\mathbf{E}_R$ . The right blue curve is the transmitted light with transmitted angle  $\theta_T$  and transmitted electric field  $\mathbf{E}_T$ . Each of the reflected and transmitted polarizations can be described in terms of an ellipticity and angle of rotation that describe the Kerr (reflected) and Faraday (transmitted) effects.

that the in-plane component  $q_y$  of the wave vector relevant for scattering is, in dimensionless units,

$$q_y a = a n_R \sin(\theta_i) \omega / (c) \sim (0.0025 \text{ THz}^{-1}) \omega \quad (113)$$

for our choice of moiré lattice constant. This is several orders of magnitude larger than the typical scale  $q_y a_0 \sim (10^{-6} \text{ THz}^{-1}) \omega$  for crystalline systems with typical lattice constants  $a_0 \sim 2$  Å. Taking the incident photon frequency to be on the order of approximately 10 THz, which should be reasonable with recent breakthroughs in terahertz spectroscopy experiments, we see  $\mathbf{q}$  in the moiré system is comparable enough to the lattice spacing to make finite wave vector corrections to optical response large enough to be experimentally measurable; although the effects of nonzero  $\mathbf{q}$  are still small, they are experimentally non-negligible in moiré systems. Although we use a very large  $n_R$  substrate appropriate to metamaterials, we see from Eq. (113) that a lower dielectric constant can be used, provided the moiré lattice constant is similarly increased (e.g., by decreasing the twist angle). Additionally, the effect

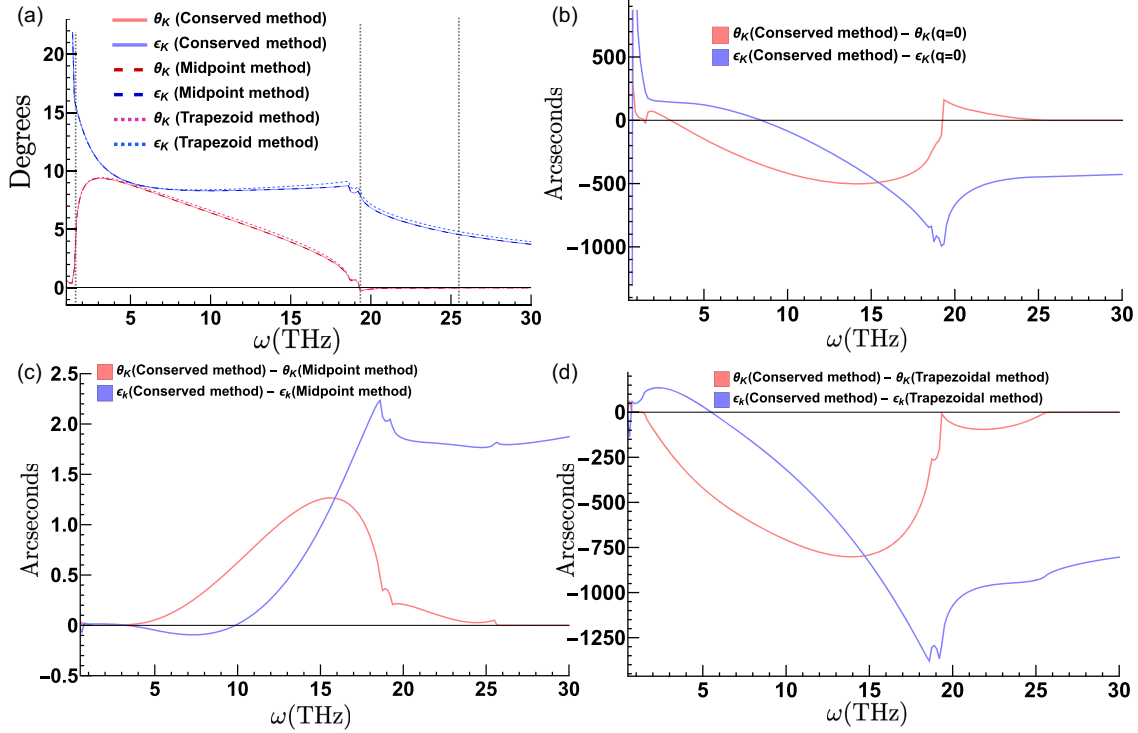


FIG. 10. Kerr effect in the moiré-Chern insulator model. (a) shows the Kerr angle  $\theta_K$  and ellipticity  $\epsilon_K$  as a function of frequency computed using the conserved current, the nonconserved midpoint current, and the nonconserved trapezoid current. (b) shows the difference between the angle and ellipticity computed using conserved current compared to the uniform approximation  $\sigma_{\mu\nu}(\mathbf{q}, \omega) \approx \sigma_{\mu\nu}(\mathbf{0}, \omega)$ , effectively showing the importance of  $\mathbf{q}$ -dependent modifications to the optical response. In (c), we show the differences between the conserved and midpoint current predictions for the Kerr angle and ellipticity. Similarly, in (d), we show the differences between the conserved and trapezoid current predictions for the Kerr angle and ellipticity. These figures use model parameters of  $M = (3\sqrt{3} - 3/4)t_2$ ,  $t = 4t_2$ ,  $t_2 = 1$  THz, and  $\phi = \pi/2$ . Vertical dashed lines denote the two topological gaps and the highest allowed transition frequency.

of the wave vector correction becomes more noticeable as the angle of incidence becomes closer to being oblique.

We can now compute the Kerr angle (rotation of the plane of polarization of the reflected wave) and ellipticity (imaginary part of the Kerr angle) for light reflected off our model of a moiré-Chern insulator. Since our sample is two-dimensional, its primary influence on the propagation of light is through the boundary conditions that enter Maxwell's equations. In particular, the conductivity tensor  $\sigma^{\mu\nu}(\omega, \mathbf{q} = q_y)$  determines the surface current at the sample, which determines the discontinuity in the magnetic field across the sample and, thus, the reflection coefficient. We give a full derivation of the Kerr angle and ellipticity in terms of the conductivity tensor in Appendix E. We can, thus, use our Kubo formula Eq. (73) in terms of the conserved current operator in Eq. (53) to compute the Kerr angle and ellipticity. Since the derivation for the Kerr angle and ellipticity involves both the longitudinal and Hall components of the conductivity tensor, we need to include both the diamagnetic and paramagnetic parts of the conductivity, as presented in Sec. IV A.

In Fig. 10(a), we show the Kerr angle  $\theta_K$  and Kerr ellipticity  $\epsilon_K$  as functions of frequency, using the Kubo

formula in Eq. (73). For comparison, we also include the analogous calculation using the nonconserved trapezoid and midpoint definitions of the current operator. As noted at the end of Sec. II, to use the nonconserved currents we introduce an approximate diamagnetic current in order to compute the corresponding  $\mathbf{q}$ -dependent diamagnetic conductivity. For the parameter values considered here, the topological gaps in the band structure occur at  $\omega = 1.5$  THz and another at  $\omega \approx 19.28$  THz, which are indicated by vertical dotted lines in Fig. 10(a). As expected, we see that the Kerr angle rises rapidly (and the ellipticity falls rapidly) near the smaller topological gap, and the Kerr angle nearly vanishes above the larger topological gap. Note also that, while the ellipticity and Kerr angles also grow at very low frequencies ( $\omega < 0.5$  THz), the intensity of the reflected light goes rapidly to zero at low frequencies. We also indicate with a vertical dotted line the frequency  $\omega = 25.60$  THz, which corresponds to the highest possible excitation energy in this system.

To highlight the non-negligibility of including wave-vector-dependent corrections to the optical response in moiré systems, in Fig. 10(b), we show the difference between the  $\mathbf{q}$ -dependent conserved current calculation

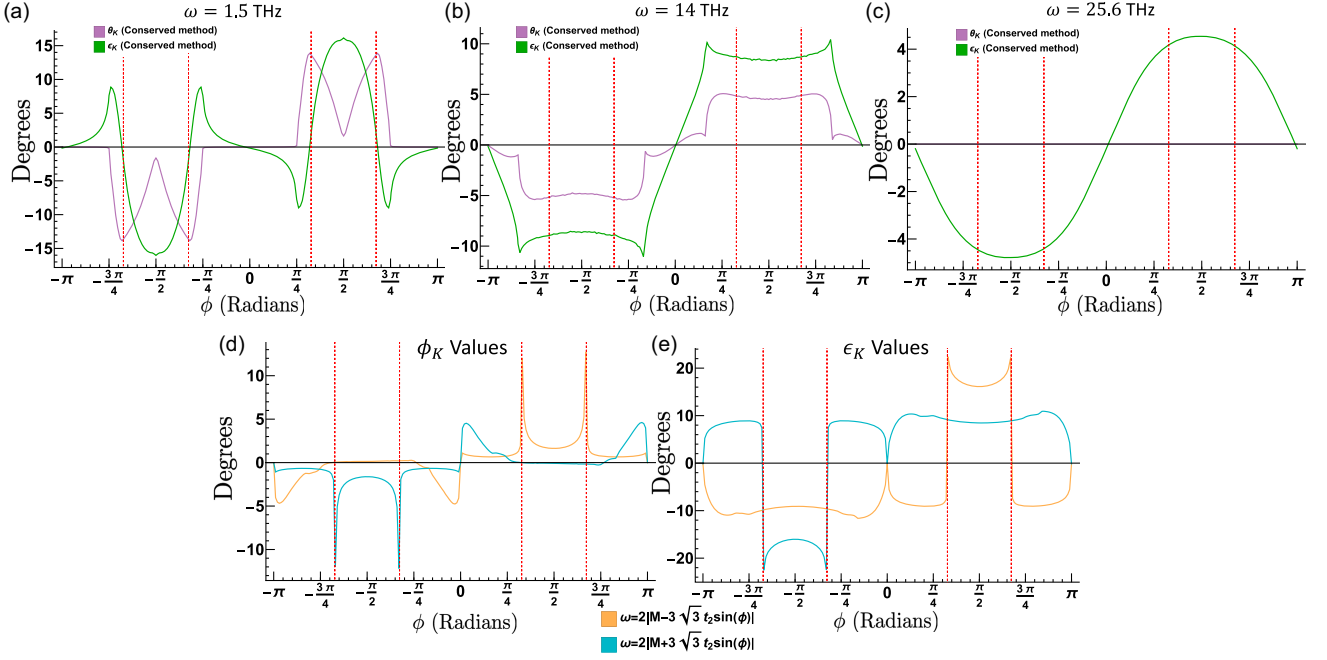


FIG. 11. Kerr angle  $\theta_K$  and ellipticity  $\epsilon_K$  as a function of magnetic flux  $\phi$  per plaquette in the moiré-Chern insulator model, computed using our conserved current in Eq. (53).  $\theta_K$  and ellipticity  $\epsilon_K$  are shown at fixed frequency  $\omega = 1.5$  THz in (a), 14 THz in (b), and 25.6 THz in (c). (d) and (e) show the Kerr angles [(d)] and Kerr ellipticities [(e)] as a function of flux when  $\omega$  is pinned to the  $\phi$ -dependent topological band gap  $\omega(\phi) = 2|M \pm 3\sqrt{3}t_2 \sin(\phi)|$  at the  $K$  (for positive  $\phi$ ) or  $K'$  (for negative  $\phi$ ) point in the Brillouin zone. The dotted red vertical lines indicate values of  $\phi$  corresponding to the topological phase boundaries shown in Fig. 7. All quantities are computed using  $M = (3\sqrt{3} - 3/4)t_2$ ,  $t = 4t_2$ , and  $t_2 = 1$  THz.

of the Kerr angle and ellipticity and the Kerr angle and ellipticity computed using the common  $\mathbf{q} \rightarrow 0$  uniform approximation to the conductivity tensor. This shows the calculational error arising from neglecting the wave vector dependence in the Kerr effect. We see that the difference varies on the order of  $10^3$  arcseconds, which shows that including spatial inhomogeneity when modeling light scattering off the material surface has quantitative effects that cannot be ignored in order to accurately describe the Kerr effect in moiré systems. Moreover, this magnitude of difference is detectable using existing experimental techniques. Thus, this further demonstrates that optical response in moiré materials is sensitive to spatial inhomogeneities in electromagnetic fields.

At the scale of Fig. 10(a), it is difficult to distinguish between the predicted value of  $\theta_K$  and  $\epsilon_K$  from the conserved and nonconserved currents. According to our analysis of the dimensionless in-plane wave vector in Eq. (113), we expect the difference in prediction between the three definitions of the current operator to be small but measurable. To demonstrate this, in Figs. 10(c) and 10(d), we show the difference between  $\theta_K$  and  $\epsilon_K$  calculated with the conserved current and the midpoint [(c)] and trapezoid [(d)] approximations to the current. We see that the differences are on the order of 1 arcsecond for the midpoint current, to as large as almost 1250 arcseconds (or about 20 arcminutes) for the trapezoid current, which are both

within reach of experimental detection. Thus, we expect that the conserved current in Eq. (53) provides a better fit to Kerr effect experiments, especially for terahertz frequencies in moiré lattices.

Next, we turn our attention to the dependence of the Kerr angle on the magnetic flux at fixed frequency. In our tight-binding model, the parameter  $\phi$  controls the magnetic flux through each plaquette; varying  $\phi$  at fixed  $M$  allows us to tune between the topological phases with Chern numbers  $-1$ ,  $0$ , and  $+1$  as indicated in Fig. 7. We examine signatures of the topological phase transition on the Kerr angle and ellipticity, restricting our attention only to predictions using the conserved current of Eq. (53).

In general, we expect the  $\theta_K$  and  $\epsilon_K$  to both be odd functions of  $\phi$ , since all of  $\theta_K$ ,  $\epsilon_K$ , and  $\phi$  are odd under time reversal. Additionally, since  $\phi = 0$  and  $\phi = \pi$  are time-reversal invariant values of the magnetic flux per plaquette, we expect the Kerr angle and ellipticity to vanish for these values of  $\phi$ . We see this reflected in Fig. 11, where we show the Kerr angles and Kerr ellipticities as functions of  $\phi$  for various values of  $\omega$ . Figures 11(a)–11(c) show  $\theta_K$  and  $\epsilon_K$  for  $\omega = 1.5$  THz (the frequency corresponding to the topological gap), 14 THz (an intermediate frequency scale in the model), and 25.6 THz (the highest possible excitation energy in the model), respectively. We show with vertical dashed lines the values of  $\phi$  corresponding to the topological phase boundaries in Fig. 7. We see that, when the

incident photon frequency is on resonance with the topological gap [Fig. 11(a)], the Kerr angle and ellipticity peak near the value of  $\phi$  corresponding to the topological phase transition. On the other hand, when the photon frequency is off resonance [Figs. 11(b) and 11(c)], the evolution of the Kerr angle and ellipticity as a function of  $\phi$  becomes more smooth; for very large frequencies [Fig. 11(c)], we see that the Kerr angle is approximately zero for all  $\phi$ , and the dependence of the ellipticity angle on  $\phi$  resembles a steepened sinusoidal function.

To further explore the sensitivity of the Kerr angle and ellipticity to the topological phase transition in our model, we compute  $\theta_K$  and  $\epsilon_K$  as functions of  $\phi$  at frequency  $\omega(\phi) = 2|M \pm 3\sqrt{3}t_2 \sin(\phi)|$ , which is on resonance with the smaller of the two topological band gaps for every  $\phi$  (shown in Fig. 7); we show these results in Figs. 11(d) and 11(e). In these figures, we examine the case when the frequency is varied in step with the size of the Dirac points' topological gaps at  $\omega = 2|M \pm 3\sqrt{3}t_2 \sin(\phi)|$ . We see that both the Kerr angle and ellipticity on resonance are large and sharply peaked at the phase boundary, with maximum values  $|\theta_K| \gtrsim 40^\circ$  and  $|\epsilon_K| \gtrsim 40^\circ$ , respectively.

Thus, in order to fully understand and quantitatively model optical experiments being performed on moiré materials, it is necessary to account for wave vector dependence of the optical conductivity. In particular, as moiré systems become larger and as fabrication techniques allow for smaller twist angles and longer-wavelength superlattice potentials, the need to understand how  $\mathbf{q}$  affects the conductivity will become more imperative for explaining the physics.

### C. Magnetic properties of insulators

While we have derived our expression for the conductivity tensor in terms of response to an electric field, gauge invariance and Maxwell's equations imply that responses to electric and magnetic fields are inextricably linked. In particular, Faraday's law Eq. (4) implies that a transverse electric field is always accompanied by a magnetic field. Rewriting Eq. (4) in Fourier space, we have that

$$B_{\mu,\mathbf{q}}(\omega) = \frac{1}{\omega} \epsilon_{\mu\nu\lambda} q_\nu E_{\lambda,\mathbf{q}}(\omega), \quad (114)$$

where  $\epsilon_{\mu\nu\lambda}$  is the totally antisymmetric Levi-Civita symbol. Additionally, the momentum-dependent conductivity tensor  $\sigma^{\mu\nu}(\omega, \mathbf{q})$  can be used to derive equilibrium susceptibilities. Recall that, in the limit  $\omega \rightarrow 0$  for generic, nonzero  $\mathbf{q}$ , our perturbing fields  $\mathbf{A}_{\mathbf{q}}(\omega)$  and  $A_{0\mathbf{q}}(\omega)$  become time independent. The electric and magnetic fields corresponding to these potentials are time independent, bounded, and oscillate in space with wave vector  $\mathbf{q}$ . Note that, in this limit, the static electric and magnetic fields interacting with the material are purely dependent on the wave vector  $\mathbf{q}$  and, therefore, offer a promising

experimental setup to probe the physics discussed in this manuscript. Thus, the Hamiltonian in the presence of a static  $\omega \rightarrow 0$  electromagnetic perturbation at nonzero  $\mathbf{q}$  has a static ground state. It follows that taking  $\omega \rightarrow 0$  at fixed nonzero  $\mathbf{q}$  leaves the system in an equilibrium state [120]. This means that there can be no charge transport, and so the ground state current is purely a *magnetization current* [121]:

$$\langle \mathbf{j}_{\mathbf{q}} \rangle = i\mathbf{q} \times \mathbf{M}_{\mathbf{q}}, \quad (115)$$

where  $\mathbf{M}_{\mathbf{q}}$  is the magnetization density (i.e., magnetic dipole moment per unit volume).

Combining Eqs. (114) and (115), we are able to use our formalism for the nonuniform conductivity  $\sigma^{\mu\nu}(\omega, \mathbf{q})$  to calculate magnetic properties of insulators. First, in Sec. V C 1, we derive a formula for the magnetic susceptibility tensor in insulators. Next, in Sec. V C 2, we derive expressions for the magnetic quadrupole moment in finite systems. Finally, in Sec. V C 3, we show that our formalism is consistent with the Streda formula, for which we provide a new derivation. We note that the results of this section are completely general, relying only on gauge invariance and the assumption that any two-particle interactions are independent of momentum; they apply equally well to interacting and noninteracting systems.

#### 1. Magnetic susceptibility in insulators

We derive an expression for the magnetic susceptibility starting from the defining relation for the linear conductivity:

$$\langle j_{\mathbf{q}}^\mu(\omega) \rangle = \sigma^{\mu\nu}(\omega, \mathbf{q}) E_{\nu,\mathbf{q}}(\omega). \quad (116)$$

While the derivations in this section do not make use of the explicit form of the current operator, they highlight constraints on the conductivity imposed by charge conservation that are satisfied only when the conserved current operator is used. We can rewrite Eq. (116) in the  $A_0 = 0$  gauge (which we use for our derivation of the diagrammatic response in Sec. III) as

$$\langle j_{\mathbf{q}}^\mu(\omega) \rangle = i\omega \sigma^{\mu\nu}(\omega, \mathbf{q}) A_{\nu,\mathbf{q}}(\omega). \quad (117)$$

Now let us suppose that our perturbing field  $A_{\nu,\mathbf{q}}$  is time independent, which in Fourier space implies that it consists of only an  $\omega = 0$  component. Defining

$$R^{\mu\nu}(\mathbf{q}) = \lim_{\omega \rightarrow 0} i\omega \sigma^{\mu\nu}(\omega, \mathbf{q}), \quad (118)$$

we have that the time-independent current response is given by

$$\langle j_{\mathbf{q}}^\mu \rangle = R^{\mu\nu}(\mathbf{q}) A_{\nu,\mathbf{q}}, \quad (119)$$



where the current and vector potential are taken at  $\omega = 0$ . From Eq. (115), we know that  $R^{\mu\nu}(\mathbf{q})A_{\nu,\mathbf{q}}$  must be expressible as the cross product of  $\mathbf{q}$  with a vector. This allows us to write

$$R^{\mu\nu}(\mathbf{q})A_{\nu,\mathbf{q}} = i\epsilon_{\mu\lambda\rho}q_\lambda\alpha^{\rho\nu}(\mathbf{q})A_{\nu,\mathbf{q}} \quad (120)$$

for some tensor  $\alpha^{\rho\nu}(\mathbf{q})$ . Furthermore, gauge invariance restricts the form of  $\alpha^{\rho\nu}(\mathbf{q})$ ; the average current can depend on  $A_{\nu,\mathbf{q}}$  only through the magnetic field  $B_{\nu,\mathbf{q}} = i\epsilon_{\nu\lambda\rho}q_\lambda A_{\rho,\mathbf{q}}$ , and so

$$\alpha^{\rho\nu}(\mathbf{q}) = i\chi^{\rho\lambda}(\mathbf{q})\epsilon_{\lambda\mu\nu}q_\mu. \quad (121)$$

Combining Eqs. (115), (119), (120), and (121), we have

$$\begin{aligned} i\epsilon_{\mu\nu\lambda}q_\nu M_{\lambda,\mathbf{q}} &= -\epsilon_{\mu\lambda\rho}q_\lambda\chi^{\rho\sigma}(\mathbf{q})\epsilon_{\sigma\delta\nu}q_\delta A_{\mathbf{q},\nu} \\ &= i\epsilon_{\mu\lambda\rho}q_\lambda\chi^{\rho\sigma}(\mathbf{q})B_{\mathbf{q},\sigma} \end{aligned} \quad (122)$$

and, hence,

$$M_{\lambda,\mathbf{q}} = \chi^{\lambda\sigma}(\mathbf{q})B_{\mathbf{q},\sigma}. \quad (123)$$

Thus, we see that  $\chi^{\lambda\sigma}(\mathbf{q})$  is the wave-vector-dependent magnetic susceptibility. Furthermore, we see from Eqs. (119)–(121) that the magnetic susceptibility is related to the conductivity tensor via

$$\lim_{\omega \rightarrow 0} i\omega\sigma^{\mu\nu}(\omega, \mathbf{q}) = -\epsilon_{\mu\lambda\rho}q_\lambda\chi^{\rho\sigma}(\mathbf{q})\epsilon_{\sigma\delta\nu}q_\delta. \quad (124)$$

From Eq. (124), we can deduce several general features of the conductivity tensor in the  $\omega \rightarrow 0$  limit. First, we see that if the magnetic susceptibility  $\chi^{\mu\nu}$  is finite, then the conductivity tensor is singular in the  $\omega \rightarrow 0$  limit. In particular, Eq. (124) shows that the magnetic susceptibility determines the weight of the  $1/\omega$  pole in the conductivity tensor as  $\omega \rightarrow 0$  at nonzero  $\mathbf{q}$ . This singularity appears only in the transverse component of the conductivity (i.e., the transverse current response to a transverse field). Counterintuitively, this shows that the transverse conductivity can be divergent as  $\omega \rightarrow 0$  even for an insulator. We also deduce for any system with a finite uniform magnetic susceptibility

$$\chi^{\mu\nu} = \lim_{\mathbf{q} \rightarrow 0} \chi^{\mu\nu}(\mathbf{q}), \quad (125)$$

the low-frequency conductivity satisfies

$$\lim_{\omega \rightarrow 0} \sigma^{\mu\nu}(\omega, \mathbf{q}) \sim \frac{i}{\omega} \epsilon_{\mu\lambda\rho}q_\lambda\chi^{\rho\sigma}(\mathbf{q})\epsilon_{\sigma\delta\nu}q_\delta + \mathcal{O}(q^3), \quad (126)$$

and so the singular part of the low-frequency transverse conductivity depends quadratically on  $\mathbf{q}$  and vanishes as  $|\mathbf{q}| \rightarrow 0$ .

To support these conclusions, we use our formalism to compute  $\lim_{\omega \rightarrow 0} \sigma^{\mu\nu}(\omega, \mathbf{q})$  for our model of a moiré-Chern

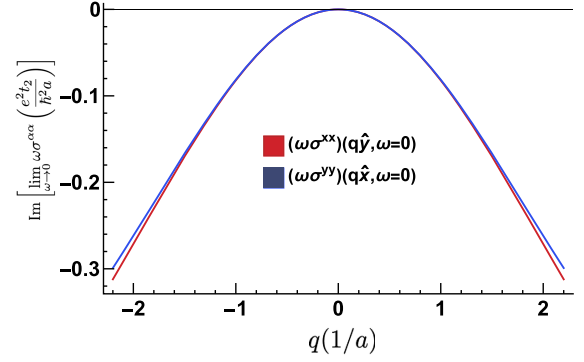


FIG. 12. Plot showing the imaginary part of the diagonal components of the transverse conductivity scaled by the frequency, in the limit of zero frequency for the moiré-Chern insulator model introduced in Eqs. (109)–(112). The real part vanishes. Notice the two curves are similarly quadratic at leading order, and higher-order curvatures separate the curves with increasing  $\mathbf{q}$ . We use model parameters of  $\phi = \pi/2$ ,  $M = (3\sqrt{3} - 3/4)t_2$ , and  $t = 4t_2$ , which puts the ground state in the topological phase at  $C = 1$  as shown in Fig. 7.

insulator described in Eqs. (109)–(112). We show the results in Fig. 12. We see that, in accordance with Eq. (124), we have that

$$\lim_{\omega \rightarrow 0} \text{Im}[\omega\sigma^{xx}(\omega, q\hat{y})] = \lim_{\omega \rightarrow 0} \text{Im}[\omega\sigma^{yy}(\omega, q\hat{x})] \sim -q^2\chi^{zz} \quad (127)$$

for small  $q$ . This figure demonstrates not only the  $q$  dependence, but also the singularity that is linear in  $\omega$ . Notably, this numerical computation does *not* assume a form of the magnetic susceptibility, since it just faithfully carries out the conductivity calculation; equality of  $\lim_{\omega \rightarrow 0} \omega\sigma^{xx}(q\hat{y}, \omega)$  and  $\lim_{\omega \rightarrow 0} \omega\sigma^{yy}(q\hat{x}, \omega)$  in the limit of small  $q$  is a reflection of gauge invariance alone.

We can also invert Eq. (124) to derive an expression for the uniform magnetic susceptibility  $\chi^{\mu\nu}$ . First, we note that, for a system that conserves energy, the uniform magnetic susceptibility must be a symmetric tensor. To derive this, we follow the logic of Ref. [122], which proved an analogous result for the tensor of elastic moduli. We can consider the change in (free) energy for a system as we slowly move the magnetic field through a closed cycle:

$$\begin{aligned} \Delta\langle H \rangle &= -\oint \mathbf{B} \cdot d\mathbf{M} \\ &= -\oint B_\mu \chi^{\mu\nu} dB_\nu \\ &= -\frac{1}{2} \oint B_\mu (\chi^{\mu\nu} - \chi^{\nu\mu}) dB_\nu. \end{aligned} \quad (128)$$

Since the total energy is a state function, its change over every closed cycle must be zero. Thus, we deduce that  $\chi^{\mu\nu} = \chi^{\nu\mu}$ .

Taking two derivatives of Eq. (124) with respect to  $\mathbf{q}$  and making use of the antisymmetry of the Levi-Civita symbol, we find after making use of the symmetry of  $\chi^{\mu\nu}$  that

$$\chi^{\mu\nu} = \lim_{\omega \rightarrow 0} \frac{-i\omega}{3} \epsilon_{\mu\lambda\rho} \epsilon_{\nu\sigma\delta} \frac{\partial^2 \sigma^{\lambda\sigma}}{\partial q_\rho \partial q_\delta} \Big|_{\mathbf{q}=0}. \quad (129)$$

Equation (129) is consistent with and generalizes expressions for the orbital susceptibility that appear in the literature [67,123–125]. It allows us to compute the magnetic susceptibility of insulators using our Kubo formula Eq. (73) in terms of the conserved current operator.

## 2. Magnetic quadrupole moments

In this section, we derive an expression for the magnetic quadrupole moment in the ground state of a finite system with zero external field using our manifestly conserved operator  $\mathbf{j}_q$  defined in Eq. (31). Using Eq. (115), we see that evaluating the average magnetic moment involves expanding our expression for the current operator in powers of  $\mathbf{q}$ ; thus, we expect our conserved current operator to give different predictions for magnetic multipole moments as compared to the nonconserved currents in Eqs. (10) or (11). Since we show in Sec. II that all definitions of the current operator agree to linear order in  $\mathbf{q}$ , we begin our analysis with the magnetic quadrupole moment. In the interest of generality, we here return to consider particles of charge  $\bar{e}$ .

We start by following Refs. [126–128] and take the average of Eq. (115) to find in components

$$\langle j_q^\mu \rangle = i\epsilon^{\mu\nu\gamma} q_\nu M_{\gamma q}. \quad (130)$$

To find the higher-order moments (quadrupole, octupole, etc.), we expand both sides of the equation in orders of the wave vector:

$$\begin{aligned} \langle j_0^\mu \rangle + \langle j_1^\mu \rangle q_\nu + \langle j_2^\mu \rangle q_\nu q_\gamma + \dots \\ = i\epsilon^{\mu\nu\gamma} q_\nu \left( M_{\gamma 0} + M_{\gamma 1}^\alpha q_\alpha + M_{\gamma 2}^{\alpha\beta} q_\alpha q_\beta + \dots \right). \end{aligned} \quad (131)$$

We point out that, in the thermodynamic limit,  $\langle j_0^\mu \rangle = 0$  by Bloch's theorem [129,130]. We now match up the left- and right-hand sides by orders of  $\mathbf{q}$ . The magnetic dipole moment  $M_{\gamma 0}$  has been extensively considered in the literature [73,127]. We focus instead first on the quadrupolar term  $\langle j_2^{\mu\nu\alpha} \rangle = i\epsilon^{\mu\nu\gamma} M_{\gamma 1}^\alpha$ . Using the antisymmetry of the Levi-Civita symbol, we can rewrite this as

$$-\frac{i}{2} \epsilon_{\mu\nu\Delta} \langle j_2^{\mu\nu\alpha} \rangle = M_{\Delta 1}^\alpha. \quad (132)$$

Since

$$\langle j_2^{\mu\nu\alpha} \rangle = \frac{1}{2} \partial_{\mathbf{q}^\nu} \partial_{\mathbf{q}^\alpha} \langle j_q^\mu \rangle \Big|_{\mathbf{q}=0}, \quad (133)$$

we can use our definition of the conserved current in Eq. (31) to find

$$\begin{aligned} \langle j_2^{\mu\nu\alpha} \rangle &= -\frac{\bar{e}}{2} \sum_i \int_0^1 d\lambda \langle (1-\lambda)^2 x_i^\nu x_i^\alpha v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) + \lambda^2 v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\nu x_i^\alpha + \lambda(1-\lambda) x_i^\nu v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\alpha + \lambda(1-\lambda) x_i^\alpha v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\nu \rangle \\ &= -\frac{\bar{e}}{6} \sum_i \left\langle \{x_i^\nu x_i^\alpha, v_i^\mu(\mathbf{p}_i, \mathbf{x}_i)\} + \frac{1}{2} x_i^\nu v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\alpha + \frac{1}{2} x_i^\alpha v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\nu \right\rangle. \end{aligned} \quad (134)$$

This implies from Eq. (134) that the magnetic quadrupole moment takes the form

$$\begin{aligned} M_{\Delta 1}^\alpha &= -\frac{i\bar{e}}{12} \epsilon_{\mu\nu\Delta} \sum_i \left\langle \{x_i^\nu x_i^\alpha, v_i^\mu(\mathbf{p}_i, \mathbf{x}_i)\} \right. \\ &\quad \left. + \frac{1}{2} x_i^\nu v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\alpha + \frac{1}{2} x_i^\alpha v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\nu \right\rangle. \end{aligned} \quad (135)$$

Given that we know our conserved current in Eq. (31) differs from the nonconserved trapezoid [(32)] and midpoint [(33)] at the order of  $\mathbf{q}^2$ , we expect Eq. (135) to differ from conventional expressions for the magnetic quadrupole moment that have appeared in the literature. Indeed, if we insert Eq. (32) for the nonconserved midpoint current  $\tilde{j}$  operator into Eq. (133), we find the alternative expression

$$\langle \tilde{j}_2^{\mu\nu\alpha} \rangle = -\frac{\bar{e}}{4} \sum_i \langle \{x_i^\nu x_i^\alpha, v_i^\mu(\mathbf{p}_i, \mathbf{x}_i)\} \rangle, \quad (136)$$

which also gives a quadrupole moment of

$$\tilde{M}_{\Delta 1}^\alpha = -\frac{i\bar{e}}{8} \epsilon_{\mu\nu\Delta} \sum_i \left\langle \{x_i^\nu x_i^\alpha, v_i^\mu(\mathbf{p}_i, \mathbf{x}_i)\} \right\rangle. \quad (137)$$

Finally, if we use instead the midpoint current  $\mathbf{j}_{\text{mid},q}$  of Eq. (33), we find the alternative expression

$$\begin{aligned} \langle j_{\text{mid},2}^{\mu\nu\alpha} \rangle &= -\frac{\bar{e}}{8} \sum_i \left\langle \{x_i^\nu x_i^\alpha, v_i^\mu(\mathbf{p}_i, \mathbf{x}_i)\} \right. \\ &\quad \left. + \langle x_i^\nu v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\alpha + x_i^\alpha v_i^\mu(\mathbf{p}_i, \mathbf{x}_i) x_i^\nu \rangle \right\rangle, \end{aligned} \quad (138)$$

yielding a quadrupole moment of

$$M_{\text{mid},\Delta 1}^{\alpha} = -\frac{i\bar{e}}{16}\epsilon_{\mu\nu\Delta}\sum_i\left\langle\left\{x_i^{\nu}x_i^{\alpha},v_i^{\mu}(\mathbf{p}_i,\mathbf{x}_i)\right\}+x_i^{\nu}v_i^{\mu}(\mathbf{p}_i,\mathbf{x}_i)x_i^{\alpha}+x_i^{\alpha}v_i^{\mu}(\mathbf{p}_i,\mathbf{x}_i)x_i^{\nu}\right\rangle. \quad (139)$$

By comparing Eqs. (135), (137), and (139), we see that the conserved current predicts a different value for the ground state magnetic quadrupole moment compared to the nonconserved currents conventionally used in the literature. This suggests that care must be exercised when computing the magnetic quadrupole moment in effective low-energy models, for which the nonconserved currents may not correspond to physically relevant observables. Furthermore, similar discrepancies will appear in the octupole magnetic moment, as well as all higher moments.

### 3. Gauge invariance and the Streda formula

Continuing with our study of response to time-independent fields, we now specialize to two dimensions and show how the Streda formula [39,40] arises as a consequence of gauge invariance and Eq. (115). While this result is somewhat tangential to our main argument, it highlights the importance of defining the  $\mathbf{q}$ -dependent conductivity using the conserved current operator. To begin, recall that although we derived the Kubo formula for the conductivity in the  $A_0 = 0$  gauge, gauge invariance requires that the conductivity tensor *also* governs the response to gradients of  $A_0$ . In particular, let us consider response to a longitudinal electric field  $\mathbf{E}_{\mathbf{q}}(\omega) = -i\mathbf{q}A_{0,\mathbf{q}}(\omega)$ . We have from Eq. (116) that

$$\langle j_{\mathbf{q}}^{\mu}(\omega) \rangle = -i\sigma^{\mu\nu}(\omega, \mathbf{q})q_{\nu}A_{0,\mathbf{q}}(\omega). \quad (140)$$

We now take the  $\omega \rightarrow 0$  limit at fixed nonzero  $\mathbf{q}$ . Recall from our discussion preceding Eq. (115) that, in this limit, the perturbing electric field is static, bounded, and spatially periodic. Thus, the system remains in a perturbed equilibrium state, and the only currents that flow are magnetization currents. Hence, we have

$$ie^{\mu\nu\lambda}q_{\nu}M_{\lambda,\mathbf{q}} = -i\lim_{\omega \rightarrow 0}\sigma^{\mu\nu}(\mathbf{q}, \omega)q_{\nu}A_{0,\mathbf{q}}. \quad (141)$$

We can expand Eq. (141) to lowest order in  $\mathbf{q}$  to obtain

$$\epsilon^{\mu\nu\lambda}M_{\lambda,0} = -\sigma^{\mu\nu}(0,0)A_0, \quad (142)$$

where we have introduced

$$\sigma^{\mu\nu}(0,0) = \lim_{\mathbf{q} \rightarrow 0} \lim_{\omega \rightarrow 0} \sigma^{\mu\lambda}(\omega, \mathbf{q}) \frac{q_{\lambda}q_{\nu}}{|\mathbf{q}|^2}. \quad (143)$$

Finally, we note that a time- and space-independent scalar potential  $A_0$  is indistinguishable from (the negative) of a uniform variation of the chemical potential  $-\delta\mu$ . Taking derivatives on both sides then yields the generalized Streda relation

$$\sigma^{\mu\nu}(0,0) = e^{\mu\nu\lambda} \frac{\partial M_{\lambda}}{\partial \mu}. \quad (144)$$

Finally, for an insulator, we know that the longitudinal component  $\sigma^{\mu\nu}(\omega, \mathbf{q})q_{\nu}$  must be analytic and vanishing as  $\mathbf{q}$  and  $\omega$  both go to zero [45]. This allows us to interchange the order of limits in Eq. (143) and identify the susceptibility  $\sigma^{\mu\nu}(0,0)$  with the dc Hall conductivity. We see then that, for an insulator, the only nonvanishing component of the response to a longitudinal electric field at zero frequency is the Hall conductivity, and the components of the Hall conductivity are equal to the derivatives of the magnetization with respect to chemical potential. We note that this derivation makes clear the importance of considering an insulating system: It is only for an insulator that we can guarantee the regularity of  $\sigma^{\mu\nu}(\omega, \mathbf{q})q_{\nu}$ .

## VI. SECOND-ORDER RESPONSE IN MOIRÉ MATERIALS

Impressive experimental advances in recent years have spurred a renewed interest in nonlinear electromagnetic responses in (topological) materials. Theoretical and experimental investigations into the nonlinear optical response of electrons in crystals has yielded new insights into band topology and geometry that can be probed using spatially uniform ( $\mathbf{q} \rightarrow 0$ ) optical fields [13–16,29–31,34,36,38,64,65,76,131–133]. So far, relatively little attention has been paid to the nonlinear response to finite  $\mathbf{q}$  electromagnetic fields. As we show in Sec. V B, in moiré materials even the wave vector dependence of optical fields may play an important part in determining the electromagnetic response.

Recently, Ref. [43] examined the wave-vector-dependent longitudinal nonlinear conductivity at low order in wave vector. Here, however, we show how our diagrammatic perturbation theory of Sec. III can be used in conjunction with our definition of the conserved current in Eq. (53) to compute longitudinal and transverse nonlinear conductivities for arbitrary wave vector. We focus primarily on the second-order response. In Sec. VI A, we write down and evaluate the Feynman diagrams for the second-order nonlinear conductivity  $\sigma^{\mu\nu\gamma}(\omega_1, \omega_2, \mathbf{q}_1, \mathbf{q}_2)$  that determines the second-order response via

$$\langle \mathbf{j}_{\mathbf{q}}^{\mu}(\omega) \rangle = \int d\omega_1 d\omega_2 d\mathbf{q}_1 d\mathbf{q}_2 \delta(\omega - \omega_1 - \omega_2) \delta(\mathbf{q} - \mathbf{q}_1 - \mathbf{q}_2) \sigma^{\mu\gamma\nu}(\omega_1, \omega_2, \mathbf{q}_1, \mathbf{q}_2) E_{\gamma,\mathbf{q}_1}(\omega_1) E_{\nu,\mathbf{q}_2}(\omega_2). \quad (145)$$

This requires using all of the diagrammatic rules of Sec. III B and involves a careful treatment of the  $\mathbf{q}$ -dependent diamagnetic current vertices from Sec. II C. Next, in Sec. VI B, we apply our results to compute the second-order response of our toy model of a moiré-Chern insulator from Eqs. (109)–(112). Motivated by experimental considerations, we focus on spatially rectified ( $\mathbf{q}_1 = -\mathbf{q}_2$ ) second-harmonic generation ( $\omega_1 = \omega_2$ ). This component of the nonlinear conductivity can be probed using transient grating spectroscopy techniques [134–138]. In such measurements, a pair of coherent lasers are made to interfere at the sample to generate a sinusoidal electric field, whose wave vector can be tuned by changing the wavelength of the beams and the angle between the beams. The nonlinear response of the system to such a finite wave vector perturbation is then determined using a probe pulse. By using light in the visible to extreme ultraviolet range, grating wave vectors on the order of  $0.001\text{--}0.1 \text{ \AA}^{-1}$  can be achieved. For moiré systems with lattice constants  $a \sim 100 \text{ \AA}$ , this means that the nonlinear response can be measured over a range of wave vectors spanning multiple moiré Brillouin zones; even for conventional materials with lattice constants on the order of  $1 \text{ \AA}$ , the transient grating wave vector can span a large portion of the first Brillouin zone. In both cases, a formalism such as ours is necessary to compute the nonlinear response of the manifestly conserved current at large  $\mathbf{q}$ .

### A. Second-order response

We may now turn to higher-order conductivities like the second-order response. We proceed in a similar fashion as in the linear response case, by using the rules and conventions set out in Secs. II C and III B.

To compute the second-order conductivity, we begin by using the diagrammatic rules in Sec. III B to write down the four relevant Feynman diagrams, shown in Fig. 13. Note that, in all cases,  $q_1$  and  $q_2$  flow “into” the fermion loop, while  $q_{12}$  flows “out” of the fermion loop. We see that diagrams Figs. 13(b) and 13(c) involve the two-photon diamagnetic current vertex, while diagram Fig. 13(a) involves the three-photon vertex. Using rule 5 from the

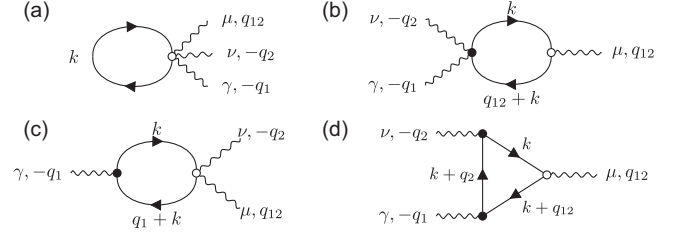


FIG. 13. Second-order response diagrams.

Feynman diagram rules in Sec. III B, we see that, since Figs. 13(b) and 13(c) are invariant under the exchange of  $(\gamma, q_1)$  with  $(\nu, q_2)$ , these diagrams enter into the expression for the conductivity with a multiplicity factor of  $1/2$ .

To this end, we examine a representative example of the input and output vertices in Fig. 13(c) in applying rules 5 and 6 before writing the full mathematical expression from the diagrams. First, the input vertex has a fermion line going into the solid dot vertex, so that  $\mathbf{k}' = \mathbf{k} + \mathbf{q}_1$ . The photon line associated with this line carries a value of  $-\mathbf{q}_1$  as well, so this input vertex goes as

$$\int d\lambda_1 \langle u_{n_1 \mathbf{k} + \mathbf{q}_1} | [\partial_{\mathbf{k}'} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}_1 - (1-\lambda_1)\mathbf{q}_1} | u_{n_2 \mathbf{k}} \rangle. \quad (146)$$

Similarly, we can apply Feynman diagram rules 5 and 6 to the output vertex of the same diagram. Since the fermion line coming into the vertex has momentum  $\mathbf{k}$ , then  $\mathbf{k}' = \mathbf{k}$ . We also see there are two vector potential lines emanating from the open dot vertex: One is going “in” with momentum  $-\mathbf{q}_2$ , and the other is going “out” with  $\mathbf{q}_{12}$ . So the output velocity vertex goes as

$$\int d\lambda'_2 d\lambda_2 \langle u_{n_2 \mathbf{k}} | [\partial_{\mathbf{k}'} \partial_{\mathbf{k}''} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} - (1-\lambda'_2)\mathbf{q}_2 + (1-\lambda_2)\mathbf{q}_{12}} | u_{n_1 \mathbf{k} + \mathbf{q}_1} \rangle. \quad (147)$$

Similarly applying the Feynman rules to each diagram in Fig. 13 and summing the results, we can write the second-order conductivity in terms of Matsubara frequencies as

$$\begin{aligned} \sigma^{\mu\nu}(i\omega_1, i\omega_2, \mathbf{q}_1, \mathbf{q}_2) = & \frac{-e^3}{(i\omega_1)(i\omega_2)} \int dk \left[ \frac{1}{2} G_{n_1}(k) v_{(3), n_1 n_1}^{\mu\nu\gamma}(\mathbf{k}, -\mathbf{q}_1, -\mathbf{q}_2, \mathbf{q}_{12}) \right. \\ & + \frac{1}{2} G_{n_1}(k + q_{12}) v_{(2), n_1 n_2}^{\gamma\nu}(\mathbf{k} + \mathbf{q}_{12}, -\mathbf{q}_1, -\mathbf{q}_2) G_{n_2}(k) v_{(1), n_2 n_1}^{\mu}(\mathbf{k}, \mathbf{q}_{12}) \\ & + G_{n_1}(k + q_1) v_{(1), n_1 n_2}^{\gamma}(\mathbf{k} + \mathbf{q}_1, -\mathbf{q}_1) G_{n_2}(k) v_{(2), n_2 n_1}^{\nu\mu}(\mathbf{k}, -\mathbf{q}_2, \mathbf{q}_{12}) \\ & + G_{n_1}(k + q_{12}) v_{(1), n_1 n_2}^{\gamma}(\mathbf{k} + \mathbf{q}_{12}, -\mathbf{q}_1) G_{n_2}(k + q_2) v_{(1), n_2 n_3}^{\nu}(\mathbf{k} + \mathbf{q}_2, -\mathbf{q}_2) G_{n_3}(k) v_{(1), n_3 n_1}^{\mu}(\mathbf{k}, \mathbf{q}_{12}) \\ & \left. + (\gamma, q_1) \leftrightarrow (\nu, q_2) \right]. \end{aligned} \quad (148)$$



Each of these diagrams must be symmetrized under the exchange of  $(\gamma, \mathbf{q}_1)$  with  $(\nu, \mathbf{q}_2)$ , since, as we see from Eq. (145), these are merely labels for components of the incident electromagnetic field. We indicate this explicit symmetrization in the last line in Eq. (148). Feynman rule 5 in Sec. III B ensures that this explicit symmetrization does

not overcount diagrams, by including multiplicity factors like the  $\frac{1}{2}$  in the second line in Eq. (148) from Fig. 13(b).

In this format, we can carry out the Matsubara frequency integrals and analytically continue back to real frequencies (using the prescription in Refs. [29,65,80] that  $i\omega_1 \rightarrow \omega_1^+$ ,  $i\omega_2 \rightarrow \omega_2^+$ , and  $\omega_{12}^+ \equiv \omega_1 + \omega_2 + 2i\eta$ ) in order to find

$$\begin{aligned} \sigma^{\mu\nu}(\omega_1, \omega_2, \mathbf{q}_1, \mathbf{q}_2) = & \frac{-e^3}{\omega_1^+ \omega_2^+} \int d\mathbf{k} d\lambda_1 d\lambda_2 d\lambda_3 \left[ \frac{1}{2} n_F(\epsilon_{n_1 \mathbf{k}}) \langle u_{n_1 \mathbf{k}} | [\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^\nu} \partial_{\mathbf{k}^\gamma} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} - (1-\lambda_1)\mathbf{q}_1 - (1-\lambda_2)\mathbf{q}_2 + (1-\lambda_3)\mathbf{q}_{12}} | u_{n_1 \mathbf{k}} \rangle \right. \\ & + \frac{1}{2} \frac{n_F(\epsilon_{n_2 \mathbf{k}}) - n_F(\epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}})}{\omega_{12}^+ + \epsilon_{n_2 \mathbf{k}} - \epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}}} \langle u_{n_1 \mathbf{k} + \mathbf{q}_{12}} | [\partial_{\mathbf{k}^\gamma} \partial_{\mathbf{k}^\nu} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}_{12} - (1-\lambda_1)\mathbf{q}_1 - (1-\lambda_2)\mathbf{q}_2} | u_{n_2 \mathbf{k}} \rangle \langle u_{n_2 \mathbf{k}} | [\partial_{\mathbf{k}^\mu} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + (1-\lambda_3)\mathbf{q}_{12}} | u_{n_1 \mathbf{k} + \mathbf{q}_{12}} \rangle \\ & + \frac{n_F(\epsilon_{n_2 \mathbf{k}}) - n_F(\epsilon_{n_1 \mathbf{k} + \mathbf{q}_1})}{\omega_1^+ + \epsilon_{n_2 \mathbf{k}} - \epsilon_{n_1 \mathbf{k} + \mathbf{q}_1}} \langle u_{n_1 \mathbf{k} + \mathbf{q}_1} | [\partial_{\mathbf{k}^\gamma} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}_1 - (1-\lambda_1)\mathbf{q}_1} | u_{n_2 \mathbf{k}} \rangle \langle u_{n_2 \mathbf{k}} | [\partial_{\mathbf{k}^\nu} \partial_{\mathbf{k}^\mu} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} - (1-\lambda_2)\mathbf{q}_2 + (1-\lambda_3)\mathbf{q}_{12}} | u_{n_1 \mathbf{k} + \mathbf{q}_1} \rangle \\ & + \left( \frac{n_F(\epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}})}{(-\omega_1^+ + \epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}} - \epsilon_{n_2 \mathbf{k} + \mathbf{q}_2})(-\omega_{12}^+ + \epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}} - \epsilon_{n_3 \mathbf{k}})} - \frac{n_F(\epsilon_{n_2 \mathbf{k} + \mathbf{q}_2})}{(-\omega_1^+ + \epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}} - \epsilon_{n_2 \mathbf{k} + \mathbf{q}_2})(-\omega_2^+ + \epsilon_{n_2 \mathbf{k} + \mathbf{q}_2} - \epsilon_{n_3 \mathbf{k}})} \right. \\ & \left. + \frac{n_F(\epsilon_{n_3 \mathbf{k}})}{(-\omega_2^+ + \epsilon_{n_2 \mathbf{k} + \mathbf{q}_2} - \epsilon_{n_3 \mathbf{k}})(-\omega_{12}^+ + \epsilon_{n_1 \mathbf{k} + \mathbf{q}_{12}} - \epsilon_{n_3 \mathbf{k}})} \right) \\ & \times \langle u_{n_1 \mathbf{k} + \mathbf{q}_{12}} | [\partial_{\mathbf{k}^\gamma} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}_{12} - (1-\lambda_1)\mathbf{q}_1} | u_{n_2 \mathbf{k} + \mathbf{q}_2} \rangle \langle u_{n_2 \mathbf{k} + \mathbf{q}_2} | [\partial_{\mathbf{k}^\nu} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}_2 - (1-\lambda_2)\mathbf{q}_2} | u_{n_3 \mathbf{k}} \rangle \\ & \times \langle u_{n_3 \mathbf{k}} | [\partial_{\mathbf{k}^\mu} H_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + (1-\lambda_3)\mathbf{q}_{12}} | u_{n_1 \mathbf{k} + \mathbf{q}_{12}} \rangle \\ & \left. + (\gamma, q_1) \leftrightarrow (\nu, q_2) \right]. \end{aligned} \quad (149)$$

This analytic expression in Eq. (149) also allows us to comment on the regularity of the conductivity at low frequencies. In the  $\mathbf{q}_1, \mathbf{q}_2 \rightarrow 0$  limit, it was shown in Refs. [29,43,80] that the diamagnetic vertices in Figs. 13(a)–13(c) serve a crucial role in regulating the low-frequency conductivity; in particular, proper definitions of the diamagnetic current are necessary to ensure that the second-order conductivity for a band insulator satisfies

$$\lim_{\omega_1 \rightarrow 0} \lim_{\omega_2 \rightarrow 0} \sigma^{\mu\nu}(\omega_1, \omega_2, 0, 0) = 0. \quad (150)$$

Using our formalism, we can now examine the behavior of the static ( $\omega_1, \omega_2 \rightarrow 0$ ) second-order conductivity as a function of  $\mathbf{q}_1$  and  $\mathbf{q}_2$ . Since we defined our diamagnetic current vertices to satisfy the generalized Ward identity in Eq. (64), we are guaranteed that for a band insulator the longitudinal components of the static second-order conductivity go to zero as the wave vectors tend to zero. On the other hand, as we show for the linear conductivity in Sec. V C 1, the transverse components of the second-order conductivity can have singularities as we take  $\omega \rightarrow 0$  at fixed  $\mathbf{q}_1$  and  $\mathbf{q}_2$ . These singularities correspond to second-order magnetic and magnetoelectric responses and are not unphysical.

## B. Second-order response in a moiré-Chern insulator: Harmonic generation in frequency and self-focusing in wave vector

Using Eq. (149), we can numerically compute the second-order conductivity as a function of frequency and wave vector for any noninteracting electron system. Typically, calculations of the second-order conductivity have focused on sum frequency generation (where the measured current oscillates as the sum of applied electric field frequencies) or difference frequency generation responses (where the measured current oscillates as the difference of applied electric field frequencies) [29,33,80,139–141]. When we allow for spatially inhomogeneous electric fields, however, the response at second order becomes more complicated. In particular, when the applied electric field varies sinusoidally in space with wave vector  $\mathbf{q}$  and in time with frequency  $\omega$ , we can consider sum frequency generation (measured current oscillating at frequency  $2\omega$ ) and “difference wave vector generation” in wave vector (measured current is spatially uniform with wave vector  $\mathbf{q} - \mathbf{q} = 0$ ). Experimentally, this could be measured in a transient grating experiment by looking at the second-harmonic generation signal.

To see an example of this in practice, we return to our model of a moiré-Chern insulator from Eqs. (109)–(112).

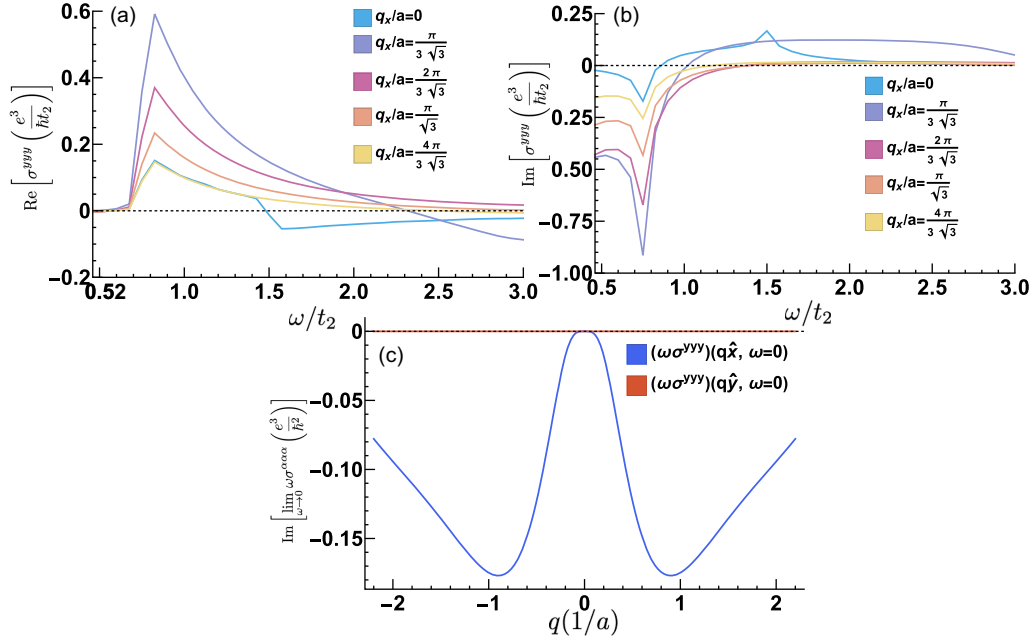


FIG. 14. Plots showing the real (a) and imaginary (b) parts of the second-order transverse conductivities as functions of the frequency,  $\sigma^{yyy}(\omega, q_x, \omega, -q_x; 2\omega, 0)$ . The chosen direction for the conductivities' indices are  $yyy$  and the wave vector is varied in  $q_x \hat{x}$ . (c) shows the imaginary parts of the transverse and longitudinal wave vectors of  $\lim_{\omega \rightarrow 0} \omega \sigma^{yyy}(\omega, \mathbf{q}, \omega, -\mathbf{q}; 2\omega, 0)$ . These figures used model parameters of  $\phi = \pi/2$ ,  $M = (3\sqrt{3} - 3/4)t_2$ , and  $t = 4t_2$ , which puts the topological phase at  $C = 1$  as shown in Fig. 7.

We can consider the “sum-in-frequency, difference-in-wave-vector” response. Since previous works [43] have presented a formalism for computing the longitudinal components of the response, we primarily focus on the transverse response  $\sigma^{yyy}(\omega, \omega, q\hat{x}, -q\hat{x})$  that quantifies the spatially uniform current  $j^y(2\omega)$  that flows in the  $y$  direction in response to a transverse  $y$ -polarized electric field. Such a response could be probed using in-plane polarized terahertz radiation at oblique incidence. Referring to Fig. 8, we see that the wave vector of incident light is parallel to the vector  $\Delta K = (4\pi/3\sqrt{3}a)$  separating the  $K$  and  $K'$  points in the moiré Brillouin zone. In Figs. 14(a) and 14(b), we plot the real and imaginary parts of  $\sigma^{yyy}(\omega, \omega, q\hat{x}, -q\hat{x})$ , respectively, as a function of frequency for five different values of  $q$  between  $q = 0$  and  $q = |\Delta K|$ . We see that for all wave vectors the real part of the response initially grows as a function of frequency, peaks at  $\omega \sim 0.8t_2 = 0.8$  THz—such that  $2\omega$  is approximately equal to the topological gap—and then decreases at larger frequency. Intriguingly, as a function of wave vector, we see that for small values of  $q$  the magnitude of the current at  $\omega \sim 0.8t_2$  initially grows as  $q$  increases, before again decreasing: The peak value of  $\sigma^{yyy}(\omega, \omega, q\hat{x}, -q\hat{x})$  is similar for  $q = 0$  and  $q = |\Delta K|$ .

To conclude, we can also study magnetoelectric response by examining the low-frequency behavior of  $\sigma^{yyy}(\omega, \omega, q\hat{x}, -q\hat{x})$ . In line with our discussion in Sec. VC 1, we expect that, as  $\omega \rightarrow 0$  for fixed  $\mathbf{q}$ , there will be a  $1/\omega$  singularity in  $\sigma^{yyy}(\omega, \omega, q\hat{x}, -q\hat{x})$  whose

weight disperses at least quadratically at small  $\mathbf{q}$ ; the weight of this pole quantifies the magnetization current that flows in response to the external magnetic and electric fields [with the magnetic field determined by Eq. (4)]. At the same time, we expect the longitudinal response  $\omega \sigma^{yyy}(\omega, \omega, q\hat{y}, -q\hat{y})$  to be regular as  $\omega \rightarrow 0$  for any fixed  $\mathbf{q}$ . We can see both properties in Fig. 14(c). We show the quantities  $\lim_{\omega \rightarrow 0} \omega \sigma^{yyy}(\omega, \omega, q\hat{x}, -q\hat{x})$  (in blue) and  $\lim_{\omega \rightarrow 0} \omega \sigma^{yyy}(\omega, \omega, q\hat{y}, -q\hat{y})$  (in red) computed for our moiré-Chern insulator model. Thus, our formalism for computing transverse nonlinear electromagnetic responses correctly captures the low-energy behavior dictated by Maxwell's equations and gauge invariance discussed in Sec. VC 1 for the linear response.

## VII. CONCLUSION

In this work, we have developed a formalism for computing spatially nonuniform (wave-vector-dependent) linear and nonlinear electromagnetic response functions for condensed matter systems. In Sec. II, we introduced a definition of the current operator that can be defined using only the velocity and position operators, independent of detailed knowledge of the microscopic form of the Hamiltonian. Furthermore, unlike the approximations Eqs. (11) and (10), our current in Eq. (31) is manifestly conserved independent of the microscopic details of the Hamiltonian. It is important to emphasize that, while Eqs. (11), (10), and (31) coincide for nonrelativistic Hamiltonians of the form of Eq. (8), only our conserved

current in Eq. (31) remains conserved when Eq. (8) is approximated by truncating the Hilbert space. For non-relativistic systems, we can thus view Eq. (31) as a conserving approximation to the total current applicable to effective models such as Wannier-based tight-binding models. This is crucial for calculations involving approximate models of nonsuperconducting systems with truncated Hilbert spaces, since physically relevant approximations must conserve charge if the total Hamiltonian also conserves charge. Additionally, unlike Eqs. (11) and (10), our conserved current (31) also remains conserved when relativistic corrections to the kinetic energy are taken into account.

Next, we showed how Ward identities could be iteratively applied to determine the diamagnetic current operator order by order in the electromagnetic vector potential. Using our conserved current as a starting point, we continued in Sec. III to develop a diagrammatic perturbation theory for computing spatially nonuniform linear and nonlinear conductivities, focusing on the case of non-interacting electrons for simplicity. Focusing first on the linear conductivity, we showed how our fully charge-conserving approach implies a generalized  $f$ -sum rule relating the density-density response function to the diamagnetic conductivity as a function of wave vector.

We also applied our formalism to compute the wave-vector-dependent Hall conductivity in a toy model of Weyl semimetal. To connect our formalism to experimentally relevant systems, in Sec. VB we introduced a model for a 2D Chern insulator in a moiré superlattice, such that the wavelength of terahertz radiation can be a non-negligible fraction of the moiré lattice scale in certain geometries. We used this model to compute the Kerr angle and ellipticity as a function of frequency for oblique incidence, showing that the effects of spatial nonuniformity are potentially measurable in the next generation of experiments. We also showed how the low-frequency transverse conductivity yields insights into the magnetic susceptibility, magnetic quadrupole moment, and Streda formula for insulating systems.

Finally, we applied our formalism to study second-order response to spatially nonuniform, time-varying electric fields in two-dimensional systems. We calculated the experimentally relevant spatially uniform second-harmonic generation current that flows in response to a transverse, spatially varying ac electric field for a model of a moiré-Chern insulator, which points toward future experimental work on transient grating nonlinear spectroscopy in moiré materials.

Our work opens up several avenues for future theoretical and experimental studies. First, while our explicit calculations in Secs. V and VI were carried out for noninteracting systems, our Feynman diagram formalism in Sec. III can naturally accommodate a treatment of interacting systems by including additional interaction vertices. Our manifestly charge-conserving approach to transverse linear and nonlinear conductivity would thus allow a consistent quantitative treatment of magnetoelectric response in models for

candidate axionic charge density wave materials such as  $(\text{TaSe}_4)_2\text{I}$  [142,143]. Along the same lines, our formalism can be systematically applied to systems with disorder. In the lowest-order approximation, the effect of disorder on the nonlinear conductivities may be phenomenologically accounted for by replacing  $\omega^+ \rightarrow \omega + i/\tau$ , where  $\tau$  is the quasiparticle lifetime [29]. Our formalism, however, can treat disorder scattering more formally by including vertices for scattering of electrons by the disorder potential, and standard diagrammatic techniques [78] for averaging over disorder may be applied. In particular, we expect the (nonlinear) conductivities to depend on both the current vertices defined in Sec. II as well as (nonlinear generalizations of) the diffuson propagator obtained by averaging over disorder. Additionally, our formalism implies additional generalized sum rule relations between nonlinear density response functions and diamagnetic current vertices, generalizing our results of Sec. IV B and making contact with Ref. [3].

As we showed that commonly used approximations for the current operator do not conserve charge within the context of effective models, our work also prompts a reexamination of wave-vector-dependent quantities computed using those approximations, such as the magnetic quadrupole moment in models of higher-order topological insulators [126,128]. Furthermore, in calculations of (nonlinear) x-ray scattering processes where both core and valence electronic states must be treated on equal footing, our response formalism based on the conserved current of Eq. (31) can be used to ensure Ward identities are obeyed even when relativistic corrections to the kinetic energy cannot be ignored [144]. Our work also motivates experimental studies of nonlinear optics in moiré systems, where the effect of spatial inhomogeneity of optical fields may be non-negligible provided samples are large enough.

Finally, recent advances in superlattice and gate engineering open the door to experimentally studying (nonlinear) response to spatially inhomogeneous electromagnetic field outside of optics. In particular, gate-tunable electronic superlattice potentials [145–147] could be modulated in time to create time-dependent electromagnetic fields. While recent theoretical [148] and experimental [149,150] progress along these lines has focused on tunable superlattices in untwisted bilayer graphene systems, applications of these techniques to twisted systems would allow the study of response to time-dependent electromagnetic fields with wave vectors comparable to the moiré lattice spacing. Thus, we emphasize that our results will be directly applicable to the next generation of transport experiments in gate-tunable superlattice devices.

Let us conclude by reemphasizing the importance of the conserved current from Eq. (31). Unlike the approximations of Eqs. (10) and (11) to the minimally coupled current in nonrelativistic systems commonly used in the literature, Eq. (31) is manifestly conserved for any model Hamiltonian. Since most Hamiltonians of interest in

condensed matter systems arise as low-energy approximations to (i.e., truncations of the Hilbert space of) complicated many-body semirelativistic systems, electromagnetic response functions computed within these models will be faithful approximations to what is experimentally measured only if the current used in the calculation is conserved *within the Hilbert space of the model*. All of Eqs. (10), (11), and (31) accomplish this task to linear order in the wave vector, ensuring that they yield the same approximations in (paramagnetic) conductivities to quadratic order in wave vector. However, only Eq. (31) is generally conserved at quadratic order in wave vector and beyond, and only Eq. (31) allows for the determination of diamagnetic current vertices via Ward identities, which are essential for properly regularizing the low-frequency conductivity. Thus, we expect that Eq. (31) is a necessary starting point for obtaining consistent approximations to response functions beyond quadratic order in wave vector.

### ACKNOWLEDGMENTS

The authors thank G. Monteiro, P. Rao, and M. Trigo for helpful discussions. The theoretical work of R. C. M. and B. B. in the development of the response theory formalism was supported by the Alfred P. Sloan Foundation and the National Science Foundation under Grant No. DMR-1945058. The numerical computation of response functions and the development of experimental proposals was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Energy Frontier Research Center for Quantum Sensing and Quantum Materials through Grant No. DE-SC0021238. This work made use of the Illinois Campus Cluster, a computing resource that is operated by the Illinois Campus Cluster Program (ICCP) in conjunction with the National Center for Supercomputing Applications (NCSA) and which is supported by funds from the University of Illinois at Urbana-Champaign.

### APPENDIX A: PROOF OF THE KARPLUS-SCHWINGER RELATION

In this appendix, we review the proof of the Karplus-Schwinger relation [Eq. (29)], which was first presented in Ref. [71]. We consider two operators  $A$  and  $B$ . Let us define a function

$$F(\tau) = e^{\tau(A+\delta t B)}. \quad (\text{A1})$$

We develop an expansion for  $F(1) = \exp(A + \delta t B)$  as a power series in  $\delta t$ . To do so, first note that

$$F'(\tau) = (A + \delta t B)F(\tau). \quad (\text{A2})$$

Defining

$$G(\tau) \equiv e^{-\tau A} F(\tau), \quad (\text{A3})$$

we find that

$$\begin{aligned} G'(\tau) &= -AG(\tau) + e^{-\tau A} F'(\tau) \\ &= -AG(\tau) + e^{-\tau A} (A + \delta t B)F(\tau) \\ &= \delta t e^{-\tau A} B e^{\tau A} G(\tau). \end{aligned} \quad (\text{A4})$$

Note that, from Eqs. (A1) and (A3), we also have that  $G(0) = 1$ . Introducing

$$C(\tau) = e^{-\tau A} B e^{\tau A} \quad (\text{A5})$$

and recasting Eq. (A4) as an integral equation, we have

$$G(\tau) = 1 + \int_0^\tau d\lambda \delta t C(\lambda) G(\lambda). \quad (\text{A6})$$

This integral equation can be solved iteratively by means of a Dyson series. We find that

$$\begin{aligned} G(\tau) &= \mathcal{P} e^{\int_0^\tau \delta t C(\lambda) d\lambda} \\ &= 1 + \delta t \int_0^\tau C(\lambda) d\lambda \\ &\quad + (\delta t)^2 \int_0^\tau d\lambda \int_0^\lambda d\lambda' C(\lambda) C(\lambda') + \mathcal{O}[(\delta t)^3], \end{aligned} \quad (\text{A7})$$

where  $\mathcal{P}$  represents the path ordering (in  $\lambda$ ) of the exponential.

Setting  $\tau = 1$ , multiplying by  $e^A$ , and using the definition of Eq. (A5) for  $C(\tau)$ , we have

$$\begin{aligned} F(1) &= e^A G(1) = e^{A+\delta t B} \\ &= e^A + \delta t \int_0^1 e^A C(\lambda) d\lambda \\ &\quad + (\delta t)^2 \int_0^1 d\lambda \int_0^\lambda d\lambda' e^A C(\lambda) C(\lambda') + \mathcal{O}[(\delta t)^3] \\ &= e^A + \delta t \int_0^1 e^{(1-\lambda)A} B e^{\lambda A} d\lambda + \mathcal{O}[(\delta t)^2], \end{aligned} \quad (\text{A8})$$

which is equal to Eq. (29).

### APPENDIX B: EXAMPLE: LONGITUDINAL AND TRANSVERSE CURRENT OPERATOR FOR SEMIRELATIVISTIC FREE ELECTRONS

As an example and verification of our expression for the current operator derived from Eq. (31), we consider the semirelativistic free-electron Hamiltonian

$$H_{\text{SR}} = \sum_i \frac{|\mathbf{p}_i|^2}{2m} + \frac{|\mathbf{p}_i|^4}{8m^3 c^2}. \quad (\text{B1})$$

By minimally coupling  $H_{\text{SR}}$  to a background vector potential, we can derive from Eq. (22) that the current operator is given by



$$\begin{aligned}
\mathbf{j}_{\min, \mathbf{q}} &= -\bar{e} \int d\mathbf{r} e^{-i\mathbf{q}\cdot\mathbf{r}} \frac{\delta H_{\text{SR}}(\mathbf{A})}{\delta \mathbf{A}(\mathbf{r})} \\
&= \bar{e} \sum_i \frac{1}{2m} (\mathbf{p}_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} + e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{p}_i) + \frac{1}{8m^3 c^2} (\mathbf{p}_i |\mathbf{p}_i|^2 e^{-i\mathbf{q}\cdot\mathbf{x}_i} + \mathbf{p}_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} |\mathbf{p}_i|^2 + |\mathbf{p}_i|^2 e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{p}_i + e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{p}_i |\mathbf{p}_i|^2) \\
&= \bar{e} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{2m} (2\mathbf{p}_i - \mathbf{q}) + \frac{1}{8m^3 c^2} (|\mathbf{p}_i|^2 \mathbf{p}_i + |\mathbf{p}_i|^2 (\mathbf{p}_i - \mathbf{q}) + |\mathbf{p}_i - \mathbf{q}|^2 \mathbf{p}_i + |\mathbf{p}_i - \mathbf{q}|^2 (\mathbf{p}_i - \mathbf{q})) \right]. \quad (\text{B2})
\end{aligned}$$

We can alternatively use our Eq. (31) to obtain the current  $\mathbf{j}_{\mathbf{q}}$ . The single-particle velocity operator for  $H_{\text{SR}}$  is

$$\mathbf{v}_i = i[H_{\text{SR}}, \mathbf{x}_i] = \frac{1}{m} \mathbf{p}_i + \frac{1}{2m^3 c^2} |\mathbf{p}_i|^2 \mathbf{p}_i. \quad (\text{B3})$$

Inserting Eq. (B3) into Eq. (31), we find that our formalism for the current operator yields

$$\begin{aligned}
\mathbf{j}_{\mathbf{q}} &= \bar{e} \sum_i \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}_i} \left( \frac{1}{m} \mathbf{p}_i + \frac{1}{2m^3 c^2} |\mathbf{p}_i|^2 \mathbf{p}_i \right) e^{-i\lambda\mathbf{q}\cdot\mathbf{x}_i} \\
&= \bar{e} \sum_i \int_0^1 d\lambda e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{m} (\mathbf{p}_i - \lambda\mathbf{q}) + \frac{1}{2m^3 c^2} |\mathbf{p}_i - \lambda\mathbf{q}|^2 (\mathbf{p}_i - \lambda\mathbf{q}) \right]. \quad (\text{B4})
\end{aligned}$$

Finally, using Eqs. (32) and (33), we have that

$$\begin{aligned}
\tilde{\mathbf{j}}_{\mathbf{q}} &= \frac{\bar{e}}{2} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} (\mathbf{v}_i(\mathbf{p}_i, \mathbf{x}_i) + \mathbf{v}_i(\mathbf{p}_i - \mathbf{q}, \mathbf{x}_i)) \\
&= \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{2m} (2\mathbf{p}_i - \mathbf{q}) + \frac{1}{4m^3 c^2} (|\mathbf{p}_i|^2 + (\mathbf{p}_i - \mathbf{q})|\mathbf{p}_i - \mathbf{q}|^2) \right] \quad (\text{B5})
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{j}_{\text{mid}, \mathbf{q}} &= \bar{e} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{v}_i(\mathbf{p}_i - \mathbf{q}/2, \mathbf{x}_i) \\
&= \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{2m} (2\mathbf{p}_i - \mathbf{q}) + \frac{1}{16m^3 c^2} (2\mathbf{p}_i - \mathbf{q}) |2\mathbf{p}_i - \mathbf{q}|^2 \right]. \quad (\text{B6})
\end{aligned}$$

Let us examine the longitudinal and transverse components of the current using the three possible definitions: Eqs. (B2), (B4), and (B5). For the longitudinal current, we find

$$\begin{aligned}
\mathbf{q} \cdot \mathbf{j}_{\mathbf{q}, \min} &= \mathbf{q} \cdot \mathbf{j}_{\mathbf{q}} = \bar{e} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{2m} (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) + \frac{1}{8m^3 c^2} (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) (2|\mathbf{p}_i|^2 - 2\mathbf{p}_i \cdot \mathbf{q} + |\mathbf{q}|^2) \right], \quad (\text{B7})
\end{aligned}$$

$$\begin{aligned}
\mathbf{q} \cdot \tilde{\mathbf{j}}_{\mathbf{q}} &= \bar{e} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{2m} (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) + \frac{1}{4m^3 c^2} (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) (|\mathbf{p}_i|^2 - \mathbf{p}_i \cdot \mathbf{q} + |\mathbf{q}|^2) \right], \quad (\text{B8})
\end{aligned}$$

$$\begin{aligned}
\mathbf{q} \cdot \mathbf{j}_{\text{mid}, \mathbf{q}} &= \bar{e} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} \left[ \frac{1}{2m} (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) + \frac{1}{16m^3 c^2} (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) (4|\mathbf{p}_i|^2 - 4\mathbf{p}_i \cdot \mathbf{q} + |\mathbf{q}|^2) \right]. \quad (\text{B9})
\end{aligned}$$

Equation (B7) shows that the longitudinal component of the current  $\mathbf{j}_{\mathbf{q}}$ , as defined in Eq. (31), agrees with the longitudinal component of the current  $\mathbf{j}_{\mathbf{q}, \min}$  defined through minimal coupling via Eq. (21); both  $\mathbf{j}_{\mathbf{q}}$  and  $\mathbf{j}_{\mathbf{q}, \min}$  satisfy the continuity equation [Eq. (16)], as mentioned in the main text. The longitudinal components of  $\tilde{\mathbf{j}}_{\mathbf{q}}$  and  $\mathbf{j}_{\text{mid}, \mathbf{q}}$  from Eqs. (B8) and (B9) are distinct, implying that the (nonconserved)  $\tilde{\mathbf{j}}_{\mathbf{q}}$  and  $\mathbf{j}_{\text{mid}, \mathbf{q}}$  as defined in Eqs. (32) and (33) *do not* satisfy the continuity equation. Subtracting Eq. (B7) from Eqs. (B8) and (B9), we find

$$\begin{aligned}
\mathbf{q} \cdot (\tilde{\mathbf{j}}_{\mathbf{q}} - \mathbf{j}_{\mathbf{q}, \min}) &= \frac{\bar{e}}{8m^3 c^4} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} |\mathbf{q}|^2 (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) \\
&\neq 0 \quad (\text{B10})
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{q} \cdot (\mathbf{j}_{\text{mid}, \mathbf{q}} - \mathbf{j}_{\mathbf{q}, \min}) &= -\frac{\bar{e}}{16m^3 c^2} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} |\mathbf{q}|^2 (2\mathbf{p}_i \cdot \mathbf{q} - |\mathbf{q}|^2) \neq 0. \quad (\text{B11})
\end{aligned}$$

This explicitly shows that  $\tilde{\mathbf{j}}_{\mathbf{q}}$  and  $\mathbf{j}_{\text{mid}, \mathbf{q}}$ , as defined in Eqs. (32) and (33), are not conserved for the semirelativistic Hamiltonian from Eq. (B1).

Let us now examine the transverse components of the currents Eqs. (B2), (B4), and (B5). Taking the cross product with  $\mathbf{q}$ , we find that the minimally coupled current satisfies

$$\mathbf{q} \times \mathbf{j}_{\mathbf{q},\min} = \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{q} \times \mathbf{p}_i \left[ \frac{1}{m} + \frac{1}{4m^3 c^2} (|\mathbf{p}_i|^2 + |\mathbf{p}_i - \mathbf{q}|^2) \right]. \quad (\text{B12})$$

On the other hand, for our  $\lambda$  integral definition of the current from Eqs. (31) and (B4), we find

$$\begin{aligned} \mathbf{q} \times \mathbf{j}_{\mathbf{q}} &= \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{q} \times \mathbf{p}_i \left[ \frac{1}{m} + \frac{1}{2m^3 c^2} \int_0^1 d\lambda |\mathbf{p}_i - \lambda \mathbf{q}|^2 \right] \\ &= \mathbf{q} \times \mathbf{j}_{\mathbf{q},\min} + \frac{1}{3m^3 c^2} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} |\mathbf{q}|^2 (\mathbf{p}_i \times \mathbf{q}). \end{aligned} \quad (\text{B13})$$

This means that, although the current  $\mathbf{j}_{\mathbf{q}}$  is conserved, its transverse components differ from that of the minimally coupled current for the semirelativistic Hamiltonian in Eq. (B1), implying a different definition for the magnetization current.

On the other hand, the transverse current for the midpoint definition of the current Eq. (33) is

$$\begin{aligned} \mathbf{q} \times \mathbf{j}_{\text{mid},\mathbf{q}} &= \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{q} \times \left[ \frac{1}{2m} (2\mathbf{p}_i - \mathbf{q}) \right. \\ &\quad \left. + \frac{1}{16m^3 c^2} (2\mathbf{p}_i - \mathbf{q}) |2\mathbf{p}_i - \mathbf{q}|^2 \right] \\ &= \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{q} \times \mathbf{p}_i \left[ \frac{1}{m} + \frac{1}{4m^3 c^2} (|\mathbf{p}_i|^2 \right. \\ &\quad \left. + |\mathbf{p}_i - \mathbf{q}|^2 - \frac{1}{2} |\mathbf{q}|^2) \right] \\ &= \mathbf{q} \times \mathbf{j}_{\mathbf{q},\min} - \frac{1}{8m^3 c^2} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} |\mathbf{q}|^2 \mathbf{q} \times \mathbf{p}_i. \end{aligned} \quad (\text{B14})$$

Curiously, if we examine the transverse component of the nonconserved trapezoid current,  $\tilde{\mathbf{j}}_{\mathbf{q}}$ , from Eq. (32), we find that for the semirelativistic system

$$\begin{aligned} \mathbf{q} \times \tilde{\mathbf{j}}_{\mathbf{q}} &= \bar{e} \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{q} \times \left[ \frac{1}{2m} (2\mathbf{p}_i - \mathbf{q}) + \frac{1}{4m^3 c^2} (\mathbf{p}_i |\mathbf{p}_i|^2 \right. \\ &\quad \left. + (\mathbf{p}_i - \mathbf{q}) |\mathbf{p}_i - \mathbf{q}|^2) \right] \\ &= \sum_i e^{-i\mathbf{q} \cdot \mathbf{x}_i} \mathbf{q} \times \mathbf{p}_i \left[ \frac{1}{m} + \frac{1}{4m^3 c^2} (|\mathbf{p}_i|^2 + |\mathbf{p}_i - \mathbf{q}|^2) \right] \\ &= \mathbf{q} \times \mathbf{j}_{\mathbf{q},\min}. \end{aligned} \quad (\text{B15})$$

Thus, for the semirelativistic Hamiltonian, the transverse component of the nonconserved current  $\tilde{\mathbf{j}}_{\mathbf{q}}$  is equal to the transverse component of the minimally coupled current. This means that for the semirelativistic Hamiltonian, Eq. (B1), we have

$$\mathbf{j}_{\mathbf{q},\min}^\mu = \frac{q^\mu q^\nu}{|\mathbf{q}|^2} j_{\mathbf{q}}^\nu + \left( \delta_{\mu\nu} - \frac{q^\mu q^\nu}{|\mathbf{q}|^2} \right) \tilde{j}_{\mathbf{q}}^\nu. \quad (\text{B16})$$

In other words, for the semirelativistic Hamiltonian in Eq. (B1), the minimally coupled current in Eq. (22) can be written entirely in terms of the velocity operator via Eqs. (31) and (32). We could, if we desired a model-dependent formulation of the current, apply Eq. (B16) to the semirelativistic Hamiltonian which is applicable to heavy elements:

$$\begin{aligned} H_{\text{SOC}} &= \sum_i \frac{|\mathbf{p}_i|^2}{2m} + \frac{|\mathbf{p}_i|^4}{8m^3 c^2} + V(\mathbf{x}_i) + \frac{1}{8m^2 c^2} \\ &\quad + \frac{1}{4m^2 c^2} \vec{\sigma} \cdot (\mathbf{p}_i \times \nabla V(\mathbf{x}_i)) + \mathcal{O}(1/c^3), \end{aligned} \quad (\text{B17})$$

where  $\vec{\sigma}$  is a vector of Pauli matrices acting on electron spin;  $\vec{\sigma} \cdot (\mathbf{p}_i \times \nabla V(\mathbf{x}_i))$  represents the spin-orbit coupling energy.

If we can assume that our Hamiltonians under study have the form of  $H_{\text{SOC}}$  in Eq. (B17), then the current defined in Eq. (B16) entirely in terms of the velocity operator can be used to compute both longitudinal and transverse responses. If we cannot assume that Eq. (B17) holds for our system of interest, then we cannot determine the transverse component of the minimally coupled current entirely in terms of the velocity operator without detailed knowledge of the full microscopic Hamiltonian. In such a situation, the only guiding principle is that we define a conserved current, such as Eq. (31). In addition, we note that the nonrelativistic approximation to the Dirac equation that yields Eq. (B17) in the absence of an external electromagnetic field does *not* give a minimally coupled Hamiltonian in the presence of a nonzero vector potential [151]. Furthermore, Eq. (B16) fails when higher-order relativistic corrections are taken into account. To see this, we can consider a toy model with the next highest power in momentum according to the semirelativistic kinetic term:

$$H' = \frac{1}{6} \sum_i |\mathbf{p}_i|^6. \quad (\text{B18})$$

Minimally coupling  $H'$  to a vector potential via Eq. (22) gives the minimally coupled current

$$\begin{aligned}
\mathbf{j}_{\mathbf{q},\min} &= \frac{\bar{e}}{6} \sum_i [e^{-i\mathbf{q}\cdot\mathbf{x}_i} |\mathbf{p}_i|^4 + \mathbf{p}_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} |\mathbf{p}_i|^4 + |\mathbf{p}_i|^2 e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{p}_i |\mathbf{p}_i|^2 + |\mathbf{p}_i|^2 \mathbf{p}_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} |\mathbf{p}_i|^2 + |\mathbf{p}_i|^4 e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{p}_i + e^{-i\mathbf{q}\cdot\mathbf{x}_i} \mathbf{p}_i |\mathbf{p}_i|^4] \\
&= \frac{\bar{e}}{6} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} [|\mathbf{p}_i|^4 \mathbf{p}_i + |\mathbf{p}_i - \mathbf{q}|^4 (\mathbf{p}_i - \mathbf{q}) + |\mathbf{p}_i - \mathbf{q}|^2 \mathbf{p}_i |\mathbf{p}_i|^2 + |\mathbf{p}_i - \mathbf{q}|^2 (\mathbf{p}_i - \mathbf{q}) |\mathbf{p}_i|^2 + |\mathbf{p}_i - \mathbf{q}|^4 \mathbf{p}_i + |\mathbf{p}_i - \mathbf{q}|^4 (\mathbf{p}_i - \mathbf{q})].
\end{aligned} \tag{B19}$$

Similarly, applying Eqs. (32) to the Hamiltonian  $H'$  in Eq. (B18) results in the nonconserved current operator

$$\tilde{\mathbf{j}}_{\mathbf{q}} = \frac{\bar{e}}{2} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} [|\mathbf{p}_i|^4 \mathbf{p}_i + |\mathbf{p}_i - \mathbf{q}|^4 (\mathbf{p}_i - \mathbf{q})]. \tag{B20}$$

Decomposing Eqs. (B19) and (B20) into longitudinal and transverse components, we find

$$\mathbf{q} \cdot \tilde{\mathbf{j}}_{\mathbf{q}} \neq \mathbf{q} \cdot \mathbf{j}_{\mathbf{q},\min}, \tag{B21}$$

such that Eq. (B20) is not conserved. Importantly, using Eq. (B20), we also find that

$$\begin{aligned}
\mathbf{q} \times \tilde{\mathbf{j}}_{\mathbf{q}} &= \mathbf{q} \times \mathbf{j}_{\mathbf{q},\min} + \frac{1}{6} \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i} (\mathbf{q} \times \mathbf{p}_i) (|\mathbf{q}|^2 - 2\mathbf{q} \cdot \mathbf{p}_i)^2 \\
&\neq 0.
\end{aligned} \tag{B22}$$

Thus, for a general Hamiltonian, neither the longitudinal nor the transverse components of the conventional current  $\tilde{\mathbf{j}}_{\mathbf{q}}$  correctly reproduce the minimally coupled current  $\mathbf{j}_{\mathbf{q},\min}$ .

### APPENDIX C: CURRENT OPERATOR IN THE TIGHT-BINDING BASIS

In many cases, we are interested in operators and dynamics projected into a (tight-binding) basis of Loewdin orbitals  $\phi_{\alpha}(\mathbf{r} - \mathbf{R} - \mathbf{r}_{\alpha})$ . Let us introduce the Bloch-Loewdin basis functions

$$\chi_{\alpha\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R})} \phi_{\alpha}(\mathbf{r} - \mathbf{R} - \mathbf{r}_{\alpha}). \tag{C1}$$

We can equally well use  $|\chi_{\alpha\mathbf{k}}\rangle$  as a basis for expanding the current operators. Taking matrix elements of the conserved current from Eq. (31) in the basis of Eq. (C1), we find

$$\begin{aligned}
\mathbf{j}_{\mathbf{q}} &= \bar{e} \sum_{\mathbf{k}\alpha\beta\gamma\eta} \int_0^1 d\lambda \langle \chi_{\alpha\mathbf{k}} | e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} \mathbf{v} e^{-i\lambda\mathbf{q}\cdot\mathbf{x}} | \chi_{\beta\mathbf{k}+\mathbf{q}} \rangle c_{\alpha\mathbf{k}}^{\dagger} c_{\beta\mathbf{k}+\mathbf{q}} \\
&= \sum_{\mathbf{k}\mathbf{k}'\mathbf{k}''\alpha\beta\gamma\eta} \int_0^1 d\lambda \langle \chi_{\alpha\mathbf{k}} | e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} | \chi_{\gamma\mathbf{k}'} \rangle \\
&\quad \times \langle \chi_{\gamma\mathbf{k}'} | \mathbf{v} | \chi_{\eta\mathbf{k}''} \rangle \langle \chi_{\eta\mathbf{k}''} | e^{-i\lambda\mathbf{q}\cdot\mathbf{x}} | \chi_{\beta\mathbf{k}+\mathbf{q}} \rangle c_{\alpha\mathbf{k}}^{\dagger} c_{\beta\mathbf{k}+\mathbf{q}}.
\end{aligned} \tag{C2}$$

Introducing the cell-periodic basis functions

$$\tilde{\chi}_{\alpha\mathbf{k}}(\mathbf{r}) = \sqrt{N} e^{-i\mathbf{k}\cdot\mathbf{r}} \chi_{\alpha\mathbf{k}}(\mathbf{r}) \tag{C3}$$

and using Eq. (48), we observe that Eq. (C2) reduces to

$$\begin{aligned}
\mathbf{j}_{\mathbf{q}} &= \bar{e} \sum_{\mathbf{k}\alpha\beta\gamma\eta} \int_0^1 d\lambda \langle \tilde{\chi}_{\alpha\mathbf{k}} | \tilde{\chi}_{\gamma\mathbf{k}+(1-\lambda)\mathbf{q}} \rangle \\
&\quad \times \langle \tilde{\chi}_{\gamma\mathbf{k}+(1-\lambda)\mathbf{q}} | \mathbf{v} | \tilde{\chi}_{\eta\mathbf{k}+(1-\lambda)\mathbf{q}} \rangle \\
&\quad \times \langle \tilde{\chi}_{\eta\mathbf{k}+(1-\lambda)\mathbf{q}} | \tilde{\chi}_{\beta\mathbf{k}+\mathbf{q}} \rangle c_{\alpha\mathbf{k}}^{\dagger} c_{\beta\mathbf{k}+\mathbf{q}}.
\end{aligned} \tag{C4}$$

We can simplify Eq. (C4) by rewriting  $\langle \tilde{\chi}_{\gamma\mathbf{k}+(1-\lambda)\mathbf{q}} | \mathbf{v} | \tilde{\chi}_{\eta\mathbf{k}+(1-\lambda)\mathbf{q}} \rangle$  in terms of the Bloch Hamiltonian. First, note that

$$\begin{aligned}
\langle \chi_{\alpha\mathbf{k}} | \mathbf{v} | \chi_{\beta\mathbf{k}} \rangle &= \langle \tilde{\chi}_{\alpha\mathbf{k}} | \partial_{\mathbf{k}} H_{\mathbf{k}} | \tilde{\chi}_{\beta\mathbf{k}} \rangle \\
&= \partial_{\mathbf{k}} \langle \tilde{\chi}_{\alpha\mathbf{k}} | H_{\mathbf{k}} | \tilde{\chi}_{\beta\mathbf{k}} \rangle - \langle \partial_{\mathbf{k}} \tilde{\chi}_{\alpha\mathbf{k}} | H_{\mathbf{k}} | \tilde{\chi}_{\beta\mathbf{k}} \rangle \\
&\quad - \langle \tilde{\chi}_{\alpha\mathbf{k}} | H_{\mathbf{k}} | \partial_{\mathbf{k}} \tilde{\chi}_{\beta\mathbf{k}} \rangle \\
&= \partial_{\mathbf{k}} H_{\mathbf{k}}^{\alpha\beta} - i[\mathcal{A}_{\mathbf{k}}, H_{\mathbf{k}}]_{\alpha\beta},
\end{aligned} \tag{C5}$$

where we have introduced

$$H_{\mathbf{k}}^{\alpha\beta} = \langle \tilde{\chi}_{\alpha\mathbf{k}} | H_{\mathbf{k}} | \tilde{\chi}_{\beta\mathbf{k}} \rangle, \tag{C6}$$

$$\mathcal{A}_{\mathbf{k}}^{\alpha\beta} = i \langle \tilde{\chi}_{\alpha\mathbf{k}} | \partial_{\mathbf{k}} \tilde{\chi}_{\beta\mathbf{k}} \rangle, \tag{C7}$$

$$[\mathcal{A}_{\mathbf{k}}, H_{\mathbf{k}}] = \sum_{\gamma} \mathcal{A}_{\mathbf{k}}^{\alpha\gamma} H_{\mathbf{k}}^{\gamma\beta} - \mathcal{H}_{\mathbf{k}}^{\alpha\gamma} \mathcal{A}_{\mathbf{k}}^{\gamma\beta}. \tag{C8}$$

In the tight-binding limit where

$$\int d\mathbf{r} \phi_{\alpha}(\mathbf{r} - \mathbf{R} - \mathbf{r}_{\alpha}) \mathbf{r} \phi_{\beta}(\mathbf{r} - \mathbf{R}' - \mathbf{r}_{\beta}) = \mathbf{r}_{\alpha} \delta_{\alpha\beta} \delta_{\mathbf{R}\mathbf{R}'}, \tag{C9}$$

we can simplify Eq. (C5) further as

$$\begin{aligned}
\partial_{\mathbf{k}} H_{\mathbf{k}}^{\alpha\beta} - i[\mathcal{A}_{\mathbf{k}}, H_{\mathbf{k}}]_{\alpha\beta} &\rightarrow \partial_{\mathbf{k}} H^{\alpha\beta} - i\mathbf{r}_{\alpha} H_{\mathbf{k}}^{\alpha\beta} + iH_{\mathbf{k}}^{\alpha\beta} \mathbf{r}_{\beta} \\
&= e^{i\mathbf{k}\cdot\mathbf{r}_{\alpha}} \partial_{\mathbf{k}} \left( e^{-i\mathbf{k}\cdot\mathbf{r}_{\alpha}} H_{\mathbf{k}}^{\alpha\beta} e^{i\mathbf{k}\cdot\mathbf{r}_{\beta}} \right) e^{-i\mathbf{k}\cdot\mathbf{r}_{\beta}} \\
&= [V(\mathbf{k})(\partial_{\mathbf{k}} h_{\mathbf{k}}) V^{\dagger}(\mathbf{k})]_{\alpha\beta}, \tag{C10}
\end{aligned}$$

where

$$V^{\alpha\beta}(\mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{r}_{\alpha}} \delta_{\alpha\beta} \tag{C11}$$

is the tight-binding embedding matrix and

$$h_{\mathbf{k}}^{\alpha\beta} = \sum_{\gamma\eta} V_{\alpha\gamma}^{\dagger}(\mathbf{k}) H_{\mathbf{k}}^{\gamma\eta} V_{\eta\beta}(\mathbf{k}) \tag{C12}$$

is the matrix Bloch Hamiltonian in the “nonperiodic” gauge with  $h_{\mathbf{k}+\mathbf{G}} = V^{\dagger}(\mathbf{G}) h_{\mathbf{k}} V(\mathbf{G})$  for any reciprocal lattice vector  $\mathbf{G}$ .

Furthermore, in the tight-binding limit, we also have

$$\begin{aligned}
\langle \tilde{\chi}_{\alpha\mathbf{k}} | \tilde{\chi}_{\beta\mathbf{k}'} \rangle &= \sum_{\mathbf{R}\mathbf{R}'} \int_{\text{cell}} d\mathbf{r} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R})} e^{-i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{R}')} \\
&\quad \times \phi_{\alpha}^*(\mathbf{r}-\mathbf{R}-\mathbf{r}_{\alpha}) \phi_{\beta}(\mathbf{r}-\mathbf{R}'-\mathbf{r}_{\beta}) \\
&\rightarrow \delta_{\alpha\beta} \int d\mathbf{r} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} \phi_{\alpha}^*(\mathbf{r}-\mathbf{r}_{\alpha}) \phi_{\beta}(\mathbf{r}-\mathbf{r}_{\beta}) \\
&\rightarrow V^{\alpha\beta}(\mathbf{k}-\mathbf{k}'). \tag{C13}
\end{aligned}$$

Notice that, in going from the second to third line in Eq. (C13), we use an even stricter form of the tight-binding limit than in Eq. (C9): We require not just the first moment but *all* moments of the position operator to be diagonal in the tight-binding basis. Nevertheless, in this strict tight-binding limit we can combine Eqs. (C4), (C10), and (C13) to find

$$\begin{aligned}
\mathbf{j}_{\mathbf{q}} &\rightarrow \bar{e} \sum_{\mathbf{k}\alpha\beta} \int_0^1 d\lambda [V((\lambda-1)\mathbf{q}) V(\mathbf{k}+(1-\lambda)\mathbf{q}) \partial_{\mathbf{k}} h_{\mathbf{k}+(1-\lambda)\mathbf{q}} V^{\dagger}(\mathbf{k}+(1-\lambda)\mathbf{q}) V(-\lambda\mathbf{q})]_{\alpha\beta} c_{\alpha\mathbf{k}}^{\dagger} c_{\beta\mathbf{k}+\mathbf{q}} \\
&= \bar{e} \sum_{\mathbf{k}\alpha\beta} \int_0^1 d\lambda [V(\mathbf{k}) \partial_{\mathbf{k}} h_{\mathbf{k}+(1-\lambda)\mathbf{q}} V^{\dagger}(\mathbf{k}+\mathbf{q})] c_{\alpha\mathbf{k}}^{\dagger} c_{\beta\mathbf{k}+\mathbf{q}} \\
&= \bar{e} \sum_{\mathbf{k}\alpha\beta} \int_0^1 d\lambda (\partial_{\mathbf{k}} h_{\mathbf{k}+(1-\lambda)\mathbf{q}})^{\alpha\beta} \bar{c}_{\alpha\mathbf{k}}^{\dagger} \bar{c}_{\beta\mathbf{k}+\mathbf{q}}, \tag{C14}
\end{aligned}$$

where we have introduced

$$\bar{c}_{\alpha\mathbf{k}} = V_{\alpha\beta}^{\dagger}(\mathbf{k}) c_{\beta\mathbf{k}} \tag{C15}$$

as the annihilation operator for the nonperiodic basis states,

$$\begin{aligned}
\tilde{\chi}_{\alpha\mathbf{k}}(\mathbf{r}) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{r}_{\alpha})} \phi_{\alpha}(\mathbf{r}-\mathbf{R}-\mathbf{r}_{\alpha}) \\
&= V_{\alpha\beta}(\mathbf{k}) \chi_{\beta\mathbf{k}}(\mathbf{r}). \tag{C16}
\end{aligned}$$

Equation (C14) allows us to compute matrix elements of the conserved current  $\mathbf{j}_{\mathbf{q}}$  in the tight-binding basis entirely in terms of the tight-binding Bloch Hamiltonian  $h^{\alpha\beta}(\mathbf{k})$ . Note that Eq. (C14) could have been anticipated from Eqs. (C4) and (C5) by noting that the nonperiodic basis states in Eq. (C16) have vanishing Berry connection in the strict tight-binding limit.

Following a completely analogous set of steps, we find for the nonconserved trapezoid current in Eq. (32) that

$$\tilde{\mathbf{j}}_{\mathbf{q}} = \frac{\bar{e}}{2} \sum_{\mathbf{k}\alpha\beta} \left( \partial_{\mathbf{k}} h_{\mathbf{k}+\mathbf{q}}^{\alpha\beta} + \partial_{\mathbf{k}} h_{\mathbf{k}}^{\alpha\beta} \right) \bar{c}_{\alpha\mathbf{k}}^{\dagger} \bar{c}_{\beta\mathbf{k}+\mathbf{q}}, \tag{C17}$$

and similarly for the nonconserved midpoint current in Eq. (33):

$$\mathbf{j}_{\text{mid},\mathbf{q}} = \bar{e} \sum_{\mathbf{k}\alpha\beta} \partial_{\mathbf{k}} h_{\mathbf{k}+\mathbf{q}/2}^{\alpha\beta} \bar{c}_{\alpha\mathbf{k}}^{\dagger} \bar{c}_{\beta\mathbf{k}+\mathbf{q}}. \tag{C18}$$

Equations (C14), (C17), and (C18) serve as our starting point for computing linear and nonlinear response coefficients from tight-binding models in the main text.

#### APPENDIX D: GENERALIZED INTEGRATION FORMULATION OF AN OPERATOR USING THE KARPLUS-SCHWINGER RELATION

In this appendix, we apply the Karplus-Schwinger relation from Eq. (60) to the specific case of the density operator  $\rho_{\mathbf{q}}$  in order to derive an iterative formula for the diamagnetic current vertices.

First, consider the commutator between a general operator  $A$  and the exponential  $e^{-i\mathbf{q}\cdot\mathbf{x}}$  that appears in the Fourier component  $\rho_{\mathbf{q}}$ . By introducing an auxiliary parameter  $\tau$ , we can write

$$\begin{aligned}
i[A, e^{-i\mathbf{q}\cdot\mathbf{x}}] &= \frac{d}{d\tau} (e^{iA\tau} e^{-i\mathbf{q}\cdot\mathbf{x}} e^{-iA\tau}) \Big|_{\tau \rightarrow 0} \\
&= \frac{d}{d\tau} e^{e^{iA\tau} (-i\mathbf{q}\cdot\mathbf{x}) e^{-iA\tau}} \Big|_{\tau \rightarrow 0} \\
&= \frac{d}{d\tau} e^{-i\mathbf{q}\cdot\mathbf{x}(\tau)} \Big|_{\tau \rightarrow 0}, \tag{D1}
\end{aligned}$$

where we have defined

$$\mathbf{x}(\tau) \equiv e^{iA\tau} \mathbf{x} e^{-iA\tau}. \tag{D2}$$



From here, we may follow our derivation in Appendix A of the Karplus-Schwinger relation to find

$$\begin{aligned} i[A, e^{-i\mathbf{q}\cdot\mathbf{x}}] &= \lim_{\delta\tau \rightarrow 0} \frac{e^{-i\mathbf{q}\cdot\mathbf{x}(\tau+\delta\tau)} - e^{-i\mathbf{q}\cdot\mathbf{x}(\tau)}}{\delta\tau} \Big|_{\tau \rightarrow 0} \\ &\approx \lim_{\delta\tau \rightarrow 0} \frac{e^{-i\mathbf{q}\cdot(\mathbf{x}(\tau)+\delta\tau[A, \mathbf{x}])} - e^{-i\mathbf{q}\cdot\mathbf{x}(\tau)}}{\delta\tau} \Big|_{\tau \rightarrow 0} \\ &= \int_0^1 d\lambda e^{-i(1-\lambda)\mathbf{q}\cdot\mathbf{x}} (\mathbf{q} \cdot [A, \mathbf{x}]) e^{-i\lambda\mathbf{q}\cdot\mathbf{x}}. \end{aligned} \quad (\text{D3})$$

Equation (D3) allows us to rewrite commutators with the density operator that appear in the generalized Ward identity of Eq. (64) in terms of iterated integrals over auxiliary variables  $\lambda$ .

We can also use Eq. (D3) in conjunction with the tight-binding expression in Appendix C to express the generalized diamagnetic current operators in Sec. II C in the tight-binding basis. In keeping with notation from Appendix C, we let an overline indicate the operator in the orbital basis. We have

$$\begin{aligned} \bar{j}_{\mathbf{q}, -\mathbf{q}_1, \dots, -\mathbf{q}_N}^{\mu\nu_1 \dots \nu_N} &= \int_0^1 d\lambda \int_0^1 d\lambda_1 \dots \int_0^1 d\lambda_N \\ &\times [\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^{\nu_1}} \dots \partial_{\mathbf{k}^{\nu_N}} h_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + (1-\lambda)\mathbf{q} - (1-\lambda_1)\mathbf{q}_1 \dots - (1-\lambda_N)\mathbf{q}_N} \bar{c}_{n\mathbf{k}+\mathbf{q}-\mathbf{q}_1 \dots -\mathbf{q}_N}^\dagger \bar{c}_{m\mathbf{k}+\mathbf{q}-\mathbf{q}_1 \dots -\mathbf{q}_N}. \end{aligned} \quad (\text{D4})$$

Furthermore, in analogy with Eq. (66), we can define the tight-binding velocity vertex

$$\bar{v}_{(N)}^{\nu_N}(\mathbf{k}, \mathbf{q}_1, \dots, \mathbf{q}_N) \equiv [\partial_{\mathbf{k}^\mu} \partial_{\mathbf{k}^{\nu_1}} \dots \partial_{\mathbf{k}^{\nu_{N-1}}} h_{\mathbf{k}}]_{\mathbf{k} \rightarrow \mathbf{k} + (1-\lambda_1)\mathbf{q}_1 \dots + (1-\lambda_N)\mathbf{q}_N}. \quad (\text{D5})$$

Note that we can express any of the diagrams and subsequent conductivities in Sec. III B in either the orbital or tight-binding bases.

## APPENDIX E: THE KERR EFFECT DERIVATION

In this appendix, we present a derivation of the Kerr effect for systems with an anisotropic dielectric tensor. In Appendix E 1, we start by analyzing electromagnetic scattering from a 3D interface. Then, in Appendix E 2, we apply these results to the experimentally relevant situation of scattering from an encapsulated 2D sample considered in Sec. V B.

### 1. Kerr effect derivation in 3D materials

Consider an interface between vacuum and a 3D semi-infinite slab of material, with the interface normal to the  $z$  axis. We take the material dielectric tensor  $\epsilon$  to have the form [152]

$$\epsilon = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & 0 \\ -\epsilon_{xy} & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{zz} \end{bmatrix}. \quad (\text{E1})$$

We also make the approximation  $\mu \sim 1$ , which is tantamount to including all magnetic response in the transverse dielectric tensor [153–155].

We now consider an incident light beam with electric field of the form

$$\mathbf{E}_{\pm I} = E_0 R(\theta_i) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} e^{-i\omega t + ik_i[y \sin(\theta_i) + z \cos(\theta_i)]}. \quad (\text{E2})$$

Here,  $R(\theta_i)$  denotes a rotation matrix about the  $x$  axis by angle  $\theta_i$  [while this has no effect on  $x$ -polarized light such as in Eq. (E2), we retain it for ease of comparison with later steps in the derivation]. Notice that this electric field is linearly polarized and propagating in the direction of  $\theta_i$ , measured from the normal of the plane of the material. The wave vector  $k_i$  and frequency  $\omega$  are related later, and  $E_0$  is the real and positive amplitude.

We next consider the reflected field [155,156]:

$$\mathbf{E}_R = E_0 \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} e^{-i\omega t + ik_R[y \sin(\theta_R) + z \cos(\theta_R)]}. \quad (\text{E3})$$

Before considering the form of the transmitted field, we consider the differential equations it must satisfy. Consider Maxwell's equations in a material [154,157,158]:

$$\nabla \cdot \mathbf{D}_T = \rho_f, \quad (\text{E4})$$

$$\nabla \times \mathbf{H}_T = \mathbf{J}_f + \frac{\partial \mathbf{D}_T}{\partial t}, \quad (\text{E5})$$

$$\nabla \cdot \mathbf{B}_T = 0, \quad (\text{E6})$$

$$\nabla \times \mathbf{E}_T = -\frac{\partial \mathbf{B}_T}{\partial t}. \quad (\text{E7})$$

Here, with the assumption  $\mu \sim 1$ , then  $\mathbf{H}_T = \mathbf{B}_T$ . Also,  $\mathbf{D}_T = \epsilon \mathbf{E}_T$ , and we take  $\rho_f = \mathbf{J}_f = 0$ . We can now combine Eqs. (E5) and (E7) to obtain a wave equation:

$$\nabla(\nabla \cdot \mathbf{E}_T) - \nabla^2 \mathbf{E}_T = -\frac{\partial \mathbf{J}_f}{\partial t} - \frac{\partial^2 \epsilon \mathbf{E}_T}{\partial t^2}. \quad (\text{E8})$$

The fields that satisfy Eq. (E8) are of the form  $t_\ell \mathbf{E}_\ell e^{-i\omega + i\mathbf{k}_\ell \cdot \mathbf{r}}$ , where  $\mathbf{E}_\ell$  is the  $\ell$ th eigenvector of  $\epsilon$  and  $|\mathbf{k}_\ell|$  is proportional to the corresponding eigenvalue. Given the form of the dielectric tensor in Eq. (E1), we can solve the eigenvalue equation  $\epsilon \mathbf{E} = n^2 \mathbf{E}$  [159], to obtain the eigenvalues

$$\begin{aligned} n_p^2 &= \frac{1}{2} \left( \epsilon_{xx} + \epsilon_{yy} + \sqrt{(\epsilon_{xx} - \epsilon_{yy})^2 - 4\epsilon_{xy}^2} \right), \\ n_m^2 &= \frac{1}{2} \left( \epsilon_{xx} + \epsilon_{yy} - \sqrt{(\epsilon_{xx} - \epsilon_{yy})^2 - 4\epsilon_{xy}^2} \right), \\ n_3^2 &= \epsilon_{zz} \end{aligned} \quad (\text{E9})$$

with normalized (in units of  $E_0$ ) eigenvectors  $\mathbf{E}_p$ ,  $\mathbf{E}_m$ , and  $\mathbf{E}_3$ , respectively [160,161]. The meaning of  $n_\ell^2$  is the refractive index (squared) in the material.

Next, let  $\theta_{T\ell}$  be the angle of propagation of each eigenmode of the transmitted wave, measured with respect to the normal. Satisfying the wave equation implies  $|\mathbf{k}_\ell|^2 = n_\ell^2 \omega^2 / c^2$ . Note that, when we assign an angle, this also changes our eigenvalue equation to  $[R(\theta_{T\ell})\epsilon R^{-1}(\theta_{T\ell})] \times R(\theta_{T\ell})\mathbf{E}_\ell = n_\ell^2 R(\theta_{T\ell})\mathbf{E}_\ell$ , but the eigenvalues remain the same even if the eigenvectors are now angle dependent [155]. Given the required form of  $|\mathbf{k}_\ell|^2$  to be a solution to Eq. (E8), we can now write down the generalized form of the solution of the transmitted electric field [155,158,160]:

$$\begin{aligned} \mathbf{E}_T &= t_p \mathbf{E}_p e^{-i\omega t + i k_{T1}[y \sin(\theta_{T1}) + z \cos(\theta_{T1})]} \\ &+ t_m \mathbf{E}_m e^{-i\omega t + i k_{T2}[y \sin(\theta_{T2}) + z \cos(\theta_{T2})]} \\ &+ t_3 \mathbf{E}_3 e^{-i\omega t + i k_{T3}[y \sin(\theta_{T3}) + z \cos(\theta_{T3})]}, \end{aligned} \quad (\text{E10})$$

where  $t_\ell$  and  $\mathbf{E}_\ell$  are angle dependent.

To solve this electromagnetic scattering problem, we can first derive Snell's law for each of the transmitted modes. Snell's law states that the argument in the exponential of all the electric fields must be equal at the boundary. Therefore, we now have a set of four equations that results in

$$k_i \sin(\theta_i) = k_R \sin(\theta_R), \quad (\text{E11})$$

$$k_i \sin(\theta_i) = k_{T1} \sin(\theta_{T1}), \quad (\text{E12})$$

$$k_i \sin(\theta_i) = k_{T2} \sin(\theta_{T2}), \quad (\text{E13})$$

$$k_i \sin(\theta_i) = k_{T3} \sin(\theta_{T3}). \quad (\text{E14})$$

It should be noted by the nature of reflection that  $k_R = -k_i$  and  $\theta_R = -\theta_i$ .

We may now consider the remaining boundary conditions derived by requiring that Eqs. (E4)–(E7) are satisfied by components of the incident, reflected, and transmitted waves at the boundary. The Snell's law equations ensure the position-dependent arguments in the exponential are the same, guaranteeing the reflection and transmission coefficients are position (and time) independent. Following Ref. [157], the boundary equations are [156]

$$[\epsilon_{\text{vacuum}} \cdot (\mathbf{E}_{\pm i} + \mathbf{E}_R) - \mathbf{D}_T] \cdot \mathbf{n} = 0, \quad (\text{E15})$$

$$[\nabla \times (\mathbf{E}_{\pm i} + \mathbf{E}_R) - \nabla \times \mathbf{E}_T] \cdot \mathbf{n} = 0, \quad (\text{E16})$$

$$[\mathbf{E}_{\pm i} + \mathbf{E}_R - \mathbf{E}_T] \times \mathbf{n} = 0, \quad (\text{E17})$$

$$[\nabla \times (\mathbf{E}_{\pm i} + \mathbf{E}_R) - \nabla \times \mathbf{E}_T] \times \mathbf{n} = 0, \quad (\text{E18})$$

where  $\mathbf{n} = [0, 0, 1]$ , the normal to the 2D material plane. We also observe, as a good check, that  $\nabla \cdot \mathbf{E}_R = 0$  and  $\nabla \cdot \mathbf{E}_T = 0$  after resolving the coefficients through the boundary conditions. Solving the transmitted coefficients in this basis results in  $t_3 = 0$ .

To solve for the Kerr angles, we first rotate our local coordinate system to the frame of the reflected wave,  $R(\theta_R)\mathbf{E}_R$ . Then we redefine the coefficients as  $r_p = (r_x + ir_y)/2$  and  $r_m = (r_y - ir_x)/2$ , which describes the reflection coefficients for the right-hand and left-hand circular polarizations. Each of these coefficients is, in general, complex so we may break up each according to complex polar coordinates, that is,  $r_p = |r_p|e^{i\alpha_p}$  and  $r_m = |r_m|e^{i\alpha_m}$ . Kerr rotation results in a phase shift between right- and left-hand circularly polarized components of the reflected wave. Furthermore, the reflected wave can have an amplitude difference between right- and left-hand circularly polarized components. This allows us to define [22,152,154,161]

$$\frac{r_p}{r_m} = \frac{|r_p|}{|r_m|} e^{i(\alpha_p - \alpha_m)} \equiv \frac{|r_p|}{|r_m|} e^{i[2(\theta_K + i\epsilon_K)]}, \quad (\text{E19})$$

where  $\theta_K$  is the Kerr rotation angle and  $\epsilon_K$  is the ellipticity,

$$\tan(\epsilon_K) = \frac{|r_p| - |r_m|}{|r_p| + |r_m|}. \quad (\text{E20})$$

Our derivation expresses the Kerr angle and ellipticity in terms of components of the dielectric tensor. This can be related to the frequency and wave-vector-dependent conductivity using [152,154]

$$\epsilon = \mathbb{I} + \frac{4\pi i}{\omega} \sigma, \quad (\text{E21})$$

where  $\mathbb{I}$  is the  $3 \times 3$  identity matrix.

## 2. Kerr effect derivation in 2D materials with a substrate

In this section, we extend our analysis of Kerr effect to 2D materials. Most of the derivation remains the same, but there are a few subtle changes. In this setup, there is a vacuum, followed by a 2D thin film, and followed by a substrate which has the permittivity  $\epsilon_R$ . We take the  $\hat{z}$  axis to be normal to our 2D film, in the direction of the vacuum.

In this setup, the conductivity of the 2D materials now enters into Maxwell's equations as a boundary condition at the interface. Therefore, the propagation speed of the transmitted wave is not influenced by the film, only by the substrate in which it propagates. That is, the eigenvalue equation for the transmitted wave must satisfy  $n_R^2 \mathbb{I}_3 \mathbf{E}_{T,2D} = \epsilon_R \mathbb{I}_3 \mathbf{E}_{T,2D}$  [152,159]. Thus, we can write the transmitted wave as

$$\mathbf{E}_{T,2D} = R(\theta_T) \cdot (t_1 \mathbf{e}_1 + t_2 \mathbf{e}_2 + t_3 \mathbf{e}_3) \times e^{-i\omega t + i[R(\theta_T) \cdot (0,0,\sqrt{\epsilon_R} k_i)] \cdot \mathbf{r}}, \quad (\text{E22})$$

where  $\mathbf{e}_1 = (1, 0, 0)$ ,  $\mathbf{e}_2 = (0, 1, 0)$ , and  $\mathbf{e}_3 = (0, 0, 1)$  are the eigenmodes of  $\epsilon_R \mathbb{I}_3$ . The form of the incident and reflected waves remain the same, respectively:

$$\mathbf{E}_{\pm I,2D} = \frac{E_0}{\sqrt{2}} R(\theta_i) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} e^{-i\omega t + i[R(\theta_i) \cdot (0,0,k_i)] \cdot \mathbf{r}}, \quad (\text{E23})$$

$$\mathbf{E}_{R,2D} = R(-\theta_i) \cdot \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} e^{-i\omega t + i[R(-\theta_i) \cdot (0,0,-k_i)] \cdot \mathbf{r}}. \quad (\text{E24})$$

The Snell's law found by comparing the exponentials is also similar to before:

$$k_i \sin(\theta_i) = -k_i \sin(\theta_R), \quad (\text{E25})$$

$$\sin(\theta_i) = \sqrt{\epsilon_R} \sin(\theta_T). \quad (\text{E26})$$

Since we have a 2D material with nonzero conductivity tensor, a surface current is generated at the interface. As such, the boundary conditions get modified to include this surface current by way of the surface conductivity,  $\mathbf{j}_s = \sigma^{2D} \cdot \mathbf{E}_{T,2D}$ , which enters into Maxwell's equations as a free current on the surface. Since  $\nabla \cdot \mathbf{E}_{R,2D} = 0$  and  $\nabla \cdot \mathbf{E}_{T,2D} = 0$  force  $t_3 = r_z = 0$ , we arrive at four equations arising from two boundary conditions [153,162]:

$$[\mathbf{E}_{\pm I,2D} + \mathbf{E}_{R,2D} - \mathbf{E}_T] \times \mathbf{n} = 0, \quad (\text{E27})$$

$$\frac{c}{i\omega} [\nabla \times (\mathbf{E}_{\pm I,2D} + \mathbf{E}_{R,2D}) - \nabla \times \mathbf{E}_{T,2D}] \times \mathbf{n} = \sigma^{2D} \cdot \mathbf{E}_{T,2D}. \quad (\text{E28})$$

Since  $\sigma^{2D}$  is a property of the 2D sample, it takes the form

$$\sigma^{2D} = \begin{bmatrix} \sigma_{xx}^{2D} & \sigma_{xy}^{2D} & 0 \\ \sigma_{yx}^{2D} & \sigma_{yy}^{2D} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (\text{E29})$$

where the bulk conductivity is related to the surface conductivity by  $\sigma^{2D} = d\sigma^{3D}$  [152]. The Kerr rotation and ellipticity can be extracted using Eqs. (E19) and (E20), respectively.

## APPENDIX F: CONSISTENCY WITH PLASMON DISPERSION

In this appendix, we examine the implications of Eqs. (31) and (73) for the dispersion of collective plasmon modes. Since the plasmon dispersion depends on the  $\mathbf{q}$  dependence of the density-density calculation, it serves as a consistency check on Eqs. (31) and the sum rule from Eq. (97).

To begin, we treat the Coulomb interaction between electrons in the random phase approximation (RPA). That is, in writing the dielectric function, only the first bubble diagram is considered, such that

$$\epsilon(\omega, \mathbf{q}) = 1 - v_c(q) \Pi(\omega, \mathbf{q}). \quad (\text{F1})$$

Here,  $v_c(\mathbf{q})$  is the Fourier transformed Coulomb interaction [in 2D,  $v_c(q) = 2\pi e^2 / \kappa q$ , and  $v_c(q) = 4\pi e^2 / \kappa q^2$  in 3D] [163,164]. The function  $\Pi(\omega, \mathbf{q})$  is the polarizability or the electric susceptibility for noninteracting electrons. The electric susceptibility is related to the conductivity through the continuity equation. First, consider a longitudinal electric field  $\mathbf{E}_q = -i\mathbf{q}A_0(\omega, \mathbf{q})$ , defined in terms of a scalar potential  $A_0$ . The current that flows in response to this field is  $\mathbf{j}(\omega, \mathbf{q}) = -i\sigma(\omega, \mathbf{q}) \cdot \mathbf{q}A_0(\omega, \mathbf{q})$ . Additionally, from the continuity equation,  $-e\partial_t n(t, \mathbf{x}) + \nabla \cdot \mathbf{j}(t, \mathbf{r}) = 0$ . Expressing the density in terms of the response function  $\Pi$ , we find [165]

$$i\omega e^2 \Pi(\omega, \mathbf{q}) + \mathbf{q} \cdot \sigma(\omega, \mathbf{q}) \cdot \mathbf{q} = 0. \quad (\text{F2})$$

Now Eq. (F2) may be substituted into Eq. (F1) to obtain the relation

$$\epsilon(\omega, \mathbf{q}) = 1 + \frac{i}{\omega e^2} v_c(\mathbf{q}) (\mathbf{q} \cdot \sigma(\omega, \mathbf{q}) \cdot \mathbf{q}). \quad (\text{F3})$$

Note that, since our conserved current in Eq. (31) satisfies the continuity equation, our formalism ensures that Eq. (F3) is obeyed. To solve for the plasmon dispersion, we set Eq. (F3) equal to zero and solve for  $\omega(\mathbf{q})$ , the plasma frequency as a function of  $\mathbf{q}$ .

Since our conserved current Eq. (31) obeys the continuity equation, we can use the conductivity Eq. (73) to

determine the RPA plasmon dispersion in generic models. By contrast, since the midpoint and trapezoid definitions of the current [Eqs. (10) and (11)] do not obey the continuity equation at large  $\mathbf{q}$ , they do not give reliable estimates of the plasmon dispersion.

- 
- [1] P. A. Martin, *Sum rules in charged fluids*, *Rev. Mod. Phys.* **60**, 1075 (1988).
  - [2] P. Nozieres and D. Pines, *Electron interaction in solids. characteristic energy loss spectrum*, *Phys. Rev.* **113**, 1254 (1959).
  - [3] H. Watanabe and M. Oshikawa, *Generalized f-sum rules and Kohn formulas on nonlinear conductivities*, *Phys. Rev. B* **102**, 165137 (2020).
  - [4] M. Graf and P. Vogl, *Electromagnetic fields and dielectric response in empirical tight-binding theory*, *Phys. Rev. B* **51**, 4940 (1995).
  - [5] E. K. U. Gross and W. Kohn, *Local density-functional theory of frequency-dependent linear response*, *Phys. Rev. Lett.* **55**, 2850 (1985).
  - [6] E. K. U. Gross, J. F. Dobson, and M. Petersilka, *Density functional theory of time-dependent phenomena*, in *Density Functional Theory II: Relativistic and Time Dependent Extensions*, edited by R. F. Nalewajski (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996), pp. 81–172, 10.1007/BFb0016643.
  - [7] Y. Yao, L. Kleinman, A. H. MacDonald, J. Sinova, T. Jungwirth, D.-S. Wang, E. Wang, and Q. Niu, *First principles calculation of anomalous Hall conductivity in ferromagnetic bcc Fe*, *Phys. Rev. Lett.* **92**, 037204 (2004).
  - [8] T. D. Kühne, J. Heske, and E. Prodan, *Disordered crystals from first principles II: Transport coefficients*, *Ann. Phys. (Amsterdam)* **421**, 168290 (2020).
  - [9] M. Veithen, X. Gonze, and P. Ghosez, *Nonlinear optical susceptibilities, Raman efficiencies, and electro-optic tensors from first-principles density functional perturbation theory*, *Phys. Rev. B* **71**, 125107 (2005).
  - [10] S. Baroni, S. De Gironcoli, A. Dal Corso, and P. Giannozzi, *Phonons and related crystal properties from density-functional perturbation theory*, *Rev. Mod. Phys.* **73**, 515 (2001).
  - [11] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Quantized Hall conductance in a two-dimensional periodic potential*, *Phys. Rev. Lett.* **49**, 405 (1982).
  - [12] Q. Niu, D. J. Thouless, and Y.-S. Wu, *Quantized Hall conductance as a topological invariant*, *Phys. Rev. B* **31**, 3372 (1985).
  - [13] C. Xiao, Z. Z. Du, and Q. Niu, *Theory of nonlinear Hall effects: Modified semiclassics from quantum kinetics*, *Phys. Rev. B* **100**, 165422 (2019).
  - [14] I. Sodemann and L. Fu, *Quantum nonlinear Hall effect induced by Berry curvature dipole in time-reversal invariant materials*, *Phys. Rev. Lett.* **115**, 216806 (2015).
  - [15] F. De Juan, A. G. Grushin, T. Morimoto, and J. E. Moore, *Quantized circular photogalvanic effect in Weyl semimetals*, *Nat. Commun.* **8**, 15995 (2017).
  - [16] F. Flicker, F. de Juan, B. Bradlyn, T. Morimoto, M. G. Vergniory, and A. G. Grushin, *Chiral optical response of multifold fermions*, *Phys. Rev. B* **98**, 155145 (2018).
  - [17] A. Alexandradinata and P. Zhu, *Anomalous shift and optical vorticity in the steady photovoltaic current*, *arXiv:2308.08596*.
  - [18] A. M. Essin, J. E. Moore, and D. Vanderbilt, *Magneto-electric polarizability and axion electrodynamics in crystalline insulators*, *Phys. Rev. Lett.* **102**, 146805 (2009).
  - [19] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, *Topological field theory of time-reversal invariant insulators*, *Phys. Rev. B* **78**, 195424 (2008).
  - [20] L. Wu, M. Salehi, N. Koirala, J. Moon, S. Oh, and N. Armitage, *Quantized Faraday and Kerr rotation and axion electrodynamics of a 3D topological insulator*, *Science* **354**, 1124 (2016).
  - [21] V. Dziom, A. Shuvaev, A. Pimenov, G. Astakhov, C. Ames, K. Bendias, J. Böttcher, G. Tkachov, E. Hankiewicz, C. Brüne *et al.*, *Observation of the universal magnetoelectric effect in a 3D topological insulator*, *Nat. Commun.* **8**, 15197 (2017).
  - [22] J. Ahn, S.-Y. Xu, and A. Vishwanath, *Theory of optical axion electrodynamics and application to the Kerr effect in topological antiferromagnets*, *Nat. Commun.* **13**, 7615 (2022).
  - [23] H. B. Nielsen and M. Ninomiya, *The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal*, *Phys. Lett. B* **130**, 389 (1983).
  - [24] X. Huang, L. Zhao, Y. Long, P. Wang, D. Chen, Z. Yang, H. Liang, M. Xue, H. Weng, Z. Fang, X. Dai, and G. Chen, *Observation of the chiral-anomaly-induced negative magnetoresistance in 3D Weyl semimetal TaAs*, *Phys. Rev. X* **5**, 031023 (2015).
  - [25] M. Hirschberger, S. Kushwaha, Z. Wang, Q. Gibson, S. Liang, C. A. Belvin, B. A. Bernevig, R. J. Cava, and N. P. Ong, *The chiral anomaly and thermopower of Weyl fermions in the half-Heusler GdPtBi*, *Nat. Mater.* **15**, 1161 (2016).
  - [26] C.-L. Zhang, S.-Y. Xu, I. Belopolski, Z. Yuan, Z. Lin, B. Tong, G. Bian, N. Alidoust, C.-C. Lee, S.-M. Huang *et al.*, *Signatures of the Adler-Bell-Jackiw chiral anomaly in a Weyl fermion semimetal*, *Nat. Commun.* **7**, 1 (2016).
  - [27] J. Xiong, S. K. Kushwaha, T. Liang, J. W. Krizan, M. Hirschberger, W. Wang, R. J. Cava, and N. P. Ong, *Evidence for the chiral anomaly in the Dirac semimetal Na<sub>3</sub>Bi*, *Science* **350**, 413 (2015).
  - [28] D. N. Basov, R. D. Averitt, D. van der Marel, M. Dressel, and K. Haule, *Electrodynamics of correlated electron materials*, *Rev. Mod. Phys.* **83**, 471 (2011).
  - [29] D. E. Parker, T. Morimoto, J. Orenstein, and J. E. Moore, *Diagrammatic approach to nonlinear optical response with application to Weyl semimetals*, *Phys. Rev. B* **99**, 045121 (2019).
  - [30] J. E. Sipe and A. I. Shkrebtii, *Second-order optical response in semiconductors*, *Phys. Rev. B* **61**, 5337 (2000).
  - [31] J. L. Cheng, J. E. Sipe, and S. W. Wu, *Third order optical nonlinearity of three dimensional massless Dirac fermions*, *ACS Photonics* **7**, 2515 (2020).



- [32] C. Aversa and J. E. Sipe, *Nonlinear optical susceptibilities of semiconductors: Results with a length-gauge analysis*, *Phys. Rev. B* **52**, 14636 (1995).
- [33] S. M. João and J. M. V. P. Lopes, *Basis-independent spectral methods for non-linear optical response in arbitrary tight-binding models*, *J. Phys. Condens. Matter* **32**, 125901 (2019).
- [34] J. Ahn, G.-Y. Guo, N. Nagaosa, and A. Vishwanath, *Riemannian geometry of resonant optical responses*, *Nat. Phys.* **18**, 290 (2022).
- [35] A. Alexandradinata, *A topological principle for photovoltaics: Shift current in intrinsically polar insulators*, *arXiv:2203.11225*.
- [36] T. Holder, D. Kaplan, and B. Yan, *Consequences of time-reversal-symmetry breaking in the light-matter interaction: Berry curvature, quantum metric, and diabatic motion*, *Phys. Rev. Res.* **2**, 033100 (2020).
- [37] J. Orenstein, J. Moore, T. Morimoto, D. Torchinsky, J. Harter, and D. Hsieh, *Topology and symmetry of quantum materials via nonlinear optical responses*, *Annu. Rev. Condens. Matter Phys.* **12**, 247 (2021).
- [38] L. Wu, S. Patankar, T. Morimoto, N. L. Nair, E. Thewalt, A. Little, J. G. Analytis, J. E. Moore, and J. Orenstein, *Giant anisotropic nonlinear optical response in transition metal mononitride Weyl semimetals*, *Nat. Phys.* **13**, 350 (2017).
- [39] P. Streda, *Theory of quantised Hall conductivity in two dimensions*, *J. Phys. C* **15**, L717 (1982).
- [40] L. Smrcka and P. Streda, *Transport coefficients in strong magnetic fields*, *J. Phys. C* **10**, 2153 (1977).
- [41] A. Malashevich, I. Souza, S. Coh, and D. Vanderbilt, *Theory of orbital magnetoelectric response*, *New J. Phys.* **12**, 053032 (2010).
- [42] S. Zhong, J. E. Moore, and I. Souza, *Gyrotropic magnetic effect and the magnetic moment on the Fermi surface*, *Phys. Rev. Lett.* **116**, 077201 (2016).
- [43] S. Gassner and E. J. Mele, *Regularized lattice theory for spatially dispersive nonlinear optical conductivities*, *Phys. Rev. B* **108**, 085403 (2023).
- [44] C. Hoyos and D. T. Son, *Hall viscosity and electromagnetic response*, *Phys. Rev. Lett.* **108**, 066805 (2012).
- [45] B. Bradlyn, M. Goldstein, and N. Read, *Kubo formulas for viscosity: Hall viscosity, Ward identities, and the relation with conductivity*, *Phys. Rev. B* **86**, 245309 (2012).
- [46] T. Holder, R. Queiroz, and A. Stern, *Unified description of the classical Hall viscosity*, *Phys. Rev. Lett.* **123**, 106801 (2019).
- [47] A. Lucas and K. C. Fong, *Hydrodynamics of electrons in graphene*, *J. Phys. Condens. Matter* **30**, 053001 (2018).
- [48] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder *et al.*, *Visualizing Poiseuille flow of hydrodynamic electrons*, *Nature (London)* **576**, 75 (2019).
- [49] L. V. Delacrétaz and A. Gromov, *Transport signatures of the Hall viscosity*, *Phys. Rev. Lett.* **119**, 226602 (2017).
- [50] A. Zhang and J.-W. Rhim, *Geometric origin of intrinsic spin Hall effect in an inhomogeneous electric field*, *Commun. Phys.* **5**, 195 (2022).
- [51] V. Kozii, A. Avdoshkin, S. Zhong, and J. E. Moore, *Intrinsic anomalous Hall conductivity in a nonuniform electric field*, *Phys. Rev. Lett.* **126**, 156602 (2021).
- [52] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, *Unconventional superconductivity in magic-angle graphene superlattices*, *Nature (London)* **556**, 43 (2018).
- [53] R. Bistritzer and A. H. MacDonald, *Moiré butterflies in twisted bilayer graphene*, *Phys. Rev. B* **84**, 035440 (2011).
- [54] A. Ciarrocchi, F. Tagarelli, A. Avsar, and A. Kis, *Excitonic devices with van der Waals heterostructures: Valleytronics meets twistrionics*, *Nat. Rev. Mater.* **7**, 449 (2022).
- [55] E. Y. Andrei, D. K. Efetov, P. Jarillo-Herrero, A. H. MacDonald, K. F. Mak, T. Senthil, E. Tutuc, A. Yazdani, and A. F. Young, *The marvels of moiré materials*, *Nat. Rev. Mater.* **6**, 201 (2021).
- [56] P. Törmä, S. Peotta, and B. A. Bernevig, *Superconductivity, superfluidity and quantum geometry in twisted multilayer systems*, *Nat. Rev. Phys.* **4**, 528 (2022).
- [57] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, Y. Ronen, H. Yoo, D. Haei Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath *et al.*, *Tunable spin-polarized correlated states in twisted double bilayer graphene*, *Nature (London)* **583**, 221 (2020).
- [58] J. M. Park, Y. Cao, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, *Tunable strongly coupled superconductivity in magic-angle twisted trilayer graphene*, *Nature (London)* **590**, 249 (2021).
- [59] Y. Zhang, T. Liu, and L. Fu, *Electronic structures, charge transfer, and charge order in twisted transition metal dichalcogenide bilayers*, *Phys. Rev. B* **103**, 155142 (2021).
- [60] J. Sung, Y. Zhou, G. Scuri, V. Zólyomi, T. I. Andersen, H. Yoo, D. S. Wild, A. Y. Joe, R. J. Gelly, H. Heo *et al.*, *Broken mirror symmetry in excitonic response of reconstructed domains in twisted MoSe<sub>2</sub>/MoSe<sub>2</sub> bilayers*, *Nat. Nanotechnol.* **15**, 750 (2020).
- [61] P. Wang, G. Yu, Y. H. Kwan, Y. Jia, S. Lei, S. Klemenz, F. A. Cevallos, R. Singha, T. Devakul, K. Watanabe *et al.*, *One-dimensional Luttinger liquids in a two-dimensional moiré lattice*, *Nature (London)* **605**, 57 (2022).
- [62] J. Liu and X. Dai, *Anomalous Hall effect, magneto-optical properties, and nonlinear optical properties of twisted graphene systems*, *npj Comput. Mater.* **6**, 57 (2020).
- [63] P. Moon and M. Koshino, *Optical absorption in twisted bilayer graphene*, *Phys. Rev. B* **87**, 205404 (2013).
- [64] G. B. Ventura, D. J. Passos, J. M. B. Lopes dos Santos, J. M. Viana Parente Lopes, and N. M. R. Peres, *Gauge covariances and nonlinear optical responses*, *Phys. Rev. B* **96**, 035431 (2017).
- [65] D. J. Passos, G. B. Ventura, J. M. Viana Parente Lopes, J. M. B. Lopes dos Santos, and N. M. R. Peres, *Nonlinear optical responses of crystalline systems: Results from a velocity gauge analysis*, *Phys. Rev. B* **97**, 235446 (2018).
- [66] S. Zhong, *Linear and Nonlinear Electromagnetic Responses in Topological Semimetals* (University of California, Berkeley, 2018).
- [67] G. Vignale, *Orbital paramagnetism of electrons in a two-dimensional lattice*, *Phys. Rev. Lett.* **67**, 358 (1991).
- [68] Óscar Pozo Ocaña and I. Souza, *Multipole theory of optical spatial dispersion in crystals*, *SciPost Phys.* **14**, 118 (2023).

- [69] P. Blaha, K. Schwarz, P. Sorantin, and S. Trickey, *Full-potential, linearized augmented plane wave programs for crystalline systems*, *Comput. Phys. Commun.* **59**, 399 (1990).
- [70] J. Zhao, P. Mai, B. Bradlyn, and P. Phillips, *Failure of topological invariants in strongly correlated matter*, *Phys. Rev. Lett.* **131**, 106601 (2023).
- [71] R. Karplus and J. Schwinger, *A note on saturation in microwave spectroscopy*, *Phys. Rev.* **73**, 1020 (1948).
- [72] D. Vanderbilt, *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge University Press, Cambridge, England, 2018).
- [73] T. Thonhauser and D. Vanderbilt, *Insulator/Chern-insulator transition in the Haldane model*, *Phys. Rev. B* **74**, 235111 (2006).
- [74] D.J. Thouless, *Wannier functions for magnetic sub-bands*, *J. Phys. C* **17**, L325 (1984).
- [75] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Advanced Book Classics (Avalon, New York, 1995).
- [76] D. Kaplan, T. Holder, and B. Yan, *Unifying semiclassics and quantum perturbation theory at nonlinear order*, *SciPost Phys.* **14**, 082 (2023).
- [77] P. Coleman, *Introduction to Many-Body Physics* (Cambridge University Press, Cambridge, England, 2015).
- [78] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction*, 1st ed. (Oxford University Press, New York, 2004).
- [79] Y.-S. Ting, B.G.-G. Chen, R. Sohn, and D. Derbes, *Quantum Field Theory: Lectures by Sydney Coleman* (World Scientific, Singapore, 2018).
- [80] R.C. McKay and B. Bradlyn, *Optical response from charge-density waves in Weyl semimetals*, *Phys. Rev. B* **104**, 155120 (2021).
- [81] A. Altland and B.D. Simons, *Condensed Matter Field Theory* (Cambridge University Press, Cambridge, England, 2010).
- [82] E.W. Fenton and G.C. Psaltakis, *Infrared conductivity from spin and charge density waves: (TMTSF) 2X*, *Solid State Commun.* **47**, 767 (1983).
- [83] K. Takano, *Diagrammatical approach to functional integral method in a one-dimensional Peierls system*, *Prog. Theor. Phys.* **68**, 1 (1982).
- [84] Note that we can generalize this rule to allow for calculation of nonzero- $\mathbf{G}$  umklapp response in Eq. (71) by imposing momentum conservation only modulo a reciprocal lattice vector.
- [85] A. A. Husain, E. W. Huang, M. Mitran, M. S. Rak, S. I. Rubeck, X. Guo, H. Yang, C. Sow, Y. Maeno, B. Uchoa, T. C. Chiang, P. E. Batson, P. W. Phillips, and P. Abbamonte, *Pines' demon observed as a 3D acoustic plasmon in  $\text{Sr}_2\text{RuO}_4$* , *Nature (London)* **621**, 66 (2023).
- [86] L. P. Kadanoff and P. C. Martin, *Hydrodynamic equations and correlation functions*, *Ann. Phys. (N.Y.)* **24**, 419 (1963).
- [87] J. F. Mendez-Valderrama, D. Mao, and D. Chowdhury, *A theory for the low-energy optical sum-rule in moiré graphene*, *arXiv:2312.03819*.
- [88] H. Weyl, *Gravitation and the electron*, *Proc. Natl. Acad. Sci. U.S.A.* **15**, 323 (1929).
- [89] J. F. Steiner, A. V. Andreev, and D. A. Pesin, *Anomalous Hall effect in type-I Weyl metals*, *Phys. Rev. Lett.* **119**, 036601 (2017).
- [90] A. A. Burkov, *Anomalous Hall effect in Weyl metals*, *Phys. Rev. Lett.* **113**, 187202 (2014).
- [91] Z.-M. Huang, J. Zhou, and S.-Q. Shen, *Topological responses from chiral anomaly in multi-Weyl semimetals*, *Phys. Rev. B* **96**, 085201 (2017).
- [92] T. M. McCormick, I. Kimchi, and N. Trivedi, *Minimal models for topological Weyl semimetals*, *Phys. Rev. B* **95**, 075133 (2017).
- [93] T. M. McCormick, R. C. McKay, and N. Trivedi, *Semi-classical theory of anomalous transport in type-II topological Weyl semimetals*, *Phys. Rev. B* **96**, 235116 (2017).
- [94] K.-Y. Yang, Y.-M. Lu, and Y. Ran, *Quantum Hall effects in a Weyl semimetal: Possible application in pyrochlore iridates*, *Phys. Rev. B* **84**, 075129 (2011).
- [95] G. Sharma, P. Goswami, and S. Tewari, *Nernst and magnetothermal conductivity in a lattice model of Weyl fermions*, *Phys. Rev. B* **93**, 035116 (2016).
- [96] B. J. Wieder, K.-S. Lin, and B. Bradlyn, *Axionic band topology in inversion-symmetric Weyl-charge-density waves*, *Phys. Rev. Res.* **2**, 042010(R) (2020).
- [97] J. Cano, B. Bradlyn, and M. G. Vergniory, *Multifold nodal points in magnetic materials*, *APL Mater.* **7**, 101125 (2019).
- [98] C. Hoyos, *Hall viscosity, topological states and effective theories*, *Int. J. Mod. Phys. B* **28**, 1430007 (2014).
- [99] G. Sharma, *Topological Phenomena in Metals and Superconductors* (Clemson University, Clemson, SC, 2016).
- [100] Y. Messica, D. B. Gutman, and P. M. Ostrovsky, *Anomalous Hall effect in disordered Weyl semimetals*, *Phys. Rev. B* **108**, 045121 (2023).
- [101] M. Kanagaraj, J. Ning, and L. He, *Topological  $\text{Co}_3\text{Sn}_2\text{S}_2$  magnetic Weyl semimetal: From fundamental understanding to diverse fields of study*, *Rev. Phys.* **8**, 100072 (2022).
- [102] F. Buccheri, R. Egger, and A. De Martino, *Transport, refraction, and interface arcs in junctions of Weyl semimetals*, *Phys. Rev. B* **106**, 045413 (2022).
- [103] I. Osterloh, P. M. Oppeneer, J. Sticht, and J. Kubler, *A theoretical study of the magneto-optical Kerr effect in  $\text{FeX}$  ( $\text{X} = \text{Co, Ni, Pd, Pt}$ )*, *J. Phys. Condens. Matter* **6**, 285 (1994).
- [104] N. P. Armitage, *Constraints on Jones transmission matrices from time-reversal invariance and discrete spatial symmetries*, *Phys. Rev. B* **90**, 035135 (2014).
- [105] J. Xia, Y. Maeno, P. T. Beyersdorf, M. M. Fejer, and A. Kapitulnik, *High resolution polar Kerr effect measurements of  $\text{Sr}_2\text{RuO}_4$ : Evidence for broken time-reversal symmetry in the superconducting state*, *Phys. Rev. Lett.* **97**, 167002 (2006).
- [106] J. Xia, E. Schemm, G. Deutscher, S. Kivelson, D. Bonn, W. Hardy, R. Liang, W. Siemons, G. Koster, M. Fejer *et al.*, *Polar Kerr effect measurements of  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ : Evidence for broken symmetry near the pseudogap temperature*, *Phys. Rev. Lett.* **100**, 127002 (2008).

- [107] E. Schemm, W. Gannon, C. Wishne, W. Halperin, and A. Kapitulnik, *Observation of broken time-reversal symmetry in the heavy-fermion superconductor UPt<sub>3</sub>*, *Science* **345**, 190 (2014).
- [108] A. Kapitulnik, J. Xia, E. Schemm, and A. Palevski, *Polar Kerr effect as probe for time-reversal symmetry breaking in unconventional superconductors*, *New J. Phys.* **11**, 055060 (2009).
- [109] D. R. Saykin, C. Farhang, E. D. Kountz, D. Chen, B. R. Ortiz, C. Shekhar, C. Felser, S. D. Wilson, R. Thomale, J. Xia *et al.*, *High resolution polar Kerr effect studies of CsV<sub>3</sub>Sb<sub>5</sub>: Tests for time-reversal symmetry breaking below the charge-order transition*, *Phys. Rev. Lett.* **131**, 016901 (2023).
- [110] X. Gong, M. Kargarian, A. Stern, D. Yue, H. Zhou, X. Jin, V. M. Galitski, V. M. Yakovenko, and J. Xia, *Time-reversal symmetry-breaking superconductivity in epitaxial bismuth/nickel bilayers*, *Sci. Adv.* **3**, e1602579 (2017).
- [111] F. D. M. Haldane, *Model for a quantum Hall effect without Landau levels: Condensed-matter realization of the “parity anomaly”*, *Phys. Rev. Lett.* **61**, 2015 (1988).
- [112] M. Mannai and S. Haddad, *Strain tuned topology in the Haldane and the modified Haldane models*, *J. Phys. Condens. Matter* **32**, 225501 (2020).
- [113] A. Filusch and H. Fehske, *Singular flat bands in the modified Haldane-Dice model*, *Physica (Amsterdam)* **659B**, 414848 (2023).
- [114] A. R. Wright, *Realising Haldane’s vision for a Chern insulator in buckled lattices*, *Sci. Rep.* **3**, 2736 (2013).
- [115] H. C. Po, L. Zou, T. Senthil, and A. Vishwanath, *Faithful tight-binding models and fragile topology of magic-angle bilayer graphene*, *Phys. Rev. B* **99**, 195455 (2019).
- [116] L. Brey, T. Stauber, T. Slipchenko, and L. Martín-Moreno, *Plasmonic Dirac cone in twisted bilayer graphene*, *Phys. Rev. Lett.* **125**, 256804 (2020).
- [117] J. Kang and O. Vafek, *Symmetry, Maximally localized Wannier states, and a low-energy model for twisted bilayer graphene narrow bands*, *Phys. Rev. X* **8**, 031088 (2018).
- [118] F. M. Faulstich, K. D. Stubbs, Q. Zhu, T. Soejima, R. Dilip, H. Zhai, R. Kim, M. P. Zaletel, G. K.-L. Chan, and L. Lin, *Interacting models for twisted bilayer graphene: A quantum chemistry approach*, *Phys. Rev. B* **107**, 235123 (2023).
- [119] M. Choi, S. H. Lee, Y. Kim, S. Kang, J. Shin, M. Kwak, K.-Y. Kang, Y.-H. Lee, N. Park, and B. Min, *A terahertz metamaterial with unnaturally high refractive index*, *Nature (London)* **470**, 369 (2011).
- [120] J. M. Luttinger, *Theory of thermal transport coefficients*, *Phys. Rev.* **135**, A1505 (1964).
- [121] N. R. Cooper, B. I. Halperin, and I. M. Ruzin, *Thermoelectric response of an interacting two-dimensional electron gas in a quantizing magnetic field*, *Phys. Rev. B* **55**, 2344 (1997).
- [122] C. Scheibner, A. Souslov, D. Banerjee, P. Surowka, W. T. M. Irvine, and V. Vitelli, *Odd elasticity*, *Nat. Phys.* **16**, 475 (2020).
- [123] F. Mauri and S. G. Louie, *Magnetic susceptibility of insulators from first principles*, *Phys. Rev. Lett.* **76**, 4246 (1996).
- [124] J. V. Bustamante, N. J. Wu, C. Fermon, M. Pannetier-Lecoeur, T. Wakamura, K. Watanabe, T. Taniguchi, T. Pellegrin, A. Bernard, S. Daddinounou, V. Bouchiat, S. Guéron, M. Ferrier, G. Montambaux, and H. Bouchiat, *Detection of graphene’s divergent orbital diamagnetism at the dirac point*, *Science* **374**, 1399 (2021).
- [125] D. O. Oriekhov, V. P. Gusynin, and V. M. Loktev, *Orbital susceptibility of t-graphene: Interplay of high-order van Hove singularities and Dirac cones*, *Phys. Rev. B* **103**, 195104 (2021).
- [126] J. Gliozzi, M. Lin, and T. L. Hughes, *Orbital magnetic quadrupole moment in higher order topological phases*, *arXiv:2211.08438*.
- [127] D. Ceresoli, T. Thonhauser, D. Vanderbilt, and R. Resta, *Orbital magnetization in crystalline solids: Multi-band insulators, Chern insulators, and metals*, *Phys. Rev. B* **74**, 024408 (2006).
- [128] D. Seleznev and D. Vanderbilt, *Towards a theory of surface orbital magnetization*, *Phys. Rev. B* **107**, 115102 (2023).
- [129] D. Bohm, *Note on a theorem of Bloch concerning possible causes of superconductivity*, *Phys. Rev.* **75**, 502 (1949).
- [130] H. Watanabe, *Bloch theorem in the presence of an additional conserved charge*, *Phys. Rev. Res.* **4**, 013043 (2022).
- [131] C. Zeng, S. Nandy, and S. Tewari, *Nonlinear transport in Weyl semimetals induced by Berry curvature dipole*, *Phys. Rev. B* **103**, 245119 (2021).
- [132] D. Rees, K. Manna, B. Lu, T. Morimoto, H. Borrmann, C. Felser, J. Moore, D. H. Torchinsky, and J. Orenstein, *Helicity-dependent photocurrents in the chiral Weyl semimetal RhSi*, *Sci. Adv.* **6**, eaba0509 (2020).
- [133] S. Kim, Y. Lv, X.-Q. Sun, C. Zhao, N. Bielinski, A. Murzabekova, K. Qu, R. A. Duncan, Q. L. Nguyen, M. Trigo *et al.*, *Observation of a massive phason in a charge-density-wave insulator*, *Nat. Mater.* **22**, 429 (2023).
- [134] C. P. Weber, J. Orenstein, B. A. Bernevig, S.-C. Zhang, J. Stephens, and D. D. Awschalom, *Nondiffusive spin dynamics in a two-dimensional electron gas*, *Phys. Rev. Lett.* **98**, 076604 (2007).
- [135] J. D. Koralek, C. P. Weber, J. Orenstein, B. A. Bernevig, S.-C. Zhang, S. Mack, and D. Awschalom, *Emergence of the persistent spin helix in semiconductor quantum wells*, *Nature (London)* **458**, 610 (2009).
- [136] D. H. Torchinsky, F. Mahmood, A. T. Bollinger, I. Božović, and N. Gedik, *Fluctuating charge-density waves in a cuprate superconductor*, *Nat. Mater.* **12**, 387 (2013).
- [137] J. R. Rouxel, D. Fainozzi, R. Mankowsky, B. Rösner, G. Seniutinas, R. Mincigrucci, S. Catalini, L. Foglia, R. Cucini, F. Döring *et al.*, *Hard x-ray transient grating spectroscopy on bismuth germanate*, *Nat. Photonics* **15**, 499 (2021).
- [138] F. Bencivenga, R. Mincigrucci, F. Capotondi, L. Foglia, D. Naumenko, A. Maznev, E. Pedersoli, A. Simoncig, F. Caporaletti, V. Chiloyan *et al.*, *Nanoscale transient gratings excited and probed by extreme ultraviolet femtosecond pulses*, *Sci. Adv.* **5**, eaaw5805 (2019).
- [139] K. Takasan, T. Morimoto, J. Orenstein, and J. E. Moore, *Current-induced second harmonic generation in*



- inversion-symmetric Dirac and Weyl semimetals*, *Phys. Rev. B* **104**, L161202 (2021).
- [140] C. H. Lee, H. H. Yap, T. Tai, G. Xu, X. Zhang, and J. Gong, *Enhanced higher harmonic generation from nodal topology*, *Phys. Rev. B* **102**, 035138 (2020).
- [141] T. Cea, C. Castellani, and L. Benfatto, *Nonlinear optical effects and third-harmonic generation in superconductors: Cooper pairs versus Higgs mode contribution*, *Phys. Rev. B* **93**, 180507(R) (2016).
- [142] J. Gooth, B. Bradlyn, S. Honnali, C. Schindler, N. Kumar, J. Noky, Y. Qi, C. Shekhar, Y. Sun, Z. Wang, B. A. Bernevig, and C. Felser, *Axionic charge-density wave in the Weyl semimetal (TaSe<sub>4</sub>)<sub>2</sub>I*, *Nature (London)* **575**, 315 (2019).
- [143] W. Shi, B. J. Wieder, H. Meyerheim, Y. Sun, Y. Zhang, Y. Li, L. Shen, Y. Qi, L. Yang, J. Jena *et al.*, *A charge-density-wave topological semimetal*, *Nat. Phys.* **17**, 381 (2021).
- [144] L. J. P. Ament, M. van Veenendaal, T. P. Devereaux, J. P. Hill, and J. van den Brink, *Resonant inelastic x-ray scattering studies of elementary excitations*, *Rev. Mod. Phys.* **83**, 705 (2011).
- [145] C. Forsythe, X. Zhou, K. Watanabe, T. Taniguchi, A. Pasupathy, P. Moon, M. Koshino, P. Kim, and C. R. Dean, *Band structure engineering of 2D materials using patterned dielectric superlattices*, *Nat. Nanotechnol.* **13**, 566 (2018).
- [146] Y. Li, S. Dietrich, C. Forsythe, T. Taniguchi, K. Watanabe, P. Moon, and C. R. Dean, *Anisotropic band flattening in graphene with one-dimensional superlattices*, *Nat. Nanotechnol.* **16**, 525 (2021).
- [147] D. Barcons Ruiz, H. Herzig Sheinfux, R. Hoffmann, I. Torre, H. Agarwal, R. K. Kumar, L. Vistoli, T. Taniguchi, K. Watanabe, A. Bachtold *et al.*, *Engineering high quality graphene superlattices via ion milled ultra-thin etching masks*, *Nat. Commun.* **13**, 6926 (2022).
- [148] Y. Zeng, T. M. Wolf, C. Huang, N. Wei, S. A. A. Ghorashi, A. H. MacDonald, and J. Cano, *Gate-tunable topological phases in superlattice modulated bilayer graphene*, *arXiv:2401.04321*.
- [149] J. Sun, S. A. A. Ghorashi, K. Watanabe, T. Taniguchi, F. Camino, J. Cano, and X. Du, *Signature of correlated insulator in electric field controlled superlattice*, *arXiv:2306.06848*.
- [150] Sayed Ali Akbar Ghorashi, A. Dunbrack, A. Abouelkomsan, J. Sun, X. Du, and J. Cano, *Topological and stacked flat bands in bilayer graphene with a superlattice potential*, *Phys. Rev. Lett.* **130**, 196201 (2023).
- [151] A. Messiah, *Quantum Mechanics* (Courier Corporation, Chelmsford, MA, 2014).
- [152] G. Catarina, N. M. R. Peres, and J. Fernández-Rossier, *Magneto-optical Kerr effect in spin split two-dimensional massive Dirac materials*, *2D Mater.* **7**, 025011 (2020).
- [153] M. Kargarian, M. Randeria, and N. Trivedi, *Theory of Kerr and Faraday rotations and linear dichroism in topological Weyl semimetals*, *Sci. Rep.* **5**, 12683 (2015).
- [154] P. M. Oppeneer, *Theory of the Magneto-Optical Kerr Effect in Ferromagnetic Compounds* (Technische Universität Dresden, Dresden, 1999).
- [155] S. Ghosh, A. Sahoo, and S. Nandy, *Theoretical investigations on Kerr and Faraday rotations in topological multi-Weyl semimetals*, *SciPost Phys.* **15**, 133 (2023).
- [156] J. Creed, *Methods for sensitive detection of magneto optic Kerr effect*, Master of Science thesis, San Jose State University, San Jose, CA, 2020.
- [157] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).
- [158] A. Wolski, *Theory of electromagnetic fields*, *arXiv:1111.4354*.
- [159] H. Ebert, *Magneto-optical effects in transition metal systems*, *Rep. Prog. Phys.* **59**, 1665 (1996).
- [160] T. Higo, H. Man, D. B. Gopman, L. Wu, T. Koretsune, O. M. J. v. t. Erve, Y. P. Kabanov, D. Rees, Y. Li, M.-T. Suzuki, S. Patankar, M. Ikhlas, C. L. Chien, R. Arita, R. D. Shull, J. Orenstein, and S. Nakatsuji, *Large magneto-optical Kerr effect and imaging of magnetic octupole domains in an antiferromagnetic metal*, *Nat. Photonics* **12**, 73 (2018).
- [161] N. W. E. McGee, *The magneto-optical Kerr effect: Theory, measurement and application*, Master thesis, Eindhoven University of Technology, Eindhoven, NB, Netherlands, 1991.
- [162] M. Shah, A. Akbar, M. Sajid, and M. Sabieh Anwar, *Transitional Faraday and Kerr effect in hybridized topological insulator thin films*, *Opt. Mater. Express* **11**, 525 (2021).
- [163] E. H. Hwang and S. Das Sarma, *Dielectric function, screening, and plasmons in two-dimensional graphene*, *Phys. Rev. B* **75**, 205418 (2007).
- [164] K. N. Mishra, S. Ashraf, and A. C. Sharma, *Ground state properties of monolayer doped graphene*, *World J. Condens. Matter Phys.* **02**, 36 (2012).
- [165] V. Jani, J. Kolorenc, and V. Picka, *Density and current response functions in strongly disordered electron systems: Diffusion, electrical conductivity and Einstein relation*, *Eur. Phys. J. B* **35**, 77 (2003).