Consequences of time-reversal-symmetry breaking in the light-matter interaction: Berry curvature, quantum metric, and adiabatic motion

Tobias Holder^{✉,*}, Daniel Kaplan[✉], and Binghai Yan[†]

Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 7610001, Israel

(Received 25 November 2019; accepted 2 June 2020; published 20 July 2020)

Nonlinear optical response is well studied in the context of semiconductors and has gained a renaissance in studies of topological materials in the recent decade. So far it mainly deals with nonmagnetic materials and it is believed to root in the Berry curvature of the material band structure. In this work we revisit the general formalism for the second-order optical response and focus on the consequences of the time-reversal-symmetry ($\mathcal{T}$) breaking, by a diagrammatic approach. We have identified three physical mechanisms to generate a DC photocurrent, i.e., the Berry curvature, a term closely related to the quantum metric, and the adiabatic motion. All three effects can be understood intuitively from the anomalous acceleration. The first two terms are respectively the antisymmetric and symmetric parts of the quantum geometric tensor. The last term is due to the dynamical antilocalization that appears from the phase accumulation between time-reversed fermion loops. Additionally, we derive the semiclassical conductivity that includes both intra- and interband effects. We find that $\mathcal{T}$ breaking can lead to a greatly enhanced nonlinear anomalous Hall effect that is beyond the contribution by the Berry curvature dipole.

DOI: [10.1103/PhysRevResearch.2.033100](https://doi.org/10.1103/PhysRevResearch.2.033100)

I. INTRODUCTION

The bulk photovoltaic effect (BPVE) [1–3] refers to the generation of a DC current from a uniform material by irradiation with strong light. Since the early 1980s, its understanding has been gradually established as the second-order optical response [2,4–8] that is closely related to the Berry phase [9], an intrinsic quantity of the material band structure. Recently it gained renewed interest in topological Weyl semimetals (WSMs) [10,11], where the Berry phase is believed to generate the giant photocurrent and also the second harmonic response [12–20]. Thus far, research on BPVE focuses on nonmagnetic materials. However, recent theoretical [21] and experimental [22–24] works on magnetic systems reveal a distinct photocurrent and second harmonic generation which cannot be merely derived from the Berry phase. Therefore, we are motivated to reexamine the second-order response theory and investigate the effects of time-reversal-symmetry ($\mathcal{T}$) breaking.

A. Overview

The nonlinear current $\mathbf{j}$ created at the second order in the incident electric field $\mathbf{E}$ of light is defined as

$$\mathbf{j}^c(\omega; \omega_1, \omega_2) = \sum_{ab} \sigma^{abc}(\omega; \omega_1, \omega_2) E^a(\omega_1) E^b(\omega_2). \quad (1)$$

*tobias.holder@weizmann.ac.il

†binghai.yan@weizmann.ac.il

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.

A focal point will be in the following, the DC-current $\mathbf{j}(0; \omega, -\omega)$ which is created in response to irradiation of light with finite frequency ω . In a clean, gapped system with $\mathcal{T}$, the BPVE is usually phrased in the form of two phenomena, respectively termed shift current and injection current. Both effects require inversion symmetry breaking, otherwise the momentum-space integral of the response function vanishes. The shift current denotes the current that is proportional to the positional shift (shift vector expressed in the form of the Berry connection) of the electron charge in an interband process. The shift current has previously been employed to successfully describe the BPVE in many compounds [2,6,25–28]. Injection current is a stronger response, which is described by a constant current creation with rate $d\mathbf{j}/dt$ [5,29–35]. In systems with finite relaxation rates, dissipative processes will limit the current injection, thus rendering the nonlinear conductivity finite, with a magnitude proportional to the lifetime of the quasiparticles. In materials with $\mathcal{T}$, the injection current will occur only in response to circular polarized light and a linear polarization can only lead to a shift current. This distinction is lost in compounds which break $\mathcal{T}$, as pointed out by Ref. [21] very recently, where both shift and injection current might arise for either linear or circular polarized light. While the nonlinear currents created by the incident light usually lead to Joule heating in the sample, it is possible in experiment to subtract this effect. In accordance with previous literature, we therefore present results only for the electrical conductivities, and not for thermoelectric response.

Compared to the shift current generation, the intuitive understanding of current injection is less clear. One of the initial objectives of this work was to understand its physical origin. An important step was the realization that a single cone in a two-band Weyl model can produce injection current at a quantized rate, which is proportional to the topological charge of a Weyl cone [19]. However, it seems unlikely for this

identification of the injection current with topological charge to translate to the generic case. Also, an analogous procedure to express the shift current with the Berry curvature has not been reported. Given these findings, it is presently unclear to which extent the presence of Berry monopoles located at the Weyl nodes determine the BPVE in WSMs. Indeed, topologically trivial semimetals like graphene have zero Berry curvature and are still known to possess a whole assortment of unique transport properties, among them an abnormally large third-order conductivity, merely due to their nodal points in the band structure and the surrounding linear Dirac cone dispersion [36–38]. A positive statement about the topological origin of the large shift current observed for example in TaAs [12] thus requires an identification of this response with an observable related to the topological nature of the material.

Recently, an intriguing nonlinear anomalous Hall effect (AHE) [39–43] was derived in the semiclassical limit, which refers to intraband transitions in the small frequency case of the BPVE. It is attributed to the existence of the Berry curvature dipole in the semiclassical description of the second-order response [16,41] and was shortly after discovered in thin films of the WSM WTe₂ [44–47]. It is, however, unclear how the mechanisms for shift and injection currents, which are from interband transitions, are related to the semiclassical motion in terms of the Berry curvature dipole.

The present discussion also touches upon a practical question of how to properly model a nonlinear response in realistic quantum systems with impurities or at finite temperature. For example, there is a longstanding question how the canonical perturbation theory is related to results from the kinetic approach [1,48–50]. In the latter approach, the current response to linear polarized light features not only a shift current but also a ballistic current [51], which originates from the momentum space asymmetry of the distribution function. In particular, in the derivation of the ballistic current, the transport time enters twice, once in the numerator to account for the build-up of the asymmetry in the distribution function, and once in the denominator as a result of phonon corrections to the electron-photon vertex. As a result, the ballistic current emerges as an order-one effect which appears completely unaccounted for in an adiabatic treatment.

In perturbation theory, a large part of the discussion has been framed in the context of two gauge choices, the length gauge [5,6] and the velocity gauge [8,52]. These names are derived from the way the applied electric field is coupled to the Hamiltonian. In the length gauge, electric field enters via a dipole energy which induces a polarization, while in the velocity gauge it is included through the electromagnetic gauge potential via minimal substitution in the momentum. Both approaches were originally conceived as complementary descriptions, but in the clean limit they yield identical results and can be related by a time-dependent gauge transformation [7,53]. Of course, also for finite quasiparticle lifetimes the calculated optical response is always independent of the chosen gauge. However, depending on the choice of gauging, different choices for the relaxation rates have been proposed in the respective formulation [52,54]. The issue is that nonlinear optical response fundamentally involves interband processes which are difficult to model semiclassically. As such, there is some ambiguity with regards to the lifetimes which enter

the expressions. While we will not attempt to solve this issue comprehensively, we offer a way to discuss both the adiabatic and the diabatic motion which arises under optical driving in connection with the so-called quantum geometric tensor [55]. Since this object has a clear physical interpretation, it is possible to retrace which finite relaxation rates enter the expressions. Here we focus on the intrinsic mechanism. We leave the discussion of extrinsic mechanisms for a future work. For time-reversal invariant systems, this has been worked out recently [56–58]. We emphasize that scattering induced processes can be comparable in size to the response due to band structure effects [39,51].

In the following we proceed to systematically explore the second-order optical response where not only inversion symmetry is absent but also the $\mathcal{T}$ is broken. In particular, we derive formulas for the semiclassical response, shift current, and injection current in terms of observable, physically distinct processes, some of which lead to previously unknown types of photocurrents as long as $\mathcal{T}$ is absent. Our starting point is the treatment of the BPVE in the diagrammatic approach [8] which can also account for finite lifetimes and offers a more immediate interpretation of the different contributions to second-order response. Still, these expressions remain opaque to a semiclassical interpretation of the physical processes as described by the perturbation theory. Therefore, we propose to understand the nonlinear response by rewriting the second-order conductivity in terms of semiclassical accelerations, separately considering instantaneous and sequential multiphoton processes, as dictated by the diagrammatics.

B. Short summary of results

As the main results we demonstrate that the second-order optical response can be understood in terms of three distinct physical processes, each of them a part of the matrix elements of the anomalous quasiparticle acceleration in a Bloch band. While it is well known that the velocity matrix elements in band eigenbasis are not simply derivatives of the eigenenergies but contain an anomalous velocity originating from the nonzero Berry curvature [59], the anomalous acceleration has not yet been discussed at this same level of detail [26,60–62]. To be precise, consider a dispersion $\hbar\epsilon_m$ with band index m and eigenstate vectors $|m\rangle$, its momentum derivative v_{mn}^a for the spatial direction a , and the Berry connection $r_{mn}^a = \langle m|\partial_{k_a}|n\rangle$. The semiclassical center of mass motion of a wave packet with position $\mathbf{r}_{\text{CM}}$ and momentum $\mathbf{k}$ moving in band m subject to an electric field $\mathbf{E}$ has the equation of motion

$$\dot{\mathbf{r}}_{\text{CM}}^a = \sum_m \partial_{k_a} \epsilon_m - \frac{e}{\hbar} E^b \Omega_{mm}^{ba}, \quad (2)$$

where the linear-response form of the anomalous velocity $\frac{e}{\hbar} E^b \Omega_{mm}^{ba}$ is due to the Berry curvature $\Omega_{mm}^{ba} = \partial^b r_{mm}^a - \partial^a r_{mm}^b$ which is usually written in the form of a vector $\Omega^c = \epsilon^{abc} \Omega_{mm}^{ab}$. As a reminder, this result is most easily derived in the adiabatic approach by employing the identification $\partial_t = (-e/\hbar) E^a \partial_{k_a}$, which follows from the semiclassical Lagrangian in minimal coupling [9]. This means that the acceleration can be found by considering the expectation value $\ddot{\mathbf{r}}_{\text{CM}}^a = -E^b \partial_{k_b} \dot{\mathbf{r}}_{\text{CM}}^a$. However, this derivation of the acceleration is problematic if bilinear couplings to the electric field are to be

TABLE I. Overview of different contributions to second-order response in terms of acceleration processes. The four sources of nonlinear current are the Berry-type geometric contribution, the diabatic motion, the effect of the quantum metric, and the Drude-type term (density of states). Checkmarks in parentheses denote the new terms which only appear when the material lacks $\mathcal{T}$. The leading order term (Leading SC) and next-leading order term (Next-leading SC) represent the semiclassical (SC) expansion at finite frequency.

	Berry curvature	Diabatic	Quantum metric	Drude
Shift current	✓		✓	
Injection current		✓		
Leading SC	(✓)	(✓)		(✓)
Next-leading SC	✓	(✓)		

retained, as this requires the equations for the semiclassical motion to be correct up to second order. Indeed, Gao *et al.* [63] reported some time ago that to second order in applied fields there appears not only a polarization dependent renormalization of the dispersion $\tilde{\varepsilon}(\mathbf{E})$, but also a field dependent positional shift $\mathbf{a}'(\mathbf{E})$ of the semiclassical wave packet, both of which enter the equation of motion for a wave packet,

$$\dot{\mathbf{r}}_{\text{CM}} = \nabla \tilde{\varepsilon} - \dot{\mathbf{k}} \times (\boldsymbol{\Omega} + \nabla \times \mathbf{a}'). \quad (3)$$

However, the discussion at the time was mostly concerned with the AHE and did not resolve these modifications with respect to second-order optical response. We remedy this situation, but for the discussion at hand it turns out to be better to directly consider the matrix elements of the second momentum derivative $w_{mn}^{ab} = \langle m | \partial_{k_a} \partial_{k_b} H_0(\mathbf{k}) | n \rangle$ of the Hamiltonian $H_0(\mathbf{k})$,

$$w_{mn}^{ab} = i \sum_l (v_{ml}^a r_{ln}^b - r_{ml}^b v_{ln}^a) + i(v_{mn}^a - v_{nm}^a) r_{mn}^b + i(\varepsilon_m - \varepsilon_n) R_{mn}^{ab} + \delta_{mn} \partial_{k_a} \partial_{k_b} \varepsilon_m, \quad (4)$$

where the symbol $R_{mn}^{ab} = (\langle i \partial_{k_a} \partial_{k_b} m | n \rangle + \langle m | i \partial_{k_a} \partial_{k_b} n \rangle) / 2$ will be explained momentarily. In Eq. (4), the last term is the normal derivative (i.e., a Drude term), while the other three give rise to the anomalous acceleration. These features of the anomalous motion appear in the second-order optical response due to the close connection between the electric field and the derivative in time. However, a straightforward translation between anomalous motion and optical response is complicated by the fact that the perturbative response formulas intrinsically contain interband processes and thus off-diagonal matrix elements in the band basis. For this reason it went so far unnoticed that second-order response actually contains three types of anomalous contributions, unlike the linear response, where only the Berry curvature makes an appearance. The phenomenology is summarized in Table I.

(i) The first term in Eq. (4) is the well-studied adiabatic acceleration of the electron wave function due to the dipole interaction with the electric field. In the second-order response, this can always be formulated as a properly symmetrized triple product of velocity matrix elements, and it corresponds to the Berry curvature dipole term in the simplified semiclassical

FIG. 1. Second-order contributions to the optical DC response. Straight lines with arrows denote quasiparticles, a straight double line connects to a current operator, and the wavy lines are couplings to the electromagnetic field. Note that diagrams (I) and (II) and (III) and (IV) are variants of each other.

treatment. Importantly, this adiabatic acceleration appears in the shift current.

(ii) The second term is a diabatic effect, which can also be written as a velocity difference between occupied and unoccupied bands. This response type can be identified as a mismatch between the derivative $\partial_{k_j} v^i$ of the quasiparticle velocity v^i and the derivative $\partial_{k_j} r^i$ of the Berry connection r^i . Speaking in terms of Feynman diagrams, the phases accumulated on momentum reversed (i.e., time reversed) paths do not cancel with each other. This phenomenon is responsible for current injection, which can therefore be viewed as the result of *dynamical*, i.e., finite-frequency antilocalization between time-reversed loops with three legs (see Fig. 1). We also show that remnants of the diabatic motion appear in the shift current as a projective correction to the Berry dipole. This correction is reminiscent of the ballistic current as found in the kinetic approach.

(iii) The third term resembles what is known as the bare quantum metric. Somewhat surprisingly, it is impossible to write higher order conductivities only as a function of velocity matrix elements and the Berry curvature. This is manifest in the third type of anomalous acceleration, which is constituted by the symmetric part of the second derivative of the wave function,

$$R_{mn}^{ab} = \frac{1}{2} (\langle i \partial_{k_a} \partial_{k_b} m | n \rangle + \langle m | i \partial_{k_a} \partial_{k_b} n \rangle), \quad (5)$$

which is closely related but not identical to the gauge-invariant quantum metric g^{ab} [55,64–68]. In essence, second-order response contains information about both the symmetric and antisymmetric part of the quantum geometric tensor (QGT),

$$\begin{aligned} Q_{mn}^{ab} &= \langle \partial^a m | 1 - \mathcal{P}_{\text{GS}} | \partial^b n \rangle \\ &= g_{mn}^{ab} - \frac{i}{2} \Omega_{mn}^{ab}, \end{aligned} \quad (6)$$

where $\mathcal{P}_{\text{GS}}$ is the projector onto the ground state manifold. In comparison, the linear response coefficients receive a contribution from the Berry curvature Ω_{mn}^{ab} only. We emphasize that both the Berry curvature and the bare quantum metric enter the second-order conductivity not directly, but in a form resembling a momentum derivative of these quantities. This is also the reason why it is necessary to consider the off-diagonal matrix elements of the acceleration, and why it is helpful to intermittently consider gauge-covariant and not exclusively gauge-invariant quantities. This notwithstanding, our final expressions for the different contributions to the nonlinear conductivity are gauge invariant.

(iv) Finally, we derive the semiclassical conductivity in the absence of $\mathcal{T}$ by employing a canonical expansion of the conductivity in powers of the applied frequency (ω). The leading order term (ω^{-2}) appears only when breaking $\mathcal{T}$ and contains three of the above four contributions, as summarized in Table I. We find that nonlinear response contains terms of geometric origin exclusively in the next-leading order in frequency (ω^{-1}). As $\mathcal{T}$ is broken, this next-leading order term is not given only by the Berry curvature dipole, but contains an additional diabatic term due to the intrinsically dissipative nature of optical transitions.

In addition, we report a robust way to obtain explicit sum rules for a general tight-binding Hamiltonian required to translate between velocity gauge and length gauge, even if the system is imperfect and quasiparticle lifetimes are finite. Our derivation shows that the sum rules should be viewed as an expansion of higher order vertices in terms of velocities and band energies. This is in agreement with the general formulation of sum rules in clean systems, where they appear in terms of vanishing nested covariant derivatives of an observable, as mandated by gauge invariance [7]. However, the latter formulation is entirely formal, leaving the resolution of the derivatives of matrix elements untouched. It is precisely this point where additional lifetime dependencies appear, thus making the sum rules one possible source of disagreement between the length and the velocity gauge formalism. In particular, the often quoted resolution of the non-Abelian Berry connection in terms of band structure parameters as $r_{mn}^a = v_{mn}^a / i\epsilon_{mn}$ does not hold if the quasiparticle lifetime is finite. In the past, another point of confusion was the use of commutator identities which are only valid for an infinite number of bands [7,69]. For example, the relation

$$[r^a, v^b] = \frac{i\hbar}{m} \delta^{ab} \quad (7)$$

generally does not hold for a tight-binding Hamiltonian. However, this problem is unrelated to gauging, it originates from the reliance on the relation $p^a = v^a m$, which is only true for an isotropic dispersion. The diagrammatic formalism does not necessitate assumptions of this kind, therefore the sum rules

presented here do not suffer from these limitations. We find that the sum rules are only effective at reducing the zero frequency singularity of the second-order conductivity if the material is time-reversal symmetric.

The remaining sections are organized as follows. In Sec. II we introduce the general formalism for nonlinear optical response for a clean system, relying mostly on the formulation in the velocity gauge. The introduction of finite lifetimes can be intuitively implemented within the diagrammatic approach. We discuss the phenomenology of this approach and also compare to the equivalent procedure in the length gauge. In Sec. III we explain the various processes which contribute to the second-order response. Some time is spent on the semiclassical picture associated with current injection, and we derive an intuitive formulation of this phenomenon in terms of a phase difference accumulated on a closed path. We calculate the semiclassical conductivity including all multiband contributions in the absence of $\mathcal{T}$. In Sec. IV we illustrate our findings by numerical calculations on a concrete example of a WSM and a Dirac semimetal. This serves as a model for a system which has zero (Abelian) Berry curvature and yet exhibits a large BPVE.

II. PERTURBATION THEORY

We briefly recollect the basic ingredients needed to capture optical response when going beyond linear approximation. In the literature this was achieved either by calculating the electric susceptibility or more directly by evaluating the current-current correlator to second order in the electric fields. These approaches are also known under the terms length gauge and velocity gauge.

A. Adiabatic formalism

The current created by the BPVE originates from the intraband motion of Bloch states in response to the irradiation. The landmark result for this effect was derived in [5] and reads

$$\begin{aligned} \sigma_{\text{DC}}^{abc}(\omega) &= \frac{e^3}{\hbar^2 \omega'} \sum_{m,n} \int_{\mathbf{k}} \frac{f_{nm} r_{nm}^a r_{mn}^b \Delta_{mn}^c}{\epsilon_{mn} - \omega - i0^+} \bigg|_{\omega'=0} \\ &\quad - \frac{e^3}{\hbar^2} \sum_{m,n} \int_{\mathbf{k}} \frac{f_{nm} r_{mn}^a r_{nm,c}^b}{\epsilon_{mn} - \omega - i0^+} \\ &\quad + (a, \omega \leftrightarrow b, -\omega). \end{aligned} \quad (8)$$

Here m, n are band indices, $\epsilon_{mn} = \epsilon_m - \epsilon_n$ is the difference of the eigenvalues of the unperturbed Hamiltonian H_0 in units of frequency, i.e., $\epsilon_m(\mathbf{k}) = \hbar^{-1} \langle m | H_0(\mathbf{k}) | m \rangle$, the difference of Fermi factors is $f_{mn} = f_m - f_n$, and finally there is the difference of velocities $\Delta_{mn}^a = v_{mn}^a - v_{nn}^a$, with $v_{mn}^a = \langle m | \partial_{k_a} H_0(\mathbf{k}) | n \rangle$. The non-Abelian Berry connection is $r_{mn}^a = i \langle m | \partial_{k_a} n \rangle$ with the generalized derivative $r_{nm,c}^b = \partial_{k_c} r_{nm}^b - i(r_{nn}^c - r_{mm}^c) r_{nm}^b$, where $r_{nn}^c = i \langle n | \partial_{k_c} n \rangle$ is the Abelian Berry connection. The integration is over the first Brillouin zone, with $\int_{\mathbf{k}} \equiv \int_{\text{FBZ}} d^3 k / (2\pi)^3$. The current created through the bulk photovoltaic effect formally corresponds to an infinite polarizability, i.e., charge imbalances are continuously created. However, the first term in Eq. (8) for the current is infinite by itself, meaning that the second-order response

may even produce an infinite current, a phenomenon termed current injection, i.e., current is continuously created. Originally, Eq. (8) was derived from a painstaking second-order calculation of the adiabatic time evolution of the instantaneous eigenstates, also known as a reduced density matrix formalism. While this approach is conceptually very clear cut, it relies on two important factors. On the one hand, the optical processes must be fast enough for lifetime effects to be negligible. This assumption becomes questionable for optical transitions involving resonances with transient states in highly excited levels. In the present context, one would expect the current injection to strongly depend on lifetime effects. Second, it is often useful to resolve the derivative $r_{nm,c}^b$ by a sum rule. The latter task was recently re-examined in [27], finding for $n \neq m$

$$r_{nm,c}^b = \frac{v_{nm}^b(v_{nn}^c - v_{nm}^c) - v_{nm}^c(v_{nn}^b - v_{mm}^b)}{-i\varepsilon_{nm}^2} + \sum_{p \neq m,n} \left(\frac{v_{np}^b v_{pm}^c}{-i\varepsilon_{nm}\varepsilon_{pm}} - \frac{v_{np}^c v_{pm}^b}{-i\varepsilon_{nm}\varepsilon_{np}} \right) + \frac{w_{nm}^{bc}}{i\varepsilon_{nm}}. \quad (9)$$

Here the second derivative $w_{mn}^{ab} = \langle m | \partial_{k_a} \partial_{k_b} H_0(\mathbf{k}) | n \rangle$ makes an appearance. This term was originally not included because a quadratic Hamiltonian implies $w_{nm}^{ab} = 0$ for $n \neq m$. Generically w is only one of several higher order vertices which appear in an order-by-order analysis [8].

Properly including lifetime effects necessitates a closer look at the theoretical underpinnings of the adiabatic formalism. The starting point is a formulation of the problem in the length gauge. Equivalently it follows from the proposition that the response of the crystal can be approximated by a dipole Hamiltonian [5]

$$H_{\text{dip}}(t) = \sum_{n,m} \int_{\mathbf{k}} \hbar \bar{\varepsilon}_{nm}(\mathbf{k}, t) c_n^\dagger(\mathbf{k}) c_m(\mathbf{k}), \quad (10)$$

where the renormalized dispersion $\bar{\varepsilon}_{nm}(\mathbf{k}, t) = \varepsilon_n \delta_{nm} - e \hbar^{-1} r_{nm}^a E_a(t)$ contains the Berry connection r . While this construction encapsulates the polarization response in a clean system, it relies on the identification of the adiabatic intraband motion with the motion induced by the electromagnetic field [9]. Namely, the crystal momentum $\mathbf{q}$ in the isolated system and the gauge invariant momentum $\mathbf{k}$ are by definition related by $\mathbf{k} = \mathbf{q} + e\mathbf{A}(t)/\hbar$. This implies $\dot{\mathbf{k}} = -e\mathbf{E}(t)/\hbar$ under the condition that $\dot{\mathbf{q}} = 0$, i.e., if the crystal momentum is conserved.

B. Diagrammatic formulation

In the velocity gauge, one instead considers the coupling to the gauge field through minimal substitution of the momentum operator $\mathbf{k} \rightarrow \mathbf{k} - e\mathbf{A}/\hbar$ in the unperturbed Hamiltonian $H_0(\mathbf{k})$. Expanding in powers of $\mathbf{A}$ yields therefore

$$H_{\text{min}}(\mathbf{k}, t) = H_0(\mathbf{k}) + \frac{e}{\hbar} \sum_a \frac{\partial H_0(\mathbf{k})}{\partial k_a} A^a(t) + \frac{e^2}{2\hbar^2} \sum_{a,b} \frac{\partial^2 H_0(\mathbf{k})}{\partial k_a \partial k_b} A^a(t) A^b(t) + \frac{e^3}{6\hbar^3} \sum_{a,b,c} \frac{\partial^3 H_0(\mathbf{k})}{\partial k_a \partial k_b \partial k_c} A^a(t) A^b(t) A^c(t) + \dots \quad (11)$$

Here we wrote everything in terms of operators and not their matrix elements, which would result in a very different expression due to the nature of the (Berry) covariant derivative in $\mathbf{k}$ space, which is defined through the connection

$$\langle m | \partial_{k_a} O | n \rangle = \frac{\partial O_{mn}}{\partial k_a} - i r_{ml}^a O_{ln} + i O_{ml} r_{ln}^a, \quad (12)$$

where O is a Hermitian operator. This connection plays a major role in the calculation of Eq. (11). The current operator follows from a variation with respect to the electromagnetic field,

$$J^c = \frac{\delta H_{\text{min}}}{\delta A^c}. \quad (13)$$

Henceforth we will refer to derivatives in $\mathbf{k}$ space by $\partial^a = \partial/\partial k_a$. The various matrix elements are therefore $v_{mn}^a = \langle m | \partial^a H_0(\mathbf{k}) | n \rangle$, $w_{mn}^{ab} = \langle m | \partial^a \partial^b H_0(\mathbf{k}) | n \rangle$, and $\langle m | \partial^a \partial^b \partial^c H_0(\mathbf{k}) | n \rangle = v_{mn}^{abc}$. We note that using the connection, Eq. (12), the velocity operator $v^a = \partial^a H_0(\mathbf{k})$ can be expanded in the eigenbasis to yield the familiar expressions for the anomalous velocity including the Berry curvature term. The same holds for all higher order derivatives. We will defer this step until the discussion of the semiclassical motion.

In total, in second order the conductivity receives contributions from four types of diagrams, two of which have two variants. They are depicted in Fig. 1. Diagrams (I) and (II) correspond to the type of response originally considered in the earliest works [2,4]. Importantly, they only carry velocity matrix elements. Diagrams (III)–(VI) contain higher derivatives in $\mathbf{k}$, which in the original formulation of Eq. (8) appear through $\partial^c r_{nm}^b$, via sum rules. In the length gauge these higher order derivatives follow from an expansion of the time evolution of the instantaneous eigenstates beyond linear approximation [5,9].

The complete nonlinear conductivity for clean system reads

$$\sigma^{ab;c}(\bar{\omega}, \omega_1, \omega_2) = \frac{-e^3}{\hbar^2 \omega_1 \omega_2} \sum_{m,n,l} \int_{\mathbf{k}} f_m u_{mn}^{abc} + f_{mn} \frac{v_{mn}^a w_{nm}^{cb}}{\omega_1 + \varepsilon_{mn}} + f_{mn} \frac{v_{mn}^b w_{nm}^{ca}}{\omega_2 + \varepsilon_{mn}} + f_{mn} \frac{w_{mn}^{ab} v_{nm}^c}{\bar{\omega} + \varepsilon_{mn}} + \left(\frac{f_{mn} v_{mn}^a v_{nl}^b v_{lm}^c}{(\omega_1 - \varepsilon_{nm})(\bar{\omega} - \varepsilon_{lm})} + \frac{f_{mn} v_{ln}^a v_{nm}^b v_{ml}^c}{(\omega_2 - \varepsilon_{mn})(\bar{\omega} - \varepsilon_{ml})} + (a, \omega_1 \leftrightarrow b, \omega_2) \right), \quad (14)$$

where $\bar{\omega} = \omega_1 + \omega_2$. We point out that Eq. (14) differs from Ref. [8] by a sign for the band energy differences in diagrams (III)–(V) [70]. As it is visible from the diagrams, the vertices w and u describe instantaneous processes, where two or three interaction events coincide. In contradistinction, diagrams (I) and (II) describe sequential interactions. We repeat that Eq. (8) was computed from the dipole Hamiltonian and necessarily assumes the quasiparticle motion to be uninterrupted. Thus it cannot differentiate between instantaneous and successive two-photon processes. In the diagrammatics, these processes are easily distinguishable from their vertex structure. This

distinction can be useful in particular to tell apart effects of the anomalous velocity from the ones produced by higher derivatives.

C. Finite relaxation rates

It is possible to prove the equivalence of length and velocity gauge in the adiabatic limit [7]. On a formal level, the equivalence only requires a proper implementation of the time-dependent unitary transformation which translates between state vectors in scalar potential gauge (length gauge) and velocity gauge.

On the other hand, it is not obvious how to include finite relaxation rates into nonlinear response formulas. Dissimilar schemes have been proposed for length and velocity gauge, and in some cases these small changes in the inserted relaxation rates have led to markedly different results for the nonlinear response. One possibility is to retain different relaxation rates for the nonequilibrium response in each order of the electromagnetic field. Generically this will produce divergences at resonances due to a mismatch in rate differences [54,71]. A careful analysis of this issue was performed recently in Ref. [52]. In the reduced density matrix formalism a relaxation rate γ is introduced through the appended equation of motion for the density matrix ρ , where ρ_0 is the equilibrium state,

$$i\hbar\partial_t\rho = [H, \rho] - i\gamma(\rho - \rho_0). \quad (15)$$

Translated into velocity gauge this entails a relaxation of the system towards a modified equilibrium distribution $\rho(\mathbf{A})$ [7]. It is not obvious to which physical situation such a relaxation mechanism corresponds to. On the other hand, starting from diagrammatics, a commonly employed device is adiabatic switching, which regularizes the poles in the energy denominators by the replacement $\omega \rightarrow \omega + i\gamma$. Most importantly, this means inserting 2γ ($3\gamma, \dots$) in the formulas for second, (third, ...) harmonic generation whenever 2ω ($3\omega, \dots$) appears. Backporting this prescription to the length gauge, the authors in [52] present a case where a highly spiked resonance peak for third harmonic generation becomes noticeably regularized by just this replacement.

Why is it so difficult to establish a unique procedure? To demonstrate the difficulties associated with the standard treatment of second-order response, let us try to interpret the following second-order term written in the length gauge ($\varepsilon_{mn} = \varepsilon_m - \varepsilon_n$) [5]

$$\sum_n \partial_{k_c} \left(\frac{r_{mn}^a v_{nm}^b}{\omega - \varepsilon_{nm}} - \frac{v_{mn}^b r_{nm}^a}{\omega - \varepsilon_{mn}} \right). \quad (16)$$

Taking $\omega = 0$, this can be shown to assume the form of the Berry dipole $\partial^c \Omega_{mn}^{ab}$ [41]. However, performing the momentum derivative first, one obtains instead

$$\sum_n \left(\frac{(\partial_{k_c} r_{mn}^a) v_{nm}^b}{-\varepsilon_{nm}} + \frac{r_{mn}^a \partial_{k_c} v_{nm}^b}{-\varepsilon_{nm}} + \frac{r_{mn}^a v_{nm}^c v_{nm}^a}{(-\varepsilon_{nm})^2} - (m \leftrightarrow n) \right). \quad (17)$$

Looking at this, it remains unclear which physical process is at play, and the reduction to the Berry dipole by any means other than reverting the derivative is contrived. These issues hold for all such terms, the presence of derivatives in the length gauge

formulation make it hard to do power counting in ω and it is equally laborious to keep track of the proper gauge invariant combinations of terms [5]. The same issues persist for any finite imaginary parts.

Fundamental to the problem is the resolution of the velocity matrix elements in the eigenbasis, the defining relation which allows us to translate response formulas between length and velocity gauge. In the adiabatic limit, one employs $r_{mn} = v_{mn}^a / i\varepsilon_{mn}$ ($m \neq n$). The origin of this is the relation

$$v_{mn}^a = \partial^a \varepsilon_m \delta_{mn} - i[r^a, H_0]_{mn}, \quad (18)$$

which is invariably true—it follows from the mere fact that the covariant derivative defines a connection in momentum space [8,9]. However, for a small but nonzero relaxation rate Γ_{mn} for excited states, the commutator in this expression is only approximately diagonal in the eigenbasis, and yields

$$v_{mn}^a = \partial^a \varepsilon_m \delta_{mn} - i\varepsilon_{nm} r_{mn}^a - \sum_l (r_{ml}^a \Gamma_{ln} - \Gamma_{ml} r_{ln}^a) / 2. \quad (19)$$

Note that Γ_{mn} is symmetric. The off-diagonal velocity matrix elements are no longer just proportional to the Berry connection. Only in a simplified picture where the relaxation is happening purely between bands m and n , i.e., where the relaxation is intraband only, it holds for $m \neq n$ that $r_{mn}^a = v_{mn}^a / i[\varepsilon_{mn} + i(\Gamma_{mm} - \Gamma_{nn})]$.

For the reasons detailed above, we employ the orthodox view of renormalized perturbation theory to approach this issue in the diagrammatic formulation with velocity gauge. Namely, we analyze when the propagators are on-shell or off-shell in the respective diagrams, and use this to introduce relaxation rates according to the self-energies associated with low-energy or high-energy bands. The approach is physically transparent and does not rely on artificially setting all relaxation rates for all states to be the same. It also sidesteps the issues associated with quasiparticle decay in a density matrix formulation, instead relying on the Green's function formalism to implement finite lifetimes. More importantly, different relaxation rates appear depending on the type of process and not on the order in perturbation theory. This also means that second harmonic generation does not involve the same relaxation rates as DC-current generation. For example, the three-point loop depicted in Fig. 2 involves two consecutive photon processes, which means the quasiparticle partially propagates in highly excited bands. However, comparing the on-shell processes ($\Omega = 0$), for the photovoltaic effect ($\omega_1 = -\omega_2$) the quasiparticle propagators are dissimilar to the case of second harmonic generation ($\omega_1 = \omega_2$). It seems consequent to take this into account by using appropriate (i.e., dissimilar) relaxation rates. In the calculation, this reasoning calls for the insertion of a constant self-energy representative of the respective pole location where the residue is taken. We stress that such a procedure is insensitive to the analytic continuation to real frequencies as the actual self-energy is still only a constant, just not the same constant in all propagators. Further details of the relaxation time approximation are summarized in Appendix A.

In short, the presented phenomenology can be seen as the replacement $\omega \rightarrow \omega + i\gamma$, with the additional stipulation that the relaxation rate Γ of the intermediate states is allowed to be different from relaxation rate $\gamma \ll \Gamma$ of the long-lived

FIG. 2. Phenomenology of on-shell processes. Diagram (a) for the DC response contains two propagators at the same frequency, while the second harmonic generation shown in (c) involves propagators at three different frequencies. (b) and (d) Two-photon transitions of the electron in terms of energy levels for either case. In the language of on-shell/off-shell processes, only for diagram (a) [i.e., excitation and deexcitation as indicated in (b)] one encounters a pole which is restricted to intraband relaxation rates near the Fermi surface.

quasiparticles close to the Fermi surface. We emphasize that it is certainly possible to systematically include the very same phenomenology in the density matrix approach with length gauge, but it seems to lack the intuitive interpretation available in the diagrammatic analysis.

While we do not specify the microscopic nature of the relaxation rates, it is important to point out that γ is expected to be given by the transport time, and not the quasiparticle lifetime. This is related to the fact that whenever γ becomes important, the relevant processes happen at the Fermi surface where the usual arguments apply for the velocity vertex renormalization [72]. In contrast, Γ is an interband quantity which is unrelated to the scattering processes happening at the Fermi surface.

III. ANOMALOUS ACCELERATION

In order to intuitively understand second-order response in the absence of $\mathcal{T}$, one might want to fall back on a semiclassical treatment of the problem.

Disregarding lifetime effects and assuming $\mathcal{T}$, the semiclassical result for the nonlinear conductivity is stated as [8,26]

$$\sigma^{ab;c} = \frac{e^3}{\hbar^2} \int_k \sum_m f_m \left(-i \frac{\partial^b \Omega_{mm}^{ca}}{\omega} \right), \quad (20)$$

where the relaxation rate can be reintroduced after the fact by the replacement $\omega \rightarrow \omega + i\gamma$. The Abelian Berry curvature is defined as $\Omega_{mm}^{ab} = \partial^a r_{mm}^b - \partial^b r_{mm}^a$.

However, this phenomenology breaks down once the result from perturbation theory is considered. First, the usual semiclassical derivation cannot account for processes which are intrinsically finite- ω , but non-negligible after the momentum

space integral. Both the shift current and the injection current are of this type. Second, nonlinear response can involve interband transitions which are easily missed in the construction of the collision integral. The situation is complicated by the fact that finite lifetimes also modify the relation between Berry connection and velocity matrix elements according to Eq. (19), as we discussed before.

A. Vertices and sum rules

The calculation proceeds as follows. We first note that the derivative of the Berry connection is $\partial^b r_{mn}^c = R_{mn}^{bc} + i[r^b, r^c]_{mn}$, where $2R_{mn}^{bc} = \langle i\partial^b \partial^c m | n \rangle + \langle m | i\partial^b \partial^c n \rangle$, as mentioned. This matrix element can be related to the quantum metric through $2g_{mn}^{ab} = 2iR_{mn}^{ab} + 2\langle m | \partial^a \partial^b n \rangle - \sum_l f_l (r_{ml}^a r_{ln}^b + r_{ml}^b r_{ln}^a)$. The Berry dipole is given by (cf. Appendix B)

$$\partial^a \Omega_{mn}^{bc} = -\frac{1}{2} [r^a, [r^b, r^c]] + i[r^b, R^{ac}] - i[r^c, R^{ab}], \quad (21)$$

where all matrix elements carry two band indices which are circled through as dictated by the commutators (i.e., $[a, b]_{mn} = \sum_l (a_{ml} b_{ln} - b_{ml} a_{ln})$). We will henceforth suppress band indices whenever possible. Furthermore, for interband terms it holds that $i\partial^c v^b = -\varepsilon \partial^c r^b - \Delta^c v^b + O(\Gamma)$. These relations, while not exact for finite relaxation rates, are perturbative in relaxation rate corrections and do not introduce any shifted poles.

Sometimes it is suggested to use the substitution $r_{mn}^a \rightarrow i v_{mn}^a / \varepsilon_{nm}$. However, as explained earlier, such is only possible in a clean system. This is noteworthy since the decomposition of w_{mn}^{ab} creates, among others, terms which are very similar to diagrams (I) and (II). Going forward, we will only employ the opposite route of identification via $v_{mn}^a \rightarrow i \varepsilon_{mn} r_{mn}^a + O(\Gamma)$, which always remains viable. This allows us to immediately and transparently characterize the role of sum rules for systems with finite lifetimes and to establish a unique procedure to translate between the different gauges.

We rewrite w^{ab} and u^{abc} by expanding the derivatives,

$$\begin{aligned} w_{mn}^{ab} &= \partial^a v^b - i[r^a, v^b] \quad \text{for } m \neq n \\ &= \frac{1}{2} i([v^a, r^b] + \Delta^a r^b + \varepsilon R^{ab} + a \leftrightarrow b) + O(\Gamma), \quad (22) \\ u_{mm}^{abc} &= \partial^a (\partial^b v^c - i[r^b, v^c]) - i[r^a, \partial^b v^c - i[r^b, v^c]] \\ &= \partial^a \partial^b v^c + i[v^c, R^{ab}] + i[v^b, R^{ca}] + i[v^a, R^{bc}] \\ &\quad + [r^b, \Delta^a r^c] + [r^a, \Delta^b r^c] - [r^a, [r^b, v^c]] + O(\Gamma). \quad (23) \end{aligned}$$

The decomposition of the vertices might still appear pretty verbose, but crucially it will allow for a straightforward identification of the physical processes. Independent of this, it makes manifest that relaxation rates are accounted for perturbatively in the velocity expansion of the higher order vertices. As an aside, note that only the diagonal component of u^{abc} was analyzed since this is the only manner in which it appears here. The off-diagonal elements of u^{abc} needed for higher orders follows straightforwardly by continuing the expansion of $\partial^a \partial^b v^c$ for off-diagonal matrix elements. Keeping the shorthand notation, the second-order response is

generally

$$\begin{aligned} \sigma^{ab;c} = & \frac{-e^3}{\hbar^2 \omega_1 \omega_2} \int_k \sum_m f_m \left(u^{abc} + \left[w^{ab}, \frac{v^c}{\bar{\omega} - \varepsilon + i\gamma} \right] \right. \\ & + \left[\frac{v^a}{\omega_1 + \varepsilon + i\Gamma}, w^{bc} \right] + \left[\frac{v^b}{\omega_2 + \varepsilon + i\Gamma}, w^{ca} \right] \\ & + \left[\left[v^b, \frac{v^c}{\varepsilon - \bar{\omega} - i\gamma} \right], \frac{v^a}{\omega_1 + \varepsilon + i\Gamma} \right] \\ & \left. + \left[\left[v^a, \frac{v^c}{\varepsilon - \bar{\omega} - i\gamma} \right], \frac{v^b}{\omega_2 + \varepsilon + i\Gamma} \right] \right)_{mm}. \quad (24) \end{aligned}$$

While all top level commutators carry the indices mm , they do not evaluate to zero since the Fermi-Dirac factor f_m is not incorporated, i.e., the expression is not a trace. Nevertheless, thanks to the Fermi factor any top level commutators actually evaluate to zero for the diagonal pieces where a denominator $\varepsilon = 0$ could otherwise lead to issues. This allows the replacement of velocity matrix elements by the Berry connection in outer commutators and in particular renders the $i\gamma$ in the first line harmless. In principle, the frequencies ω_1 and ω_2 in front of the integral also acquire an imaginary part according to adiabatic switching ($\omega \rightarrow \omega + i\gamma$), but this replacement does not lead to any further complications and will henceforth be implicitly assumed for clarity of presentation.

The usual simplifications for a gapped and time-reversal symmetric system resulting in Eq. (20) are implemented using the sum rules Eqs. (22) and (23), which present the proper extensions of the classical sum rules for nonquadratic Hamiltonians in systems with finite quasiparticle lifetimes [5,7,69]. $\mathcal{T}$ dictates that velocity matrix elements fulfill $v_{mn}^a(\mathbf{k}) = -v_{mn}^a(-\mathbf{k})^*$, while the non-Abelian Berry connection transforms like $r_{mn}^a(\mathbf{k}) = r_{mn}^a(-\mathbf{k})^*$, and $R_{mn}^{ab}(\mathbf{k}) = -R_{mn}^{ab}(-\mathbf{k})^*$. For a gapped system the DC conductivity follows from the simultaneous expansion in small $\bar{\omega} \ll \omega_1, \omega_2$; we will also keep track of finite lifetimes whenever necessary. The integrand of the leading order term of size ω^{-2} is explicitly ($\delta_{mn}^c = r_{mm}^c - r_{nn}^c$)

$$\begin{aligned} u^{abc} - i[w^{ab}, r^c] + i[r^a, w^{bc}] + i[r^b, w^{ca}] \\ - [[v^a, r^c], r^b] - [[v^b, r^c], r^a] - [v^a \delta^c, r^b] - [v^b \delta^c, r^a] \\ + \frac{1}{\gamma} [v^a \Delta^c, r^b] + \frac{1}{\gamma} [v^b \Delta^c, r^a]. \quad (25) \end{aligned}$$

Here we used that $[v^a, v^c/(\varepsilon - i\gamma)] \approx i[v^a, r^c] + iv^a \delta^c - iv^a \Delta^c/\gamma$. We reiterate that sum rules basically constitute expansions of the vertices w and u in terms of acceleration (w) and jerk (u). Accordingly, the sum rules contain commutators with exactly the required number of odd components under momentum inversion. For example, it is

$$[v^c(-\mathbf{k}), R^{ab}(-\mathbf{k})] = (-1)^2 [v^c(\mathbf{k})^*, R^{ab}(\mathbf{k})^*] \quad (26)$$

$$= -[v^c(\mathbf{k}), R^{ab}(\mathbf{k})], \quad (27)$$

$$\begin{aligned} [r^a(-\mathbf{k}), \Delta^c(-\mathbf{k})r^b(-\mathbf{k})] &= -[r^a(\mathbf{k})^*, \Delta^c(\mathbf{k})r^b(\mathbf{k})^*] \\ &= -[r^a(\mathbf{k}), \Delta^c(\mathbf{k})r^b(\mathbf{k})]. \quad (28) \end{aligned}$$

Unlike in linear response, the Drude-type term $\partial^a \partial^b v^c$ is odd under momentum inversion and equally cancels. In essence,

in the presence of $\mathcal{T}$, higher order vertices do not contribute at leading order in ω to $\sigma^{ab;c}$. However, from the definition of the vertices ($w_{mn}^{ab} = \langle m | \partial^a \partial^b H | n \rangle$ and $u_{mn}^{abc} = \langle m | \partial^a \partial^b \partial^c H | n \rangle$), the same conclusion can be drawn immediately from their symmetry properties under momentum inversion, thus sidestepping the need for sum rules.

B. Finite frequency terms

The last two lines in Eq. (24) require more care as they contain a relaxation rate in the nested denominator, where a finite γ is not zeroed out *a priori*. While Eq. (25) suggests that both terms proportional to γ^{-1} cancel each other, this is not necessarily the case as the limit $\omega \rightarrow 0$ does neither commute with the k -space integration nor with the limit $\Gamma \rightarrow 0$. Thus, while Eq. (25) correctly captures the response at $\omega = 0$, it misses not only the injection current, but also the shift current, both are by construction finite frequency responses. Their incorporation is of course straightforward by keeping all imaginary parts in Eq. (24). Denominators containing Γ have a nonzero imaginary part for $\varepsilon = \omega_{1,2}$ and lead to the shift current, while denominators with γ lead to the injection current. Sipe and Shkrebti [5] discussed the order of limits $\gamma \rightarrow 0$, and then $\bar{\omega} \rightarrow 0$. The divergence then manifests in a finite current injection $dJ/dt \neq 0$. Another alternative is to discuss the DC case as the limit $\bar{\omega} \rightarrow 0$, followed by $\gamma \rightarrow 0$ [8]. In this context it is important to keep in mind that any finite quasiparticle lifetime will limit the phenomenon of current injection to irradiation times shorter than the lifetime. At longer times, the current instead saturates at a large but finite current. For simplicity, we continue to call a current proportional to γ^{-1} injection current, even when keeping γ finite.

The injection current is due to the conductivity

$$\sigma_{(i)}^{ab;c} = \frac{e^3}{\hbar^2 \omega^2} \int_k \sum_m f_m \frac{[v^a \Delta^c, v^b \mathcal{D}]_{mm}}{i\gamma}, \quad (29)$$

$$\mathcal{D}_{mn} = \frac{1}{\varepsilon_{mn} + \omega + i\Gamma} + \frac{1}{-\varepsilon_{mn} - \omega + i\Gamma}, \quad (30)$$

In the clean limit the last term simplifies to the injection current in the form written in Eq. (8). In particular, for applied frequencies $\omega \gg \Gamma$, $\mathcal{D}$ assumes the form of δ functions. This means that current injection will occur invariably for any ordering of the three limits $\bar{\omega} \rightarrow 0$, $\gamma \rightarrow 0$, and $\omega \rightarrow 0$. However, completing the expansion under the assumption that $\omega \ll \Gamma$, the result is instead

$$\sigma_{(i)}^{ab;c} = \frac{e^3}{\hbar^2 \omega^2} \int_k \sum_m f_m [r^a \Delta^c, r^b]_{mm} \frac{-2\Gamma}{\gamma}. \quad (31)$$

This expression is finite as long as $\Gamma/\gamma < \infty$, meaning that the effect of current injection ceases in the strict zero frequency limit. We caution that the parameter regime $\omega \ll \Gamma$ is not well controlled in the present formulation, which somewhat hides the effects of Pauli blocking and disregards possible localization effects, both of which might further decrease the current. However, this does not affect our observation that it requires not only time-reversal breaking, but also well defined quasiparticles ($\omega \gg \Gamma$) to prevent the cancellation of the time-reversed paths. The effect of current

FIG. 3. Comparison of the lowest order diagrams which contain time-reversed paths for the case of (a) impurity scattering and (b) electromagnetic interaction. Dashed lines denote disorder lines, the red colored propagators are coherently counterpropagating. The disorder problem does not have any lower order diagrams of this kind because the disorder average removes any correlations between self-energy terms of different propagators, making it necessary for the disorder to connect two fermion lines in order for it to have a coherent effect.

injection for ($\omega \gg \Gamma$) is reminiscent of antilocalization in a disordered system with spin-orbit coupling. In systems with antilocalization, pairs of counterpropagating paths will not interfere destructively due to the spin-momentum locking, leading to an increased value of the linear conductivity. Upon removing spin-orbit coupling, both paths acquire opposite Aharonov-Bohm phases and the destructive interference between them leads to a decrease in conductivity. Notably, weak antilocalization persists in the strict adiabatic limit. In contrast, the effect of current injection discussed here relies on applied frequencies which exceed the inverse lifetime of excited states ($\omega \gg \Gamma$), and is thus intrinsically a dynamical effect, which does not extend to the adiabatic limit where all frequencies approach zero uniformly.

The common property which we want to draw attention to is the following: Both in weak localization and for the dynamical effect discussed here the leading order contribution to the conductivity is (fully or partially) neutralized due to destructive interference with another term where a retarded propagator is exchanged for an advanced one, *unless* the phase of the electron wave function is rotated nontrivially (for example due to time-reversal breaking). Crucially, the lowest order where this happens for weak antilocalization is the crossed disorder diagram, but for second-order optical response it is the three-leg loop (Fig. 3). The effect first appears in different diagrams due to the fact that self-energy corrections from disorder interactions are not coherent between time-reversed paths, whereas the interaction with an external plane wave is coherent. Another difference is that the treatment of localization effects require the summation of the Diffuson and Cooperon [72], while the finite-frequency transitions considered here are perturbative in the applied electric field.

Current injection could therefore be described as an effect of *dynamical* antilocalization. It is tied to the appearance of

fermion loops with more than two vertices, meaning similar effects should be present in nonlinear response functions of any order, given some suitable order of limits for the external frequencies.

Using the expression for the clean case, the δ functions allow the substitution $\omega \rightarrow -\varepsilon_{mn}$ in the global coefficient. The result coincides with the corresponding expression in Eq. (8),

$$\sigma_{(i)}^{ab;c} = \frac{-e^3}{\hbar^2} \int_k \sum_m f_m \frac{[r^a \Delta^c, r^b \mathcal{D}]_{mm}}{i\gamma}. \quad (32)$$

Then the following suggestive formulation of the commutator is possible:

$$[r^a \Delta^c, r^b \mathcal{D}] = [v^a, \partial^c r^b \mathcal{D}] - [r^a, \partial^c v^b \mathcal{D}], \quad (33)$$

where we used $\partial^c v^b = i\varepsilon \partial^c r^b + i\Delta^c r^b$. We therefore interpret the injection current as a mismatch between $v^a \partial^c r^b$ which contains the derivative of the Berry connection $\partial^c r^b$, and $r^a \partial^c v^b$ due to the kinematic acceleration $\partial^c v^b$, each of them integrated over the manifold where $\varepsilon_{mn} = \pm\omega$. This reinforces the dynamical nature of the injection current, which appears due to the different accelerations acting on a band wave function and on a quasiparticle. It is not possible to formulate either acceleration purely in terms of the Berry curvature, with the exception of a two-band model, where Δ_{mn}^c can trivially be moved outside of the commutator [19].

The same treatment can be applied to the shift current by analyzing resonances for $\Gamma \rightarrow 0$, which yields

$$\begin{aligned} \sigma_{(s)}^{ab;c} = \frac{e^3}{\hbar^2 \omega^2} \int_k \sum_m f_m & ([i\pi \delta(\omega + \varepsilon) v^a, \frac{\varepsilon}{2} [r^b, r^c] \\ & - i v^b \delta^c + i \varepsilon R^{bc}] + [i\pi \delta(\omega + \varepsilon) v^b, \frac{\varepsilon}{2} [r^a, r^c] \\ & - i v^a \delta^c + i \varepsilon R^{ac}])_{mm}. \end{aligned} \quad (34)$$

This expression for the shift current is identical to the established formulation in terms of a shift vector [2]. For linear polarized light it is once again helpful to insert the band energy differences (i.e., replace $\omega \rightarrow \pm\varepsilon_{mn}$), to obtain for finite relaxation rates

$$\begin{aligned} \sigma_{(s)}^{aa;c} = \frac{e^3}{\hbar^2} \int_k \sum_m f_m & (-[D' r^a, R^{ca}] \\ & + i[D' r^a, \frac{1}{2}[r^a, r^c] + \delta^c r^a])_{mm}. \end{aligned} \quad (35)$$

where we used

$$D'_{mn} = \frac{-1}{\varepsilon_{mn} + \omega + i\Gamma} + \frac{-1}{\varepsilon_{mn} - \omega + i\Gamma}. \quad (36)$$

The shift current can thus be viewed as a geometric effect with two sources, the quantum metric and the motion due to the Berry dipole, but with the intraband acceleration projected out. For an explanation of this second part, we recall that $\delta^c r^a$ appears as a consequence of replacing the velocity v^c in Eq. (24) by the Berry connection, compensating the overcounting that happens with respect to the intraband motion. This intraband motion is actually diabatic and is exactly what gives rise to the injection current; consequently the projected quantity $[r^c, r^a] + \delta^c r^a$ comprises only the remaining adiabatic processes. We emphasize that the obtained expression for the shift current is in agreement with the previous literature

(e.g., [5]). At the same time, the presented derivation of a counterterm for the diabatic motion which forms part of the shift current allows us to connect this term with the ballistic current from the kinetic approach: Just like the ballistic current, it carries a factor $1/\gamma$ from the injection current, which is canceled by the expansion in γ/ε [39,51]. This confirms that the theory of second-order optical response recovers such an order-one effect in the shift current, an important consistency check. We point out that in the literature, the shift current is sometimes further subdivided into ballistic current j_b , shift current due to excitation j_{ex} , and the shift current due to recombination j_{rec} [51]. This is necessary if the source of the shift is not intrinsic, i.e., not due to the asymmetry in the band structure but due to impurity scattering. In this case, the lifetimes entering Eq. (35) can be dissimilar so that the sum of these three currents is no longer uniquely determined by the dispersion. The same observation was made by considering the side jump mechanism as a source of electron shift [39]. As a study of microscopics is beyond the scope of our work, it is not possible to check whether all results of the kinetic approach are recovered in the shift current formalism for finite lifetimes. This is left for future work.

In summary, borrowing from the language of antilocalization, current injection occurs because there is a mismatch between adiabatic motion and finite frequency motion, which by the order of limits which defines optical response cannot cancel. We repeat that $\mathcal{T}$ enforces that current injection appears only in response to circular polarized light. The shift current is in turn the sum of all remaining finite frequency processes, except for the type of intraband motion which already gives rise to the injection current. The fundamental difference between shift and injection current can also be discussed in terms of the units of the involved integrands. While the matrix elements in the shift current Eq. (35) naturally form a quantity with the units of a Berry dipole, this is not true for the injection current, because there is an additional relaxation time in the denominator. This distinction will also prove useful in the semiclassical expressions.

C. Anomalous semiclassical terms

Finally, we derive the semiclassical expansion in the absence of $\mathcal{T}$. To repeat, in this case Eqs. (22) and (23) are no longer helpful in canceling even powers in the ω dependence. Up to order ω^{-1} , the regular terms are

$$\sigma^{ab;c} = \frac{e^3}{\hbar^2} \int_k \sum_m f_m \left(\frac{S_{-2}}{\omega^2} + \frac{S_{-1}}{\omega} + \dots \right), \quad (37)$$

$$S_{-2} = \partial^a \partial^b v^c - [r^a, \Delta^c r^b] - [r^c, [v^a, r^b]] + [r^c, [r^a, v^b]] \quad (38)$$

$$S_{-1} = -i\partial^c \Omega^{ab} - 2 \left[\frac{r^a}{\varepsilon}, \Delta^c r^b \right]. \quad (39)$$

As explained, injection and shift current do not appear here as they are not amenable to an expansion in small ω . Also, the leading correction for finite lifetimes is given by the replacement $\omega \rightarrow \omega + i\gamma$. From the systematics of expression (24), one can immediately deduce that only the term of order ω^{-1} can have a purely geometric origin. All other orders

necessarily depend on velocity matrix elements or the dispersion relation. Importantly, for $\mathcal{T}$ we recover immediately that $S_{-2} = 0$, and S_{-1} becomes an only function of the Berry dipole $\partial^c \Omega^{ab}$. This latter term is the well-known nonlinear AHE [26,39,41]. We emphasize that in the absence of $\mathcal{T}$, there is also an AHE at linear order in the electric field. In this case it can be difficult to distinguish linear from nonlinear contributions. At least for the intrinsic contributions to the AHE discussed here, the leading order nonlinear effect as given by S_{-2} scales with γ^{-2} , which is distinct from the linear AHE, which is of order one. Alternatively, in materials which break both inversion symmetry ($\mathcal{P}$) and $\mathcal{T}$ but preserve the combined symmetry $\mathcal{PT}$, the Berry curvature is identically zero everywhere, which means that the linear AHE vanishes.

The second-order response Eq. (24) thus decomposes into two pieces given by Eqs. (29) and (35) with a singular frequency dependence and the regular terms Eq. (37). Compared to previous results, the reported formulas have no further derivatives which act on the propagators, they explicate the frequency dependence and they phrase the response in terms of four physical processes, each of which is, in principle, measurable independently. For the limit $\omega \rightarrow 0$, our semiclassical expressions recover the lifetime dependence which was found for the nonlinear anomalous Hall conductivity as calculated in the Boltzmann kinetic approach [39,62]. As an aside, we point out that the most often used formula for shift and injection current, Eq. (8), does not lend itself to a straightforward generalization for finite lifetimes, as the replacement of applied frequencies by band energy differences is only possible in the limit $\omega \gg \gamma$. Instead, the correct starting point in the length gauge are expressions containing the covariant derivative, which can be found for example in [7,52].

In the absence of a gap, the semiclassical expansion Eq. (37) for small ω becomes questionable. This is, however, not a problem for the wide range of Dirac and Weyl semimetals which have undertilted cones: The regions in k space in the direct vicinity of the Dirac points do not contribute significantly to the nonlinear conductivity, rendering them irrelevant. In contrast, in type II Weyl semimetals this does not hold, hence they exhibit a very different phenomenology [73].

D. Conductivity tensor

At this point it is helpful to recall that sometimes several components of the nonlinear conductivity $\sigma^{ab;c}$ are combined for a given current response, depending on the applied electric fields. For example, the BPVE for linear polarized light with polarization direction along the x, y diagonal is given by

$$j^c = \frac{1}{2} (\sigma^{xx;c} + \sigma^{yy;c} + \sigma^{xy;c} + \sigma^{yx;c}) |E|^2. \quad (40)$$

In the same fashion, light with anticlockwise polarization in the x, y plane leads to the current

$$j^c = \frac{1}{2} (\sigma^{xx;c} - \sigma^{yy;c} + \sigma^{xy;c} - \sigma^{yx;c}) |E|^2. \quad (41)$$

We will employ some of these properties in the next section to recast the discussion in a more familiar fashion in terms of the symmetries of real and imaginary parts of the matrix elements.

IV. DC CURRENT RESPONSE

In the following we explore the DC-current creation from a single Weyl cone and relate it to the general phenomenology we explained in the previous section. Of particular interest will of course be the injection current, but we also make some comments regarding the remaining contributions which combine to the shift current. Diagrams (III)–(VI) contain the contribution from curvature effects in the intraband motion, they do not contribute to the injection current and will therefore be disregarded in this section.

Since the Abelian Berry curvature is strictly zero in a space-time inversion (combined symmetry by the inversion symmetry and $\mathcal{T}$, $\mathcal{PT}$) symmetric material, the contribution from the Berry dipole vanishes. We utilize this special case to investigate the DC response when it is dominated by acceleration terms unrelated to the Berry curvature.

To repeat, the second-order optical response is given by the six diagrams depicted in Fig. 1. The DC-current response is $\sigma^{ab;c}(0, \omega, -\omega) \equiv \sigma_{\text{DC}}^{ab;c}$. Additionally, for the benefit of familiarity, in this section we will fall back on the standard notation in band indices instead of the very compact notation of the previous section. Without loss of generality, consider in the following a coordinate system which is aligned with the direction of the incident linear polarized light. Then for the circular polarization, the indices a and b of the electric field are dissimilar while for the linear polarized light it is $a = b$. For any other choice of coordinates, the following statements still hold for the respective current response, but not individually for the matrix elements. It follows in either case that

$$\sigma_{\text{circ}}^{ab;c} = \sum_{m,n,l} \int_k f_{nl} \text{Im}[A_{nml}^{abc}(\omega) - A_{nml}^{abc}(-\omega)], \quad (42)$$

$$\sigma_{\text{lin}}^{aa;c} = \sum_{m,n,l} \int_k f_{nl} \text{Re}[A_{nml}^{aac}(\omega) + A_{nml}^{aac}(-\omega)]. \quad (43)$$

Furthermore,

$$A_{nml}^{abc}(\Omega) = \frac{e^3}{\hbar^2 \Omega^2} \frac{N_{nml}^{abc}}{(\varepsilon_{mn} - i\gamma)(\Omega - \varepsilon_{nl} - i\Gamma)}, \quad (44)$$

where $N_{nml}^{abc} = v_{nl}^a v_{lm}^b v_{nm}^c$. Note that the intraband relaxation rate γ is usually significantly smaller than the interband relaxation rate Γ . The formulation as real and imaginary part follows from the observation that the incoming electric fields either fulfill $E^a(\omega)E^b(-\omega) = E^a(-\omega)E^b(\omega)$ (linear polarization) or $E^a(\omega)E^b(-\omega) = -E^a(-\omega)E^b(\omega)$ (circular polarization). We will occasionally use the shorthand $\Sigma_{\pm}^{ab;c}(\omega)$ with $\sigma_{\text{lin}}^{ab;c} = \text{Re} \Sigma_{+}^{ab;c}(\omega)$ and $\sigma_{\text{circ}}^{ab;c} = \text{Im} \Sigma_{-}^{ab;c}(\omega)$.

A. Cancellations from time-reversal symmetry

In the presence of $\mathcal{T}$, energy eigenvalues are inversion symmetric in the Brillouin zone $\varepsilon_n(\mathbf{k}) = \varepsilon_n(-\mathbf{k})$ and velocity matrix elements fulfill $v_{nm}^a(\mathbf{k}) = -v_{nm}^a(-\mathbf{k})$. In this case, both injection current and shift current benefit from cancellations between $\mathbf{k}$ and $-\mathbf{k}$. The important term to keep track of is $\varepsilon_{mn} - i\gamma$, which in the limit $\gamma \rightarrow 0$ is manifestly divergent for $m = n$. We therefore proceed to write out the real and imaginary parts of the energy denominators $D_{mn}^{-1} = \varepsilon_{mn} - i\gamma$ and $D_{nl}^{-1}(\omega) = \Omega - \varepsilon_{nl} - i\Gamma$. Since the nonlinear

DC response is proportional to $E_a(-\omega)E_b(\omega)$ one can replace $a \leftrightarrow b$ and $\omega \leftrightarrow -\omega$ in the expression for A , meaning that for an inversion symmetric Brillouin zone the second-order response can be written as proportional to

$$N_{nml}^{abc} - sN_{mnl}^{abc} = V_s, \quad (45)$$

where $s = +1$ ($s = -1$) for linear (circular) polarized light. For the various combination possibilities of real and imaginary parts of J^{lin} , this implies

$$\text{Re} D_{mn} \text{Re} D_{nl}(\omega) \text{Re} V_1 = 0, \quad (46)$$

$$-\text{Im} D_{mn} \text{Re} D_{nl}(\omega) \text{Im} V_1 = 0, \quad (47)$$

$$-\text{Re} D_{mn} \text{Im} D_{nl}(\omega) \text{Im} V_1 \neq 0, \quad (48)$$

$$-\text{Im} D_{mn} \text{Im} D_{nl}(\omega) \text{Re} V_1 = 0, \quad (49)$$

while J^{circ} has only one cancellation,

$$\text{Re} D_{mn} \text{Re} D_{nl}(\omega) \text{Im} V_{-1} \neq 0, \quad (50)$$

$$\text{Im} D_{mn} \text{Re} D_{nl}(\omega) \text{Re} V_{-1} = 0, \quad (51)$$

$$\text{Re} D_{mn} \text{Im} D_{nl}(\omega) \text{Re} V_{-1} \neq 0, \quad (52)$$

$$-\text{Im} D_{mn} \text{Im} D_{nl}(\omega) \text{Im} V_{-1} \neq 0. \quad (53)$$

In Eqs. (46) and (49) we made use of the fact that $a = b$ for linear polarized light, which implies $\text{Re} V_1 = 0$. Equation (47) follows from the fact that $\text{Im} D_{mn} \neq 0$ requires $m = n$, which entails $\text{Im} V_1 = 0$, and Eq. (51) is zero due to the antisymmetry in the band summation for $m = n$.

B. Space-time inversion symmetry

In the absence of $\mathcal{T}$, the aforementioned cancellations do no longer hold and a comparable behavior (injection current) is expected for both linear and circular polarization. Using the conventional argumentation, without $\mathcal{T}$ any combination of real and imaginary parts of the denominators is potentially nonzero and has to be reconsidered. This is of course notwithstanding crystal symmetries, which can still render certain components zero. As we discussed, the well-known formula for the shift current and linear polarization,

$$\sigma_{\text{lin}}^{aa;c} \sim \sum_{n,m} f_{nm} \text{Im}[r_{mn}^a r_{nm;c}^a] \delta(\varepsilon_{nm} - \omega), \quad (54)$$

is no longer the most relevant part of the response function once $\mathcal{T}$ is broken. This term originates from Eq. (48) only and evaluates to zero for $m = n$. Instead, resonances involving $m = n$ provide the dominant contribution to the BPVE. It is then useful to separate two-band contributions containing $\text{Im} D_{mn} = \gamma^{-1} \delta_{mn}$ from three-band ones with $\text{Re} D_{mn} = \varepsilon_{mn}^{-1}(1 - \delta_{mn})$, the latter of which are suppressed by the energy denominator ε_{mn} .

In the presence of $\mathcal{PT}$ symmetry, only the real part of the product of momentum matrix elements N_{nml}^{abc} contributes. Imposing $\mathcal{PT}$ and focusing on the dominant two-band contributions, the response therefore features the following term:

$$-\sum_{m,n,l} f_{nl} \text{Im} D_{mn} \text{Im} D_{nl}(\omega) N_{nml}^{aac} \pm (\omega \rightarrow -\omega). \quad (55)$$

Note that for a $\mathcal{T}$ material, only the imaginary part of the numerator N_{nml} contributes to the DC current, while in the presence of $\mathcal{PT}$ symmetry only the real part does. In this sense, both are complementary. Since $\mathcal{T}$ is broken, a symmetrization with respect to $\mathbf{k}$ is no longer useful. Instead, one should consider each Dirac cone separately. It was previously noted that semimetals have zero shift current if the cone has a linear dispersion [18,73]. This is related to the triple product of velocities, which is antisymmetric on a radial shell which concentrically encloses a perfect Dirac point.

To be concrete, suppose a tight-binding model with $\mathcal{PT}$ symmetry, constituted of two pairs of degenerate bands. Given a trivial implementation of inversion symmetry, the most general $\mathcal{PT}$ -symmetric four-band Hamiltonian is [74]

$$H_{PT} = g_{x0}(\mathbf{k})\tau_x + g_{z0}(\mathbf{k})\tau_z + g_{yx}(\mathbf{k})\tau_y\sigma_x + g_{yy}(\mathbf{k})\tau_y\sigma_y + g_{yz}(\mathbf{k})\tau_y\sigma_z, \quad (56)$$

where σ and τ are Pauli matrices of spin and orbital degrees. If the orbital makeup requires a nontrivial implementation of inversion, some spatial indices will be permuted. Due to the high symmetry in this four-band model, for $g_{x0} = g_{yx} = g_{yy} = 0$, the basis can be chosen pairwise such that one can express dispersion and momentum matrix elements referring to only one upper and one lower band index. In the general case one obtains not one but three structurally identical pieces corresponding to the three different interband transitions. Inserting the degeneracies into the general expression, and using that $\omega_{nn} = 0$, $\omega_{mn} = -\omega_{nm}$, $v_{nm}^a = v_{mn}^{a*}$, one obtains for the nonlinear DC conductivity explicitly

$$\begin{aligned} \Sigma_{\pm}^{abc}(\omega) &= \frac{e^3}{\hbar^2\omega^2} \int_{\mathbf{k}} \frac{f_{du}}{-i\gamma} \left(\frac{\text{Re}(v_{du}^a v_{ud}^b) v_{dd}^c}{\varepsilon_{du} - \omega - i\Gamma} + \frac{\text{Re}(v_{du}^a v_{ud}^b) v_{uu}^c}{\varepsilon_{du} + \omega + i\Gamma} \right) \\ &+ \int_{\mathbf{k}} f_{du} \left(\frac{\text{Re}(v_{du}^a v_{ud}^c) v_{uu}^b}{(\varepsilon_{du} - i\gamma)(\varepsilon_{du} - \omega - i\Gamma)} \right. \\ &\quad \left. - \frac{\text{Re}(v_{du}^a v_{ud}^c) v_{dd}^b}{(\varepsilon_{du} + i\gamma)(\varepsilon_{du} + \omega + i\Gamma)} \right) \\ &\pm (\omega \rightarrow -\omega) + \dots, \end{aligned} \quad (57)$$

where the upper (lower) band is labeled u (d). The dots indicate that there are analogous terms including the other three pairings of bands (involving the $\mathcal{T}$ partners $\bar{u}$, $\bar{d}$), where $v_{ud} \neq v_{\bar{u}\bar{d}} \neq v_{u\bar{d}} \neq v_{\bar{u}d}$. The problem thus decomposes in four pieces which somewhat resembles the two-band case when $\mathcal{T}$ is preserved [27]. However, now there are both two-band and three-band terms.

C. Single cone

Concentrating on one band touching point at a time, the local two-band Hamiltonian corresponding to band (d, u) is $H(\mathbf{k}) = \hbar g_i \sigma_i$, with a generic momentum dependencies $g_i = v_F(k_i + \alpha_{ijl} k_j k_l)$; analogous expressions with dissimilar coefficients exist for the other combinations of bands at the remaining Weyl points. Keeping only the two-band expression for linear polarization, the injection current

becomes

$$\begin{aligned} \sigma_{DC}^{aa;c}(\omega) &= -\frac{2e^3}{\hbar^2\omega^2\gamma} \int_{\mathbf{k}} f_{du} |v_{du}^a|^2 (v_{dd}^c - v_{uu}^c) \\ &\times \text{Im} \left(\frac{1}{\varepsilon_{du} - \omega - i\Gamma} + \frac{1}{\varepsilon_{du} + \omega + i\Gamma} \right). \end{aligned} \quad (58)$$

The velocity matrix elements are given by

$$\text{Re}(v_{du}^a v_{ud}^b) v_{dd}^c = \hbar^3 \sum_{ijl} \partial^a g_i \partial^b g_j \partial^c g_l \frac{g_i g_j g_l}{\varepsilon_i \varepsilon_j \varepsilon_l} (\delta_{ij} - 1). \quad (59)$$

Expanding in the nonlinearity α , the zeroth term vanishes, as expected. At the next order, since $i \neq j$ in Eq. (59), the only contribution is ($\Gamma \rightarrow 0$)

$$\begin{aligned} \sigma_{DC}^{aa;c}(\omega) &= -\frac{2e^3(v_F\hbar)^3}{\hbar^2\omega^2\gamma} \int_{\mathbf{k}} 8\delta_{cl}\delta_{ai}\alpha_{iaq}k_q \frac{k_i k_j k_l}{|\mathbf{k}|^3} \\ &\times (\delta_{ij} - 1)\delta(2v_F|\mathbf{k}| - \omega) \end{aligned} \quad (60)$$

so the two cases of linear polarization are

$$\begin{aligned} \sigma_{DC}^{aa;a}(\omega) &= \frac{16e^3(v_F\hbar)^3}{\hbar^2\omega^2\gamma} \int_{\mathbf{k}} \sum_{i \neq a} \alpha_{iai} \frac{k_a^2 k_i^2}{|\mathbf{k}|^3} \delta(2v_F|\mathbf{k}| - \omega) \\ &= \frac{2}{3\pi} \frac{e^3}{\hbar^2} \sum_{i \neq a} \frac{\alpha_{iai}}{\gamma} \omega, \end{aligned} \quad (61)$$

$$\begin{aligned} \sigma_{DC}^{aa;c}(\omega) &= \frac{16e^3(v_F\hbar)^3}{\hbar^2\omega\gamma} \int_{\mathbf{k}} \alpha_{caa} \frac{k_a^2 k_c^2}{|\mathbf{k}|^3} \delta(2v_F|\mathbf{k}| - \omega) \\ &= \frac{2}{3\pi} \frac{e^3}{\hbar^2} \frac{\alpha_{caa}}{\gamma} \omega \quad \text{for } c \neq a. \end{aligned} \quad (62)$$

For finite interband relaxation rates ($\Gamma \neq 0$), there is a crossover at frequency $\omega_c \sim (\Gamma\Lambda^2)^{1/3}$, with Λ the bandwidth of the system. Below this ω_c , $\sigma_{DC}^{aa;c}$ acquires a residual term of the size $\sim \Gamma\Lambda^2/\gamma\omega^2$.

Barring any symmetry cancellations, for transitions near the Dirac point we thus obtain the estimate

$$\sigma_{DC}^{aa;c}(\omega) = c_1 \frac{\Gamma\Lambda^2}{\omega^2\gamma} + \left(\frac{c_2}{\gamma} + c_3 \right) \omega, \quad (63)$$

where c_1, c_2, c_3 are coefficients depending on the nonlinearity in the Dirac cone dispersion. The nonidentical Dirac cone close to $-\mathbf{k}$ will contribute a similar term with numerically different coefficients, but this does not change the qualitative behavior. Additionally, the values of c_1 and c_3 are not exclusively determined by diagrams (I) and (II) but will also receive a contribution from the remaining second-order diagrams.

In summary, in the case that both $\mathcal{T}$ and inversion are preserved, the second-order DC response will vanish. With T present but P being absent, what matters is the imaginary part of the product of three velocities (N_{nml}^{abc}), with P present but without T , both the real and imaginary part of N_{nml}^{abc} matters, and in the absence of both P and T but with $\mathcal{T}$ symmetry only the real part of N_{nml}^{abc} remains. For the $\mathcal{T}$ -symmetric case, weak $\mathcal{T}$ breaking will therefore result in a strong suppression of the nonlinear DC conductivity due to the nearly complete cancellation of the contributions from $\mathbf{k}$ with the ones from $-\mathbf{k}$. A strong response is therefore expected from intrinsically

$\mathcal{T}$ -breaking materials where the $\mathcal{T}$ is not broken perturbatively. At leading order the size of the injection current is dependent on the band structure near the Dirac cone through the combination α/γ , which means that a Dirac cone with a large nonlinear components of type α_{cac} or α_{caa} appears to be the most promising candidate for measuring a large DC current.

Regarding the contributions from diagrams (III)–(VI) which we safely discarded, the situation becomes quite different if $\mathcal{T}$ is preserved, rendering diagrams (I) and (II) of size $O(1)$ instead of $O(\gamma^{-1})$. Then, all diagrams have a similar sized contribution to the DC current [27].

D. Minimal four-band model

We corroborate the findings of the Sec. IV C with an explicit calculation of the BPVE in a $\mathcal{PT}$ -symmetric four-band model. As a comparison, a similar four-band model exhibiting $\mathcal{T}$ with comparable dispersion and density of states is chosen. They are a representative of the BPVE which arises from two-photon processes in a Dirac cone close to the nodal point. Of course, at a higher frequency resonances will appear which are particular to the chosen band structure. To describe the general photoresponse beyond the injection current, we should consider all the diagrams in Fig. 1, which corresponds to Eq. (24).

As a tight-binding Hamiltonian we choose

$$\begin{aligned} \sigma_{\text{DC}}^{aa;c}(\omega) &= c_1 \frac{\Gamma \Lambda^2}{\omega^2 \gamma} + \left(\frac{c_2}{\gamma} + c_3 \right) \omega, \\ H &= \hbar t \{ \sigma_y \sin k_y + s_x \sigma_z \sin k_z \\ &\quad + \sigma_x s_0 [\cos k_x + m_y (1 - \cos k_y) + m_z (1 - \cos k_z)] \\ &\quad + \Delta [s_x \sigma_y \cos k_y + s_z \sin k_z] \\ &\quad + \delta [s_y \sigma_z \cos k_x + s_z \sigma_z (\cos k_x + \sin k_z)] \}. \end{aligned} \quad (64)$$

The parameters m_y, m_z produce anisotropy, Δ breaks inversion symmetry, and δ induces time-reversal and inversion breaking, such that while the Hamiltonian does not commute with either $\mathcal{P}$ or $\mathcal{T}$, their combined operation $\mathcal{PT}$ is a symmetry. The spectrum is then of two doubly degenerate bands. In the following we compare the nonlinear conductivity between the choice of parameters $\Delta = 0, \delta \neq 0$, implementing $\mathcal{T}$ and the choice $\Delta \neq 0, \delta = 0$ for $\mathcal{PT}$ symmetry. The band structures have the form shown in Fig. 4. Both cases lead to a very similar total density of states (DOS) [Fig. 4(c)], with a clear separation between the low energy Dirac cone physics and the high energy Lifshitz point where the cones connect.

The Hamiltonian Eq. (63) is generic in the sense that it encodes the minimal number of four Weyl cones in case of $\mathcal{T}$ and the minimal number of two doubly degenerate Dirac cones in the $\mathcal{PT}$ case. The momentum dependencies are not generic, they are chosen in reference to a minimal model which was proposed for the time-reversal-symmetric material TaAs, which has a cubic Brillouin zone. In Appendix C the steps are listed which lead to Eq. (63), suffice to say here that the only nontrivial part is to keep enough structure in the momentum dependencies such that there are no unwanted additional symmetries like mirrors present. There is one speciality added to the $\mathcal{PT}$ -symmetric dispersion which is to not

FIG. 4. Band structure in the (k_x, k_y) plane at $k_z = 0$, for the $\mathcal{PT}$ -symmetric Hamiltonian (a) and with the $\mathcal{T}$ -symmetric one (b). The parameters are chosen as $m_y = 1, m_z = 5, (\Delta, \delta) = (0, 0.5)$ (for $\mathcal{PT}$) and $(\Delta, \delta) = (0.5, 0)$ (for $\mathcal{T}$). (c) Comparison of the density of states in both cases as a function of chemical potential.

include a term which would correspond to a nonlinearity in the dispersion of the type α_{zyy} in the vicinity of the Dirac points. This allows us to verify the claim of Eq. (61) that the injection current is sourced from exactly one nonlinear dispersion term at next-to-leading order. In other words, the expected injection current for our $\mathcal{PT}$ -symmetric example for the DC-conductivity $\sigma^{yy;z}$ has size $O(\omega^2/\gamma)$ instead of $O(\omega/\gamma)$. For the numerical evaluation we repeat that diagrams (III)–(VI) are negligibly small, so that it is enough to consider diagrams (I) and (II), with varying but finite relaxation rates $\gamma = \Gamma$. The second-order DC-conductivity $\sigma^{yy;z}$ is shown in Fig. 5 as a function of frequency. As expected, the injection current persists for large frequencies, crossing over into a shift current for small ω . We emphasize the difference in magnitude of roughly 10^2 between $\mathcal{T}$ and $\mathcal{PT}$, in spite of the very comparable band structure, bandwidth, and relaxation rates used for both models.

V. CONCLUSIONS

For many years, studies of the bulk photovoltaic effect were restricted to the evaluation of the complicated perturbative expressions. Here we used the diagrammatic approach to describe second-order optical response for materials without time-reversal symmetry and also with finite relaxation times in

FIG. 5. Comparison of the modulus of $\sigma^{yz}(0, \omega, -\omega)$ between the $\mathcal{PT}$ -symmetric case (a) and the one with $\mathcal{T}$ (b). For improved visibility, we show the product $\sigma^{yz}\omega^2$. For $\mathcal{PT}$, the nonlinear DC conductivity approaches a ω^{-2} dependence for small frequencies which is independent of the quasiparticle lifetime γ . At frequencies above a threshold proportional to Γ , the current creation is only capped by the lifetime $1/\gamma$. For $\mathcal{T}$, the current vanishes linearly with ω^1 for small frequencies, save for a residual term of size $\gamma\omega^{-2}$ for the smallest frequencies. At frequencies comparable to the bandwidth, the spectrum is sensitive to nesting effects, and thus to the band structure. Note that the apparent erratic behavior for these large values of ω is due to the logarithmic scale, not numerical artifacts. Parameter values are the same as in Fig. 4.

terms of four distinct physical processes. The central element was the analysis of the anomalous acceleration, which appears naturally in the velocity gauge. The resulting analytical structure [e.g., Eqs. (24), (29), and (35)] of the expressions is substantially more amenable to analytical manipulations, but can also be directly implemented in numerics.

We derived explicit sum rules for nonquadratic Hamiltonians and demonstrated that their cancellation properties are equivalently accounted for in the diagrammatic formulation of the velocity gauge.

Contrary to the semiclassical intuition, the presence of a Berry dipole is not a necessary requirement for a large BPVE. Instead, the topological content of the second-order response is encoded in expressions which involve both the

antisymmetric and the symmetric part of the quantum geometric tensor. In particular, we showed that the shift current can be written as the sum of a piece containing the quantum metric and certain elements of the Berry dipole. The shift current is therefore of topological origin, but not in the usual sense that it would be proportional to the Berry curvature. Also, it does not evaluate to a quantized value in a topological insulator, and it contains a nonuniversal piece which we connected to the ballistic current. In agreement with earlier work, we recover the Berry dipole in the semiclassical expansion, but alongside find a number of other terms which only vanish in the presence of $\mathcal{T}$. This means that the semiclassical current response is of topological origin only in time-reversal-symmetric compounds. From the same analysis, we conclude that the injection current is generically not of topological origin, instead being attributed to *dynamical antilocalization* between counterpropagating momentum space trajectories.

We believe that the same analysis is feasible for third-order conductivities, with a larger number of physically distinct processes emerging for the matrix elements involving third derivatives (“jerk”).

The principles outlined here should prove helpful in the search for materials with tailor-made nonlinear optical response. For example, we outlined that current injection is enhanced in semimetals with broken time-reversal symmetry when they exhibit a large nonlinear term in their dispersion near the Dirac cone.

ACKNOWLEDGMENTS

We thank S. Fischer, O. Golan, J. Hofmann, R. Thorgren, A. Stern, M. Mittenzweig, Y. Zhang, H. Ishizuka, N. Nagaosa, L. Wu, F. de Juan Sanz, A. G. Grushin, and D. E. Parker for helpful discussion. B.Y. acknowledges the financial support by the Willner Family Leadership Institute for the Weizmann Institute of Science, the Benozio Endowment Fund for the Advancement of Science, Ruth and Herman Albert Scholars Program for New Scientists, and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant No. 815869).

APPENDIX A: RELAXATION RATES

Here we give some further details on the nature of the finite relaxation rates γ and Γ in the three-leg diagrams. In a Fermi liquid, lifetime effects enter the propagator through the retarded self-energy $\Sigma^R(\mathbf{k}, \omega)$, according to $\Sigma_m^R(\mathbf{k}, \omega) = G_m^{R-1}(\mathbf{k}, \omega) - G_m^{R-1}(\mathbf{k}, \omega)$, where $G_m^{R-1}(\mathbf{k}, \omega) = \omega + i0^+ - \epsilon_m(\mathbf{k})$ is the bare retarded Green’s function. In the relaxation time approximation, the self-energy is projected onto the Fermi surface and averaged, thus replacing it by a constant. This procedure is ubiquitously used for ground state calculations. Due to the multiband nature of the optical response, here the reference point is not always the Fermi surface but the respective band where the electron is traveling, i.e., $\langle \Sigma_m^R(\mathbf{k}, \omega) \rangle_k = i/2\tau_m$. After the contour integration in the loop frequency Ω , the triple product of propagators

$\mathcal{G}^R(\mathbf{k}, \Omega) \mathcal{G}^{R/A}(\mathbf{k}, \Omega + \omega_1) \mathcal{G}^A(\mathbf{k}, \Omega + \omega_1 + \omega_2)$ leads to several terms with two denominators. In these, three frequency combinations appear, ω_1 , ω_2 , and $\omega_1 + \omega_2$, each of them with the imaginary part $1/2\tau_m + 1/2\tau_n$. The only question is which combinations of band indices are the relevant ones for the zeros of the real part of the denominators.

If we consider the case that $\omega_1 = -\omega_2$, two Greens functions are at the same energy, as shown in Fig. 2 in the main text. The denominators with $\omega_{1,2}$ have a pole at $\omega_{1,2} = \varepsilon_{mn}$, where one band index is occupied, and one is unoccupied. This leads to the interband relaxation rate $\Gamma = 1/2\tau_c + 1/2\tau_v$, where the conduction band with index c and the valence band with index v are chosen so that $|\omega_{1,2}| = \varepsilon_{cv}$. Other combination of band indices have different interband relaxation rates, but they are not resonant, so that it is always true that $|\omega_{1,2} - \varepsilon_{mn}| \gg \Gamma$. For this reason, in a given denominator, it is enough to specify one relaxation rate. For the combination $\omega_1 + \omega_2 = 0$, the pole is realized for $\varepsilon_{mn} = 0$, which means that the process is either intraband or the bands m and n are degenerate. In both cases, the relaxation rate is given by $\gamma = 1/2\tau_v + 1/2\tau_v$, which only depends on the microscopic quasiparticle lifetime in the valence band(s). Again, the other combination of band indices in this specific denominator are insensitive to the precise value of the relaxation rate, as it holds again that $|\omega_1 + \omega_2 - \varepsilon_{mn}| \gg \gamma$.

APPENDIX B: DERIVATIVES OF THE BERRY CONNECTION

In the following we establish the general procedure how to easily derive the sum rules necessary for a translation between length gauge and velocity gauge. The first step is to expand derivatives of the matrix elements r_{mn}^a into commutators. To our knowledge this has not been documented so far. We use the definition of the momentum space derivative for a multiband system,

$$\langle m | (\partial^a O) | n \rangle = \partial^a O_{mn} - i r_{ml}^a O_{ln} + i O_{ml} r_{ln}^a. \quad (\text{B1})$$

This represents a known generalization of the usual Berry connection. Note that this non-Abelian connection is covariant under gauge transformations, not invariant. The momentum derivative on the left-hand side can be evaluated directly as long as the operator is given in a basis which is momentum independent, which applies in particular to a tight-binding Hamiltonian. However, for r_{mn}^a the left-hand side is of little use as the derivative cannot act on anything except the wave function, which means we do not know how to write the operator in any other basis than the canonical band basis. Therefore, we only evaluate the right-hand side and use this to define the matrix elements on the left-hand side. An important role is played by the second derivative $\langle m | i \partial^b \partial^c n \rangle - \langle i \partial^b \partial^c m | n \rangle = 2R_{mn}^{bc} = 2R_{mn}^{cb}$. One can equivalently rewrite this by using

$$\begin{aligned} \partial^b \partial^c (\langle m | n \rangle) &= 0 = \langle \partial^b m | \partial^c n \rangle + \langle \partial^c m | \partial^b n \rangle \\ &+ \langle m | \partial^b \partial^c n \rangle + \langle \partial^b \partial^c m | n \rangle, \end{aligned} \quad (\text{B2})$$

$$\langle m | \partial^b \partial^c n \rangle + \langle \partial^b \partial^c m | n \rangle = \langle \partial^b m | \partial^c n \rangle + \langle \partial^c m | \partial^b n \rangle. \quad (\text{B3})$$

We further observe that

$$\langle m | \partial^b \partial^c n \rangle = \langle m | \partial^b \left(\sum_l | l \rangle \langle l | \partial^c n \right) \rangle \quad (\text{B4})$$

$$\begin{aligned} &= \sum_l \langle m | \partial^b l \rangle \langle l | \partial^c n \rangle + \langle \partial^b m | \partial^c n \rangle \\ &+ \langle m | \partial^b \partial^c n \rangle, \end{aligned} \quad (\text{B5})$$

$$\langle \partial^b m | \partial^c n \rangle = \sum_l r_{ml}^b r_{ln}^c. \quad (\text{B6})$$

In other words, the quantity $\langle m | \partial^b \partial^c n \rangle + \langle \partial^b \partial^c m | n \rangle$ is essentially the symmetrized matrix element $\langle \partial^b m | \partial^c n \rangle$, a quantity known as the (bare) quantum metric [66]. Importantly, the diagonal matrix elements of the bare quantum metric are not gauge invariant. The proper gauge invariant distance measure is instead introduced via the quantum geometric tensor Q_{mn}^{ab} . The quantum geometric tensor is defined with the help of the projector $\mathcal{P}_{GS} = \sum_l^N | l \rangle \langle l |$ onto the N -dimensional ground state in band basis, and reads [65]

$$Q_{mn}^{bc} = \langle \partial^b m | 1 - \mathcal{P}_{GS} | \partial^c n \rangle \quad (\text{B7})$$

$$= \sum_l r_{ml}^b r_{ln}^c - \sum_l r_{ml}^b r_{ln}^c. \quad (\text{B8})$$

Here we used that $\langle \partial^b m | n \rangle = -\langle m | \partial^b n \rangle$. The non-Abelian quantum metric g^{ab} is the symmetric part of Q^{ab} , while the non-Abelian Berry curvature is given by the antisymmetric part, i.e., $Q_{mn}^{ab} = g_{mn}^{ab} - \frac{i}{2} \Omega_{mn}^{ab}$. Therefore, by definition the matrix elements $2iR_{mn}^{ab} = \langle \partial_{k_a} \partial_{k_b} m | n \rangle - \langle m | \partial_{k_a} \partial_{k_b} n \rangle$ are related to the quantum metric by

$$2g_{mn}^{ab} = 2iR_{mn}^{ab} + 2\langle m | \partial^a \partial^b m \rangle - \sum_l f_l (r_{ml}^b r_{lm}^c + r_{ml}^c r_{lm}^b), \quad (\text{B9})$$

where we used the Fermi factor f_l to perform the sum over the ground state manifold. One might want to go about and immediately construct the response in terms of the gauge-invariant quantities g and Ω . However, as we show in the main text, this is not recommendable. Instead, the matrix elements which appear in the conductivity are, by construction, always diagonal matrix elements of commutators of Hermitian operators. As such, while the constituents are not gauge invariant, the commutators are. The same observation was reached previously on the basis of general properties of the fiber bundle in Bloch basis [8].

We decompose the momentum derivative in symmetric and antisymmetric parts, yielding $R^{bc} = \partial^b r^c - \frac{i}{2} [r^b, r^c]$. Equivalently, this means,

$$\partial^b r^c - \partial^c r^b = i[r^b, r^c] \quad (\text{B10})$$

$$2R^{bc} = \partial^b r^c + \partial^c r^b. \quad (\text{B11})$$

As usual, the Berry curvature is $\Omega^{bc} = \partial^b r^c - \partial^c r^b = i[r^b, r^c]$. Repeating the same procedure for the Berry curvature dipole results in the diagonal components

$$\partial^a \Omega^{bc} = -\frac{1}{2} [r^a, [r^b, r^c]] + i[r^b, R^{ac}] - i[r^c, R^{ab}], \quad (\text{B12})$$

where we used the identity $0 = [r^b, [r^c, r^a]] + [r^c, [r^a, r^b]] + [r^a, [r^b, r^c]]$. The second step involves manipulations of objects containing velocity matrix elements v_{mn}^a and their derivatives. We use the relation that for off-diagonal matrix elements $\partial^a v^b = i\varepsilon \partial^a r^b + i\Delta^a v^b + O(\Gamma)$ and, as previously established, $\partial^b r^c = R^{bc} + \frac{i}{2}[r^b, r^c]$. For example, for the operator $\partial^a \partial^b H_0(\mathbf{k})$ this yields immediately

$$\begin{aligned} w_{mn}^{ab} &= \partial^a v^b - i[r^a, v^b] \quad m \neq n \\ &= i\varepsilon(R^{ab} + \frac{i}{2}[r^a, r^b]) + i\Delta^a r^b - i[r^a, v^b] + O(\Gamma). \end{aligned} \quad (\text{B13})$$

Under the reasonable assumption that the dispersion relation is mostly smooth, it is $w^{ab} = w^{ba}$. In the main text, this was used to write the result more concisely. However, it also means that

$$\begin{aligned} -\frac{\varepsilon}{2}[r^a, r^b] + i\Delta^a r^b - i[r^a, v^b] \\ = -\frac{\varepsilon}{2}[r^b, r^a] + i\Delta^b r^a - i[r^b, v^a] \end{aligned} \quad (\text{B14})$$

$$i\varepsilon[r^a, r^b] = [r^a, v^b] + [v^a, r^b] - \Delta^a r^b + \Delta^b r^a. \quad (\text{B15})$$

Therefore, also

$$\begin{aligned} -[v^c, [r^a, r^b]] &= [r^c, [r^a, v^b]] + [r^c, [v^a, r^b]] \\ &\quad - [r^c, \Delta^a r^b] + [r^c, \Delta^b r^a] \end{aligned} \quad (\text{B16})$$

In the same manner, we conclude that

$$\begin{aligned} u^{abc} &= \partial^a \partial^b v^c - i[\partial^a r^b, v^c] - i[r^b, \partial^a v^c] \\ &\quad - i[r^a, \partial^b v^c] - [r^a, [r^b, v^c]] \quad (m = n) \\ &= \partial^a \partial^b v^c - [r^a, [r^b, v^c]] + [r^b, \Delta^a r^c] + [r^a, \Delta^b r^c] \\ &\quad - \frac{1}{2}[v^c, [r^a, r^b]] - \frac{1}{2}[v^a, [r^b, r^c]] + \frac{1}{2}[v^b, [r^c, r^a]] \\ &\quad + i[v^c, R^{ab}] + i[v^b, R^{ca}] + i[v^a, R^{bc}] + O(\Gamma). \end{aligned} \quad (\text{B17})$$

By using that $[r^a, [r^b, v^c]] + [r^b, [v^c, r^a]] + [v^c, [r^a, r^b]] = 0$, this becomes

$$\begin{aligned} u^{abc} &= \partial^a \partial^b v^c - [r^a, [r^b, v^c]] + [r^b, \Delta^a r^c] + [r^a, \Delta^b r^c] \\ &\quad + i[v^c, R^{ab}] + i[v^b, R^{ca}] + i[v^a, R^{bc}] \end{aligned} \quad (\text{B18})$$

Additionally it holds that

$$u^{abc} - u^{bac} = 0 = -[r^b, [v^c, r^a]] + [r^a, [v^c, r^b]] \quad (\text{B19})$$

$$\begin{aligned} [v^c, [r^a, r^b]] &= [r^c, [r^a, v^b]] + [r^c, [v^a, r^b]] \\ &\quad - [r^c, \Delta^a r^b] + [r^c, \Delta^b r^a] = 0. \end{aligned} \quad (\text{B20})$$

We emphasize that we do not need all of these relations, but they allow for the sum rules to be rearranged in a condensed format of highly symmetric or cyclical expressions using only velocity matrix elements and wave function derivatives of increasingly higher order depending on the order considered for the nonlinear conductivity. In the main text we show how these various terms can be interpreted physically.

APPENDIX C: COMMENTS ON THE FOUR-BAND MODEL

As a stand-in for a generic four-band model we make use of the tight-binding Hamiltonian previously proposed for TaAs, which reads with normalized hopping

$$\begin{aligned} H &= [\cos k_x + m_y(1 - \cos k_y) + m_z(1 - \cos k_z)]\sigma_x \\ &\quad + \sin k_y \sigma_y + \Delta \cos k_y \sigma_y s_x + \sin k_z \sigma_z s_x. \end{aligned} \quad (\text{C1})$$

Here σ are orbital degrees of freedom and s spin. Δ breaks inversion. Denoting complex conjugation by $\mathcal{K}$ and the replacement $\mathbf{k} \rightarrow -\mathbf{k}$ by $\mathcal{M}$, it holds generally for the time-reversal operator in momentum space

$$\mathcal{T} = i s_y \mathcal{K} \mathcal{M}. \quad (\text{C2})$$

From the Hamiltonian we further deduce the appropriate form of the inversion operator to

$$\mathcal{P} = \sigma_x \mathcal{M}. \quad (\text{C3})$$

This is because just by inversion we flip the signs of the sin terms, which can be rectified by flipping exactly σ_y and σ_z , something achieved by σ_x . Other models will have different $\mathcal{P}$ operators.

The original model also has two mirror symmetries, implemented by

$$\mathcal{X}_x = \mathcal{M}_x s_x, \quad (\text{C4})$$

$$\mathcal{X}_y = \mathcal{M}_y \sigma_x s_y. \quad (\text{C5})$$

This is undesirable because the mirror planes can lead to additional cancellations which might skew the result against what we are trying to demonstrate, i.e., that a $\mathcal{PT}$ -symmetric material can have larger response than a otherwise very similar $\mathcal{T}$ material. For this reason, we break the mirrors by adding $\Delta \sin k_z s_z$ to the time-reversal-symmetric Hamiltonian.

The operator for $\mathcal{PT}$ reads $\mathcal{PT} = i \sigma_x s_y \mathcal{K}$. This means that altogether there are the following five elements in (pseudo-) spin space which preserve this combined symmetry $\mathcal{PT}$

$$\sigma_x \quad \sigma_y \quad \sigma_z s_x \quad \sigma_z s_y \quad \sigma_z s_z. \quad (\text{C6})$$

All of them can also preserve $\mathcal{P}$ and $\mathcal{T}$ separately given the corresponding sign change is compensated by an appropriate momentum dependence.

In close analogy with the $\mathcal{T}$ case, for the $\mathcal{PT}$ -symmetric Hamiltonian we thus choose

$$\begin{aligned} H &= [\cos k_x + m_y(1 - \cos k_y) + m_z(1 - \cos k_z)]\sigma_x \\ &\quad + \sin k_y \sigma_y + \sin k_z \sigma_z s_x + \delta[\cos k_x s_y \sigma_z \\ &\quad + (\cos k_x + \sin k_z) s_z \sigma_z], \end{aligned} \quad (\text{C7})$$

where the momentum dependence of the last term allows for a Dirac point along the x axis.

-
- [1] V. I. Belinicher and B. I. Sturman, The photogalvanic effect in media lacking a center of symmetry, *Sov. Phys. Usp.* **23**, 199 (1980).
 [2] R. von Baltz and W. Kraut, Theory of the bulk photovoltaic effect in pure crystals, *Phys. Rev. B* **23**, 5590 (1981).

- [3] R. W. Boyd, *Nonlinear Optics* (Elsevier, San Diego, CA, 2003).
 [4] W. Kraut and R. von Baltz, Anomalous bulk photovoltaic effect in ferroelectrics: A quadratic response theory, *Phys. Rev. B* **19**, 1548 (1979).

- [5] J. E. Sipe and A. I. Shkrebtii, Second-order optical response in semiconductors, *Phys. Rev. B* **61**, 5337 (2000).
- [6] S. M. Young and A. M. Rappe, First Principles Calculation of the Shift Current Photovoltaic Effect in Ferroelectrics, *Phys. Rev. Lett.* **109**, 116601 (2012).
- [7] 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).
- [8] 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).
- [9] D. Xiao, M.-C. Chang, and Q. Niu, Berry phase effects on electronic properties, *Rev. Mod. Phys.* **82**, 1959 (2010).
- [10] B. Yan and C. Felser, Topological materials: Weyl semimetals, *Annu. Rev. Condens. Matter Phys.* **8**, 337 (2017).
- [11] N. P. Armitage, E. J. Mele, and A. Vishwanath, Weyl and Dirac semimetals in three-dimensional solids, *Rev. Mod. Phys.* **90**, 015001 (2018).
- [12] 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 monophosphide Weyl semimetals, *Nat. Phys.* **13**, 350 (2017).
- [13] Q. Ma, S.-Y. Xu, C.-K. Chan, C.-L. Zhang, G. Chang, Y. Lin, W. Xie, T. Palacios, H. Lin, S. Jia, P. A. Lee, P. Jarillo-Herrero, and N. Gedik, Direct optical detection of Weyl fermion chirality in a topological semimetal, *Nat. Phys.* **13**, 842 (2017).
- [14] G. B. Osterhoudt, L. K. Diebel, M. J. Gray, X. Yang, J. Stanco, X. Huang, B. Shen, N. Ni, P. J. Moll, Y. Ran *et al.*, Colossal mid-infrared bulk photovoltaic effect in a type-I Weyl semimetal, *Nat. Mater.* **18**, 471 (2019).
- [15] P. Hosur and X. Qi, Recent developments in transport phenomena in Weyl semimetals, *C. R. Phys.* **14**, 857 (2013).
- [16] T. Morimoto and N. Nagaosa, Topological nature of nonlinear optical effects in solids, *Sci. Adv.* **2**, e1501524 (2016).
- [17] K. Taguchi, T. Imaeda, M. Sato, and Y. Tanaka, Photovoltaic chiral magnetic effect in Weyl semimetals, *Phys. Rev. B* **93**, 201202(R) (2016).
- [18] C.-K. Chan, N. H. Lindner, G. Refael, and P. A. Lee, Photocurrents in Weyl semimetals, *Phys. Rev. B* **95**, 041104(R) (2017).
- [19] F. de Juan, A. G. Grushin, T. Morimoto, and J. E. Moore, Quantized circular photogalvanic effect in Weyl semimetals, *Nat. Commun.* **8**, 15995 (2017).
- [20] Y. Zhang, H. Ishizuka, J. van den Brink, C. Felser, B. Yan, and N. Nagaosa, Photogalvanic effect in Weyl semimetals from first principles, *Phys. Rev. B* **97**, 241118(R) (2018).
- [21] Y. Zhang, T. Holder, H. Ishizuka, F. de Juan, N. Nagaosa, C. Felser, and B. Yan, Switchable magnetic bulk photovoltaic effect in the two-dimensional magnet CrI₃, *Nat. Commun.* **10**, 3783 (2019).
- [22] S. D. Ganichev, E. L. Ivchenko, V. V. Bel'kov, S. A. Tarasenko, M. Sollinger, D. Weiss, W. Wegscheider, and W. Prettl, Spin-galvanic effect, *Nature (London)* **417**, 153 (2002).
- [23] S. D. Ganichev, V. V. Bel'kov, S. A. Tarasenko, S. N. Danilov, S. Giglberger, C. Hoffmann, E. L. Ivchenko, D. Weiss, W. Wegscheider, C. Gerl, D. Schuh, J. Stahl, J. de Boeck, G. Borghs, and W. Prettl, Zero-bias spin separation, *Nat. Phys.* **2**, 609 (2006).
- [24] Z. Sun, Y. Yi, T. Song, G. Clark, B. Huang, Y. Shan, S. Wu, D. Huang, C. Gao, Z. Chen, M. McGuire, T. Cao, D. Xiao, W.-T. Liu, W. Yao, X. Xu, and S. Wu, Giant nonreciprocal second-harmonic generation from antiferromagnetic bilayer CrI₃, *Nature (London)* **572**, 497 (2019).
- [25] S. M. Young, F. Zheng, and A. M. Rappe, First-Principles Calculation of the Bulk Photovoltaic Effect in Bismuth Ferrite, *Phys. Rev. Lett.* **109**, 236601 (2012).
- [26] T. Morimoto, S. Zhong, J. Orenstein, and J. E. Moore, Semi-classical theory of nonlinear magneto-optical responses with applications to topological Dirac/Weyl semimetals, *Phys. Rev. B* **94**, 245121 (2016).
- [27] A. M. Cook, B. M. Fregoso, F. de Juan, S. Coh, and J. E. Moore, Design principles for shift current photovoltaics, *Nat. Commun.* **8**, 14176 (2017).
- [28] L. Z. Tan, F. Zheng, S. M. Young, F. Wang, S. Liu, and A. M. Rappe, Shift current bulk photovoltaic effect in polar materials—hybrid and oxide perovskites and beyond, *npj Comput. Math.* **2**, 16026 (2016).
- [29] P. Olbrich, S. A. Tarasenko, C. Reitmaier, J. Karch, D. Plohm, Z. D. Kvon, and S. D. Ganichev, Observation of the orbital circular photogalvanic effect, *Phys. Rev. B* **79**, 121302(R) (2009).
- [30] H. Yuan, X. Wang, B. Lian, H. Zhang, X. Fang, B. Shen, G. Xu, Y. Xu, S.-C. Zhang, H. Y. Hwang, and Y. Cui, Generation and electric control of spin-valley-coupled circular photogalvanic current in WSe₂, *Nat. Nanotechnol.* **9**, 851 (2014).
- [31] J. W. McIver, D. Hsieh, H. Steinberg, P. Jarillo-Herrero, and N. Gedik, Control over topological insulator photocurrents with light polarization, *Nat. Nanotechnol.* **7**, 96 (2012).
- [32] K. N. Okada, N. Ogawa, R. Yoshimi, A. Tsukazaki, K. S. Takahashi, M. Kawasaki, and Y. Tokura, Enhanced photogalvanic current in topological insulators via Fermi energy tuning, *Phys. Rev. B* **93**, 081403(R) (2016).
- [33] E. J. König, H. Y. Xie, D. A. Pesin, and A. Levchenko, Photogalvanic effect in Weyl semimetals, *Phys. Rev. B* **96**, 075123 (2017).
- [34] L. E. Golub and E. L. Ivchenko, Circular and magnetoinduced photocurrents in Weyl semimetals, *Phys. Rev. B* **98**, 075305 (2018).
- [35] F. de Juan, Y. Zhang, T. Morimoto, Y. Sun, J. E. Moore, and A. G. Grushin, Difference frequency generation in topological semimetals, *Phys. Rev. Research* **2**, 012017 (2020).
- [36] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, The electronic properties of graphene, *Rev. Mod. Phys.* **81**, 109 (2009).
- [37] X. Chen, L. Fidkowski, and A. Vishwanath, Symmetry enforced non-Abelian topological order at the surface of a topological insulator, *Phys. Rev. B* **89**, 165132 (2014).
- [38] W. Chen, A. V. Andreev, and A. Levchenko, Boltzmann-Langevin theory of Coulomb drag, *Phys. Rev. B* **91**, 245405 (2015).
- [39] E. Deyo, L. E. Golub, E. L. Ivchenko, and B. Spivak, Semiclassical theory of the photogalvanic effect in non-centrosymmetric systems, *arXiv:0904.1917*.
- [40] J. E. Moore and J. Orenstein, Confinement-Induced Berry Phase and Helicity-Dependent Photocurrents, *Phys. Rev. Lett.* **105**, 026805 (2010).

- [41] 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).
- [42] T. Low, Y. Jiang, and F. Guinea, Topological currents in black phosphorus with broken inversion symmetry, *Phys. Rev. B* **92**, 235447 (2015).
- [43] S. S. Tsirkin, P. A. Puente, and I. Souza, Gyrotropic effects in trigonal tellurium studied from first principles, *Phys. Rev. B* **97**, 035158 (2018).
- [44] Y. Zhang, Y. Sun, and B. Yan, Berry curvature dipole in Weyl semimetal materials: An *ab initio* study, *Phys. Rev. B* **97**, 041101(R) (2018).
- [45] Y. Zhang, J. van den Brink, C. Felser, and B. Yan, Electrically tuneable nonlinear anomalous Hall effect in two-dimensional transition-metal dichalcogenides WTe_2 and MoTe_2 , *2D Mater.* **5**, 044001 (2018).
- [46] Q. Ma, S.-Y. Xu, H. Shen, D. MacNeill, V. Fatemi, T.-R. Chang, A. M. Mier Valdivia, S. Wu, Z. Du, C.-H. Hsu, S. Fang, Q. D. Gibson, K. Watanabe, T. Taniguchi, R. J. Cava, E. Kaxiras, H.-Z. Lu, H. Lin, L. Fu, N. Gedik, and P. Jarillo-Herrero, Observation of the nonlinear Hall effect under time-reversal-symmetric conditions, *Nature (London)* **565**, 337 (2019).
- [47] K. Kang, T. Li, E. Sohn, J. Shan, and K. F. Mak, Nonlinear anomalous Hall effect in few-layer WTe_2 , *Nat. Mater.* **18**, 324 (2019).
- [48] B. I. Sturman and V. M. Fridkin, *The Photovoltaic and Photo-refractive Effects in Noncentrosymmetric Materials* (Gordon and Breach, Philadelphia, 1992).
- [49] E. Ivchenko and G. Pikus, *Superlattices and Other Heterostructures: Symmetry and Optical Phenomena* (Springer, Berlin, 1997).
- [50] E. Ivchenko, *Optical Spectroscopy of Semiconductor Nanostructures* (Alpha Science, Harrow, UK, 2005).
- [51] B. Sturman, Ballistic and shift currents in the bulk photovoltaic effect theory, *Phys. Usp.* **63**, 407 (2020).
- [52] D. J. Passos, G. B. Ventura, J. M. Viana Parente Lopes, J. M. B. Lopez 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).
- [53] A. Taghizadeh, F. Hipolito, and T. G. Pedersen, Linear and nonlinear optical response of crystals using length and velocity gauges: Effect of basis truncation, *Phys. Rev. B* **96**, 195413 (2017).
- [54] S. A. Mikhailov, Quantum theory of the third-order nonlinear electrodynamic effects of graphene, *Phys. Rev. B* **93**, 085403 (2016).
- [55] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, Geometry and non-adiabatic response in quantum and classical systems, *Phys. Rep.* **697**, 1 (2017).
- [56] H. Isobe, S.-Y. Xu, and L. Fu, High-frequency rectification via chiral Bloch electrons, *Sci. Adv.* **6**, eaay2497 (2020).
- [57] Z. Z. Du, C. M. Wang, S. Li, H.-Z. Lu, and X. C. Xie, Disorder-induced nonlinear Hall effect with time-reversal symmetry, *Nat. Commun.* **10**, 3047 (2019).
- [58] E. J. König, M. Dzero, A. Levchenko, and D. A. Pesin, Gyrotropic Hall effect in Berry-curved materials, *Phys. Rev. B* **99**, 155404 (2019).
- [59] M.-C. Chang and Q. Niu, Berry phase, hyperorbits, and the Hofstadter spectrum: Semiclassical dynamics in magnetic Bloch bands, *Phys. Rev. B* **53**, 7010 (1996).
- [60] G. Sundaram and Q. Niu, Wave-packet dynamics in slowly perturbed crystals: Gradient corrections and Berry-phase effects, *Phys. Rev. B* **59**, 14915 (1999).
- [61] O. Matsyshyn and I. Sodemann, The Nonlinear Hall Acceleration and the Quantum Rectification Sum Rule, *Phys. Rev. Lett.* **123**, 246602 (2019).
- [62] Y. Gao, Semiclassical dynamics and nonlinear charge current, *Front. Phys.* **14**, 33404 (2019).
- [63] Y. Gao, S. A. Yang, and Q. Niu, Field Induced Positional Shift of Bloch Electrons and Its Dynamical Implications, *Phys. Rev. Lett.* **112**, 166601 (2014).
- [64] J. Anandan and Y. Aharonov, Geometry of Quantum Evolution, *Phys. Rev. Lett.* **65**, 1697 (1990).
- [65] Y.-Q. Ma, S. Chen, H. Fan, and W.-M. Liu, Abelian and non-Abelian quantum geometric tensor, *Phys. Rev. B* **81**, 245129 (2010).
- [66] R. Cheng, Quantum geometric tensor (Fubini-study metric) in simple quantum system: A pedagogical introduction, [arXiv:1012.1337](https://arxiv.org/abs/1012.1337).
- [67] T. Ozawa and N. Goldman, Extracting the quantum metric tensor through periodic driving, *Phys. Rev. B* **97**, 201117(R) (2018).
- [68] D.-J. Zhang, Q.-h. Wang, and J. Gong, Quantum geometric tensor in PT-symmetric quantum mechanics, *Phys. Rev. A* **99**, 042104 (2019).
- [69] C. Aversa and J. E. Sipe, Nonlinear optical susceptibilities of semiconductors: Results with a length-gauge analysis, *Phys. Rev. B* **52**, 14636 (1995).
- [70] D. Parker (private communication).
- [71] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Third order optical nonlinearity of graphene, *New J. Phys.* **16**, 053014 (2014).
- [72] P. A. Lee and T. V. Ramakrishnan, Disordered electronic systems, *Rev. Mod. Phys.* **57**, 287 (1985).
- [73] X. Yang, K. Burch, and Y. Ran, Divergent bulk photovoltaic effect in Weyl semimetals, [arXiv:1712.09363](https://arxiv.org/abs/1712.09363).
- [74] P. Tang, Q. Zhou, G. Xu, and S.-C. Zhang, Dirac fermions in an antiferromagnetic semimetal, *Nat. Phys.* **12**, 1100 (2016).