Exploring the equivalence of causality-based and quantum mechanics-based sum rules for harmonic generation in nonlinear optical materials

Theodoros T. Koutserimpas,[1,3] Hao Li,[2] Owen D. Miller,[2] and Francesco Monticone[1,4]

[1] *School of Electrical and Computer Engineering, Cornell University, Ithaca, New York 14853, USA*
[2] *Department of Applied Physics and Energy Sciences Institute, Yale University, New Haven, Connecticut 06511, USA*
[3] *tkoutserimpas@cornell.edu*
[4] *francesco.monticone@cornell.edu*

Abstract: The Kramers-Kronig relations and various oscillator strength sum rules represent strong constraints on the physical response of materials. In this work, taking inspiration from the well-established equivalence between f-sum rules and Thomas-Reiche-Kuhn sum rules in linear optics, we explore the connection between causality-based and quantum-mechanics-based sum rules in the context of nonlinear optical processes. Specifically, by considering the sum-over-states expression for the second harmonic generation susceptibility, we deduce a new representation basis for the imaginary part of this susceptibility and we use it to derive, from causality-based integral sum rules, a new set of discrete sum rules that the transition dipole moments must satisfy. As in the case of the Thomas-Reiche-Kuhn sum rules, we also show that these results can alternatively be derived through an independent quantum mechanical analysis. Finally, we consider the implications of the derived sum rules for the second-harmonic-generation susceptibility of two- and three-level systems and, more broadly, we discuss the possible significance and challenges of using these results for the goal of identifying fundamental limits to the response of nonlinear optical materials.

1. Introduction

Any physical response must satisfy the principle of causality [1]. In linear optics, this principle dictates that the electric susceptibility of a material, $\chi^{(1)}(\omega)$, is holomorphic in the upper half of the complex $\omega-$plane, where ω is the angular frequency of the electromagnetic excitation. As a result of this property, and assuming the susceptibility is sufficiently well-behaved (square integrable), $\text{Re}\{\chi^{(1)}(\omega)\}$ and $\text{Im}\{\chi^{(1)}(\omega)\}$ are connected through integral relations known as the Kramers-Kronig relation [2]. However, causality does not always lead to analyticity in nonlinear optics, as seen for example in the nonlinear process of self-induced change in refractive index [3,4].

Nevertheless, Kramers-Kronig relations can be established for many nonlinear optical responses of interest, including any process in which all input frequencies are fixed (corresponding to fixed pump beams, e.g.), except one (the "probe" beam) [5]. For cases in which the input frequencies are mutually dependent and varied at the same time, the Bassani-Scandolo theorem [6] gives conditions under which Kramers-Kronig relations can be derived. This theorem relies on the selection of a one-dimensional space embedded in the n-dimensional space of the interdependent frequency variables. Notably, harmonic generation, of any order, satisfies the Bassani-Scandolo theorem and Kramers-Kronig relations can be derived from the holomorphic properties of the corresponding susceptibility [5].

A prominent nonlinear process satisfying Kramers-Kronig relations is second-harmonic generation (SHG), described by a second-order susceptibility $\chi^{(2)}(\omega, \omega)$, which results in the doubling of the incident frequency of light. Dipoles of noncentrosymmetric materials can oscillate under the action of the incident field as asymmetric anharmonic oscillators, radiating at 2ω, twice the incident frequency ω, in addition to their usual radiation at ω. This nonlinear process is employed in high-power lasing at specific wavelengths, as well as various applications in microscopy, ultra-short pulse measurement, and materials characterization [7–12]. Despite being one of the most important processes in nonlinear optics, there is still limited understanding of the fundamental constraints to harmonic generation in terms of either its resonance properties (e.g., optimal number and location of resonances) or the involved transition dipole moments. The most important known constraint on transition dipole moments is the so-called Thomas-Reiche-Kuhn (TRK) (or oscillator strength) sum rule [13]. While the TRK sum rule is discrete and derived from quantum-mechanical arguments, it is equivalent to the causality-based integral f−sum rule, which relates the integral of the imaginary part of the linear susceptibility, over the entire electromagnetic spectrum, to the total electron density [5,14]. One can derive the TRK sum rule from the f−sum rule by discretizing the integral through a physically motivated oscillator representation of the linear susceptibility, as further discussed in Section (3) and in [15], where the discretized causality-based sum rule was used to bound the linear optical response of any transparent medium. The equivalence between these sum rules, derived from very different arguments, is nontrivial and surprising, with important implications for the analysis and design of new materials and metamaterials with extreme response [16].

In this work, considering the case of SHG, we explore a new connection between causality-based integral sum rules and quantum-mechanics-based discrete sum rules, unveiling new constraints on the transition dipole moments and various implications for the SHG process. This is achieved by considering the Kramers-Kronig relations enforced by the principle of causality for the SHG susceptibility, $\chi^{(2)}(\omega, \omega)$, in combination with a new representation basis for the imaginary part of the susceptibility deduced from the perturbation solution to the Schrödinger equation. The validity of the new set of discrete sum rules derived here is also confirmed by an independent quantum mechanical analysis. As in the case of TRK and f−sum rules, two very different approaches lead to identical results.

In the following, the Kramers-Kronig relations and associated integral sum rules for the SHG susceptibility are briefly reviewed in Section (2). In Section (3), a new representation basis for the imaginary part of the SHG susceptibility is established and used to derive new discrete sum rules on the transition dipole moments. In Section (4), these sum rules are related to an independent quantum mechanical analysis, further proving their validity. In Section (5), examples of two-level and three-level systems are given, discussing various implications of the derived sum rules for the SHG process in these systems. Finally, in Section (6) we summarize our results, discuss the significance of our findings, and highlight a few open questions.

2. Kramers-Kronig relations and integral sum rules for $\mathrm{Im}\left\{\chi^{(2)}(\omega, \omega)\right\}$

Scandolo and Bassani [17] presented an elegant proof showing that $\chi^{(2)}(\omega, \omega)$, $\omega^2 \chi^{(2)}(\omega, \omega)$, and $\omega^4 \chi^{(2)}(\omega, \omega)$ are holomorphic in the upper half of the complex frequency plane and exhibit an asymptotic high-frequency decay faster than ω^{-1}. As a result, several Kramers-Kronig relations for the SHG susceptibility can be derived (for brevity, only the relations for the real part of $\chi^{(2)}(\omega, \omega)$ are shown here):

$$\omega^{2m}\mathrm{Re}\left\{\chi^{(2)}(\omega, \omega)\right\} = \frac{2}{\pi}\fint_0^\infty \frac{\omega'^{2m+1}\mathrm{Im}\left\{\chi^{(2)}(\omega', \omega')\right\}}{\omega'^2 - \omega^2}d\omega', \tag{1}$$

where $m = 0, 1, 2$ and $\fint$ indicates the principal value integral. In the same work [17], the asymptotic behavior of the SHG susceptibility was also used to determine the following integral

sum rules:

$$\int_0^\infty \omega \text{Im}\left\{\chi^{(2)}(\omega, \omega)\right\} d\omega = 0, \tag{2}$$

$$\int_0^\infty \omega^3 \text{Im}\left\{\chi^{(2)}(\omega, \omega)\right\} d\omega = 0, \tag{3}$$

$$\int_0^\infty \omega^5 \text{Im}\left\{\chi^{(2)}(\omega, \omega)\right\} d\omega = -\frac{\pi}{16}\frac{e^3 N}{m^3}\left\langle\frac{\partial^3 V}{\partial x^3}\right\rangle_0, \tag{4}$$

where N represents the electron density, e is the electron charge, m is the electron mass, $V(x)$ denotes the potential experienced by the electrons, and the averaging of the third derivative in Eq. (4) is conducted over the ground state of the system. Interestingly, it is clear that the imaginary part of $\chi^{(2)}(\omega, \omega)$ has to change sign along the electromagnetic spectrum in order to satisfy Eqs. (2)–(4). In contrast to the imaginary part of the linear susceptibility, however, $\text{Im}\left\{\chi^{(2)}(\omega, \omega)\right\}$ does not directly correlate with optical absorption or gain. Instead, it denotes a phase relationship between nonlinear polarization and applied fields, without necessarily implying any time-averaged absorbed or gained power [18].

In addition to Eqs. (1)–(4), Kramers-Kronig relations that calculate $\text{Im}\left\{\chi^{(2)}(\omega, \omega)\right\}$ from $\text{Re}\left\{\chi^{(2)}(\omega, \omega)\right\}$ can also be obtained, and an extra set of sum rules for $\text{Re}\left\{\chi^{(2)}(\omega, \omega)\right\}$ was also established in [17]. These theoretical results have been used to assist in the analysis of experimental data, facilitating the connection between the phase and amplitude of the susceptibilities [19,20], and providing insight into the possible presence of additional contributions outside the measured frequency range.

3. Basis for $\text{Im}\left\{\chi^{(2)}(\omega, \omega)\right\}$ and new discrete sum rules

A perturbative solution to the Schröndiger equation gives the well-established sum-over-states expression for the nonlinear SHG susceptibility [21,22]

$$\chi^{(2)}(\omega, \omega) = \frac{Ne^3}{\hbar^2}\sum_{n,m}{}'\left[\frac{x_{0n}\bar{x}_{nm}x_{m0}}{(\omega_{n0} - 2\omega - i\gamma_n)(\omega_{m0} - \omega - i\gamma_m)} + \frac{x_{0n}\bar{x}_{nm}x_{m0}}{(\omega_{n0} + \omega + i\gamma_n)(\omega_{m0} + 2\omega + i\gamma_m)} + \frac{x_{0n}\bar{x}_{nm}x_{m0}}{(\omega_{n0} + \omega + i\gamma_n)(\omega_{m0} - \omega - i\gamma_m)}\right], \tag{5}$$

where x represents the position operator, $x_{nm} = \langle n|x|m\rangle$ denotes its (n, m) element, $\bar{x} = x - x_{00}\mathbb{I}$ (i.e., $\bar{x}$ is the same as the position operator for its off-diagonal elements and the difference of the position operator with x_{00} for its diagonal elements), and the notation $\sum'$ indicates that the ground state is excluded from the summation over the states, i.e., the summation is conducted solely over the excited states. The energy level differences are given by $E_{n0} = \hbar\omega_{n0} = \hbar(\omega_n - \omega_0)$. As usually done, since the dipole moment of the molecule is proportional to the position ($\mu = -ex$), we refer to x_{nm} as the transition dipole moments and to x_{nn} as the dipole moment of the excited state. γ_n is a linewidth parameter (decay rate) that phenomenologically models various broadening/damping mechanisms [4]. We assume that the system is initially in its ground state, denoted as $|0\rangle$.

Following [23], using fractional decomposition, one can then rewrite Eq. (5) in a more compact, yet still general, form:

$$\chi^{(2)}(\omega, \omega) = \sum_n{}'\left[\frac{\alpha_1^{(n)}}{(\omega_{n0} - \omega - i\gamma_n)} + \frac{\alpha_2^{(n)}}{(\omega_{n0} - 2\omega - i\gamma_n)}\right] + \sum_{n'}{}'\frac{\alpha_3^{(n')}}{(\omega_{n'0} - \omega - i\gamma_{n'})^2} + (\omega \to -\omega)^*. \tag{6}$$

Here, n' denotes doubly resonant conditions where either $\omega_{mn'} = \omega_{n'0}$ or $\omega_{m0} = \omega_{0n'}$ (note that, if "0" denotes the ground state, the latter resonant condition would correspond to a negative

transition frequency, and hence to an "antiresonance" that can only be induced in an active system; however, our analysis will be limited to passive systems). $(\omega \rightarrow -\omega)^*$ denotes the sum of the negative-frequency and complex-conjugate functions as calculated from the first three terms of (6). The coefficients $\alpha_1^{(n)}$ and $\alpha_2^{(n)}$ are the nonlinear oscillator strengths associated with single- and two-photon resonances at ω_{n0}, respectively, and $\alpha_3^{(n')}$ is the oscillator strength associated with the specific condition of a double-resonance. These nonlinear oscillator strengths are given by

$$\alpha_1^{(n)} = \frac{Ne^3}{\hbar^2} \sum_m{}' \frac{\mathbb{1}_{\omega_{mn} \neq \omega_{n0}}}{(\omega_{m0} - 2\omega_{n0})} x_{0m}\bar{x}_{mn}x_{n0} + \frac{Ne^3}{\hbar^2} \sum_m{}' \frac{\mathbb{1}_{\omega_{m0} \neq -\omega_{n0}}}{(\omega_{m0} + \omega_{n0})} x_{0m}\bar{x}_{mn}x_{n0}, \tag{7}$$

$$\alpha_2^{(n)} = \frac{2Ne^3}{\hbar^2} \sum_m{}' \frac{\mathbb{1}_{\omega_{nm} \neq \omega_{m0}}}{(2\omega_{m0} - \omega_{n0})} x_{0n}\bar{x}_{nm}x_{m0}, \tag{8}$$

$$\alpha_3^{(n')} = \frac{Ne^3}{2\hbar^2} \mathbb{1}_{\omega_{mn'} = \omega_{n'0}} x_{0m}\bar{x}_{mn'}x_{n'0} - \frac{Ne^3}{\hbar^2} \mathbb{1}_{\omega_{m0} = -\omega_{n'0}} x_{0m}\bar{x}_{mn'}x_{n'0}. \tag{9}$$

The algebraic manipulation giving Eq. (6) from Eq. (5) was initially introduced in [23], albeit with the goal to study a single doubly resonant system (namely, with only one term for each of the summations in (6)). Instead, here we are interested in how the general form of Eq. (6) suggests a possible basis to discretize the integral sum rules discussed in Section (2). To this end, in the limit of vanishing linewidth, by inspecting Eq. (6) one can deduce the following expression for the imaginary part of the nonlinear susceptibility,

$$\text{Im}\{\chi^{(2)}(\omega, \omega)\} = \sum_n{}' \lim_{\gamma_n \rightarrow 0^+} \left(\text{Re}\{\alpha_1^{(n)}\} f_{\gamma_n}(\omega - \omega_{n0}) + \frac{\text{Re}\{\alpha_2^{(n)}\}}{2} f_{\gamma_n/2}(\omega - \omega_{n0}/2) \right)$$
$$+ \sum_{n'}{}' \lim_{\gamma_{n'} \rightarrow 0^+} \left(\text{Re}\{\alpha_3^{(n')}\} f'_{\gamma_{n'}}(\omega - \omega_{n'0}) \right), \tag{10}$$

where $f_{\gamma_n}(\omega)$ is the Lorentzian distribution: $f_{\gamma_n}(\omega) = \frac{\gamma_n}{\omega^2 + \gamma_n^2}$, and f'_γ is its frequency derivative. Note that the $1/2$ weight and the modified broadening $\gamma_n/2$ of the Lorentzian associated with two-photon resonances is in full agreement with the permutation symmetry of the nonlinear susceptibility [4]. Additionally, it is easily found that $\lim_{\gamma_n \rightarrow 0^+} f_{\gamma_n, \gamma_n/2}(\omega) = \pi\delta(\omega)$, (see, e.g., [24]) and $\lim_{\gamma_n \rightarrow 0^+} f'_{\gamma_n}(\omega) = \pi\delta'(\omega)$, where δ is the Dirac–δ distribution and δ' its frequency derivative. One can then verify that substituting (10) into the Kramers-Kronig relation (1) yields (6) for the case of $\gamma_n = 0$. We also note that, in writing Eq. (10), we made the assumption that only the real part of the coefficients $\alpha_1^{(n)}$, $\alpha_2^{(n)}$ and $\alpha_3^{(n')}$ enter this expression, essentially neglecting the effect of the transition dipole moments' phases on the imaginary part of the susceptibility. Rather surprisingly, while this assumption is strictly valid only in the off-resonance case, $\omega \ll \omega_{n0}$ (which is important, per se, in many ultrafast applications), it ultimately leads to sum rules that are completely general, as proven by an independent quantum mechanical analysis in Section (4).

The considered expression for the imaginary part of the SHG susceptibility, in the limit of vanishing linewidth, involves delta functions and their frequency derivative. Specifically, there are two delta functions for each resonance frequency, ω_{n0}, corresponding to single-photon and two-photon resonances, as well as the possible inclusion of the derivative of a delta function for the doubly resonant case. This is intriguingly different from the linear case, where the imaginary part of the linear susceptibility, in the limit of zero loss, can be expressed as a sum of only delta functions [14]. In [15], this fact was used to derive a general representation for the linear susceptibility by discretizing the Kramers-Kronig relation for the real part of $\chi^{(1)}$ using delta functions as localized basis functions for the imaginary part of $\chi^{(1)}$. Since the coefficients of this representation are proportional to the linear oscillator strengths, it was then possible to derive

an upper bound for $\text{Re}\left\{\chi^{(1)}\right\}$ by using the $f-$sum rule. While the same mathematical trick, using a basis of delta functions, could be applied to discretize the nonlinear Kramers-Kronig relation in Eq. (1) (any collocation methods with localized basis functions would work), this would not give much insight, as the resulting coefficients of this representation would not be directly relatable to the nonlinear oscillator strengths and transition dipole moments, in contrast with the linear case. In fact, the general expression for the imaginary part of $\chi^{(2)}(\omega, \omega)$, given by Eq. (10), suggests that a more natural choice for localized basis functions for $\chi^{(2)}(\omega, \omega)$ should include both delta functions and their derivatives, such that the coefficients resulting from this discretization would now be directly related to the nonlinear oscillator strengths, as shown by Eqs. (7),(8),(9). Most importantly, this approach can also yield new sum rules on the transition dipole moments. Specifically, inserting the basis implied by Eq. (10) into the causality-based integral sum rules in Eqs. (2)–(4) yields a discretized version of these sum rules in terms of the nonlinear oscillator strengths,

$$\sum_n{}' \left(\omega_{n0}\text{Re}\{\alpha_1^{(n)}\} + \frac{\omega_{n0}}{4}\text{Re}\{\alpha_2^{(n)}\} \right) - \sum_{n'}{}' \text{Re}\{\alpha_3^{(n')}\} = 0, \tag{11}$$

$$\sum_n{}' \left(\omega_{n0}^3\text{Re}\{\alpha_1^{(n)}\} + \frac{\omega_{n0}^3}{16}\text{Re}\{\alpha_2^{(n)}\} \right) - \sum_{n'}{}' 3\omega_{n'0}^2\text{Re}\{\alpha_3^{(n')}\} = 0, \tag{12}$$

$$\sum_n{}' \left(\omega_{n0}^5\text{Re}\{\alpha_1^{(n)}\} + \frac{\omega_{n0}^5}{64}\text{Re}\{\alpha_2^{(n)}\} \right) - \sum_{n'}{}' 5\omega_{n'0}^4\text{Re}\{\alpha_3^{(n')}\} = -\frac{e^3N}{16m^3}\left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0. \tag{13}$$

While passivity implies that the transition frequencies are positive, $\omega_{n0}>0$, for all energy levels, it does not constrain the real part of the nonlinear oscillator strengths $\text{Re}\{\alpha\}$ to take only positive values, in contrast with the linear case where the oscillator strengths are always positive. If we then combine Eqs. (7)–(9) with (11)–(13) and after some algebraic manipulations, we obtain a new set of discrete sum rules for the transition dipole moments:

$$\sum_{n,m}{}' \frac{\omega_{m0} - \omega_{n0}}{\omega_{m0} + \omega_{n0}} \text{Re}\{x_{0m}\bar{x}_{mn}x_{n0}\} = 0, \tag{14}$$

$$\sum_{n,m}{}' \frac{\omega_{m0}^3 + 3\omega_{m0}^2\omega_{n0} + 6\omega_{n0}^2\omega_{m0} - 4\omega_{n0}^3}{\omega_{m0} + \omega_{n0}} \text{Re}\{x_{0m}\bar{x}_{mn}x_{n0}\} = 0, \tag{15}$$

$$\sum_{n,m}{}' \frac{\omega_{m0}^5 + 3\omega_{m0}^4\omega_{n0} + 6\omega_{m0}^3\omega_{n0}^2 + 12\omega_{m0}^2\omega_{n0}^3 + 24\omega_{n0}^4\omega_{m0} - 16\omega_{n0}^5}{\omega_{m0} + \omega_{n0}} \text{Re}\{x_{0m}\bar{x}_{mn}x_{n0}\} = \frac{2\hbar^2}{m^3}\left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0. \tag{16}$$

Although (14) is automatically fulfilled (all $n = m$ terms are zero, whereas for $n \neq m$ the pairs (n, m) and (m, n) cancel each other), the second and third sum rules (15), (16) provide a set of new constraints that the transition dipole moments must satisfy to be consistent, ultimately, with the principle of causality, provided that our discretization approach is justified. In the next section, we show that the new sum rules (15), (16) can, in fact, be derived from an independent quantum mechanical analysis, confirming the deep connection between causality-based and quantum-mechanics-based constraints also in the nonlinear optical case, and proving that our proposed basis to represent the imaginary part of the susceptibility is a physically sensible choice to discretize causality-based integral relations.

4. Independent quantum mechanical derivation

In this Section, we connect the derived sum rules, as established by the principle of causality and the sum-over-states expression of the SHG susceptibility, with an independent quantum-mechanical analysis. To this end, we employ generalized quantum mechanical sum rules, based on the operator theory of quantum mechanics.

4.1. Quantum mechanical analysis of sum rule (15)

Sum rule (15) can be alternatively obtained from the generalized Thomas-Reiche-Kuhn (TRK) sum rules, which were derived for the calculation of fundamental limits to the off-resonance nonlinear optical response [25–28] and were additionally used for predicting the frequency dispersion of the hyperpolarizabilities spectrum [29] and, very recently, for the derivation of fundamental bounds on resonant nonlinear optical responses in multi-quantum-well systems [30]. These generalized TRK sum rules can be written as

$$\sum_n \left[\omega_n - \tfrac{1}{2}(\omega_l + \omega_q) \right] x_{ln} x_{nq} = \tfrac{Z\hbar}{2m} \delta_{lq}, \tag{17}$$

where δ_{lq} is the Kronecker delta function and Z is the number of electrons in the system. We take $q = 0$ and consider all possible values of $l \in \mathbb{Z}^+ \backslash \{0\}$ in (17). By multiplying both sides by $\omega_{n0} x_{n0}$ and summing over all $n \in \mathbb{Z}^+ \backslash \{0\}$, we obtain

$$\sum_n{}' \omega_{n0}^2 x_{0n} \bar{x}_{nn} x_{n0} + \sum_{n,l \neq n}{}' (4\omega_{n0}\omega_{l0} - \omega_{n0}^2 - \omega_{l0}^2) x_{0n} x_{nl} x_{l0} = 0. \tag{18}$$

Summing (18) with its index-interchanged expression, and taking into account that $x_{ij} = x_{ji}^*$, gives

$$\sum_n{}' \omega_{n0}^2 x_{0n} \bar{x}_{nn} x_{n0} + \sum_{n,l \neq n}{}' (4\omega_{n0}\omega_{l0} - \omega_{n0}^2 - \omega_{l0}^2) \mathrm{Re}(x_{0n} x_{nl} x_{l0}) = 0. \tag{19}$$

Equation (19) is simply a rearrangement of (15), as it can be seen by separating the $n = m$ term from the other terms of the sum in (15).

4.2. Quantum mechanical analysis of sum rule (16)

The independent derivation of sum rule (16) is more involved. First, we consider the following commutation relations between momentum and Hamiltonian: $[p, H] = -i\hbar \frac{\partial V}{\partial x}$, $[p, [p, H]] = (-i\hbar)^2 \frac{\partial^2 V}{\partial x^2}$ and $[p, [p, [p, H]]] = (-i\hbar)^3 \frac{\partial^3 V}{\partial x^3}$, where $V(x)$ is the potential function of the system. Additionally, $[p, [p, [p, H]]] = -3p[p, H]p$.

Similar to [31], we consider a generalized sum rule:

$$\sum_{l,m} (E_l - E_m) \langle 0|p|l \rangle \langle l|p|m \rangle \langle m|p|0 \rangle$$
$$= \sum_{l,m} \langle 0|p|l \rangle \langle l|[H,p]|m \rangle \langle m|p|0 \rangle \tag{20}$$
$$= \langle 0|p[H,p]p|0 \rangle.$$

Then, since $[x, H] = \left[x, \frac{p^2}{2m} \right] = \frac{i\hbar}{m} p$, we have: $\langle n|[x, H]|m \rangle = x_{nm} E_{mn} = \frac{i\hbar}{m} p_{nm}$. Hence, we get $\langle 0| [p, [p, [p, H]]] |0 \rangle = (-i\hbar)^3 \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0$ and

$$\frac{\hbar^2}{m^3} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0 = 3 \sum_{l,m} \omega_{lm}^2 \omega_{l0} \omega_{m0} \mathrm{Re}(x_{0l} x_{lm} x_{m0}). \tag{21}$$

Our objective here is to demonstrate the equivalence between (16) and (21). To achieve this, we perform the following algebraic manipulations to (16). The sum rule (16) can be expressed as:

$$\frac{2\hbar^2}{m^3} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0 = \frac{1}{2} \left(\sum_{n,m \neq n}{}' A_{mn} + \sum_n{}' B_{nn} + \sum_m{}' B_{mm} \right), \tag{22}$$

where $A_{nm} = (-15\omega_{m0}^4 - 15\omega_{n0}^4 + 42\omega_{m0}^3 \omega_{n0} + 42\omega_{n0}^3 \omega_{m0} - 24\omega_{m0}^2 \omega_{n0}^2) \mathrm{Re}(x_{0m} \bar{x}_{mn} x_{n0})$ and $B_{nn} = 15\omega_{n0}^4 x_{0n} \bar{x}_{nn} x_{n0}$.

Again using similar algebraic manipulations as before, we take the TRK sum rules (17), for $q = 0$ and all possible values of $l \in \mathbb{Z}^+ \backslash \{0\}$ and multiply both sides by $\omega_{n0}^3 x_{n0}$. We denote $C_{ln} = -15\omega_{n0}^3 (\omega_{l0} + \omega_{ln}) \text{Re}(x_{0l} x_{ln} x_{n0})$, so that $B_{nn} = \sum_{l \neq n} C_{ln}$. Then (22) becomes:

$$\frac{2\hbar^2}{m^3} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0 = \frac{1}{2} \sum_{n,m \neq n} {}' (A_{mn} + C_{mn} + C_{nm}), \tag{23}$$

where each of the summed terms of (23) can be simplified to $A_{mn} + C_{mn} + C_{nm} = 12\omega_{n0}\omega_{m0}\omega_{nm}^2$ $\text{Re}(x_{0n} x_{nm} x_{m0})$. The equivalence of (21) with (23) is then clear.

Rather strikingly, these results show that, just as in the linear case for the standard TRK sum rule, these new sum rules can be derived independently from either a quantum mechanical analysis or from causality considerations combined with a suitable discretization of the relevant integrals to connect the discretization coefficients to the transition dipole moments (10). We stress that, although the integral sum rules have been derived from the causality properties of the SHG susceptibility, their discrete form in terms of transition dipole moments, which rely on generic quantum mechanical operator theory [see Eqs. (17) and (21)], is general and not limited to any specific nonlinear process. Some intriguing consequences of these sum rules are discussed next.

5. Examples

5.1. Two-level approximation

It is often convenient to approximate the nonlinear response of a material using a two-level system [4]. In this case, the system can be described in its simplest form as having only two levels: a ground state and an excited state. Despite its simplicity, this model has proven valuable for understanding susceptibility trends in many-level systems [22]. In a many-level system, the transition dipole moment to the first excited state can reach its maximum value when transitions to all states beyond the first excited state are zero, per Eq. (17) (when p and q are both zero, the sum rule yields: $|x_{10}|^2 = \frac{\hbar}{2m\omega_{10}} Z - \sum_{n=2}^{\infty} \frac{\omega_{n0}}{\omega_{10}} |x_{n0}|^2$, which is maximized when the second term on the right-hand-side is zero). However, this does not necessarily imply that a two-level system exhibits the strongest possible nonlinear response compared to higher-level systems, as the nonlinear oscillator strengths are not simply proportional to $|x_{n0}|^2$ [see Eqs. (7)–(9)] and are not constrained by sign restrictions.

Under the two–level approximation and assuming negligible linewidth/damping, the SHG susceptibility is given by [22]

$$\chi^{(2)}(\omega, \omega) = \frac{3Ne^3}{\hbar^2} \frac{\omega_{10}^2 \bar{x}_{11} |x_{10}|^2}{(\omega_{10}^2 - \omega^2)(\omega_{10}^2 - 4\omega^2)}. \tag{24}$$

Interestingly, using the sum rule (16), we can relate the transition dipole moments to the expected value of the third derivative of the potential function: $\bar{x}_{11}|x_{10}|^2 = \frac{2\hbar^2}{15m^3\omega_{10}^4} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0$. This allows us to determine a new expression for the SHG susceptibility within the two-level approximation:

$$\chi^{(2)}(\omega, \omega) = \frac{2Ne^3}{5m^3\omega_{10}^2} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0 \frac{1}{(\omega_{10}^2 - \omega^2)(\omega_{10}^2 - 4\omega^2)}. \tag{25}$$

This model suggests an interesting scaling law with respect to the transition frequency that is not immediately evident in an expression like Eq. (24) due to the dependence of the dipole moments on the transition frequencies. In the low-frequency limit, far from resonance, the

susceptibility scales very rapidly with respect to the transition frequency, $\chi^{(2)}(0,0) \propto \omega_{10}^{-6}$. We note however that, while the two-level approximation can be a useful simplification, such a model for the nonlinear susceptibility does not have the correct high-frequency asymptotic behavior and does not satisfy all the relevant Kramers-Kronig relations and sum rules; therefore, it cannot rigorously represent a causal SHG susceptibility and cannot be used to derive universal physical bounds and general scaling laws. Specifically, the sum rule (15) is not satisfied, and the high-frequency asymptotic behavior follows $\sim \omega^{-4}$, whereas the expected, physical asymptotic behavior for the SHG susceptibility should be $\sim \omega^{-6}$ [5,17].

5.2. Doubly resonant three-level system

A three-level system is the simplest model for which the SHG susceptibility can fully satisfy the relevant Kramers-Kronig relations, and therefore causality. In a scenario where the levels are equally spaced, resulting in resonance frequencies ω_{10} and $\omega_{20} = 2\omega_{10}$, the TRK sum rules (17) give: $\bar{x}_{22} = 0$ and $x_{01}\bar{x}_{11} = -3x_{02}x_{21}$. The nonlinear oscillator strength at ω_{20} is therefore zero, and the susceptibility doesn't present a resonance at this frequency. If the transition frequencies are positive (e.g., in a passive system), one can then identify two possible resonant conditions: one when the applied field frequency coincides with the first transition frequency $\omega = \omega_{10}$, leading to a double-resonance, as depicted in Fig. 1(a), and a second one when $\omega = \omega_{10}/2$, corresponding to a two-photon resonance, as depicted in Fig. 1(b). A plot of the absolute value of the imaginary part of the SHG susceptibility for a doubly resonant three-level system is depicted in Fig. 1(c). This consists of two Lorentzian functions at the transition frequency, ω_{10}, and at half this frequency, $\omega_{10}/2$, as well as the derivative of a Lorentzian function at ω_{10}, corresponding to the double resonance phenomenon, consistent with our discussion in Section 3.
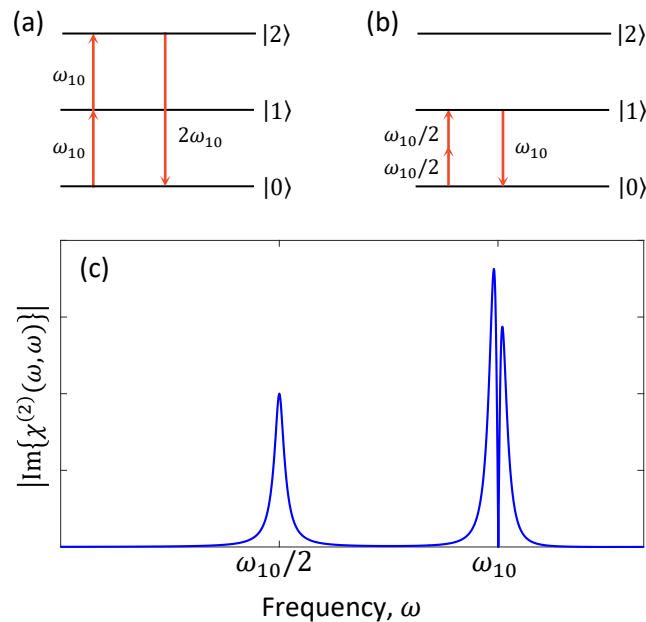

Fig. 1. (a) Schematic representation of level transitions related to the double-resonance at $\omega = \omega_{10}$. (b) Schematic representation of level transitions related to the two-photon resonance at $\omega = \omega_{10}/2$. (c) Absolute value of the imaginary part of the second-harmonic-generation susceptibility for a doubly resonant three–level system.

For a doubly resonant three-level system with vanishing linewidth, the SHG susceptibility given by Eq. (6) can then be simplified to:

$$\chi^{(2)}(\omega, \omega) = \frac{\alpha_1}{\omega_{10} - \omega} + \frac{\alpha_2}{\omega_{10} - 2\omega} + \frac{\alpha_3}{(\omega_{10} - \omega)^2} + (\omega \to -\omega)^*, \tag{26}$$

where $\alpha_1 = -\frac{Ne^3}{2\hbar^2} \frac{|x_{01}|^2 \bar{x}_{11}}{\omega_{10}} + \frac{Ne^3}{3\hbar^2} \frac{x_{01}x_{12}x_{20}}{\omega_{10}}$, $\alpha_2 = \frac{2Ne^3}{\hbar^2} \frac{|x_{01}|^2 \bar{x}_{11}}{\omega_{10}} + \frac{2Ne^3}{3\hbar^2} \frac{x_{01}x_{12}x_{20}}{\omega_{10}}$ and $\alpha_3 = \frac{Ne^3}{2\hbar^2} x_{02}x_{21}x_{10}$. One can then use the sum rule (16) to obtain: $|x_{01}|^2 \bar{x}_{11} = -\frac{\hbar^2}{4m^3\omega_{10}^4} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0$ and $x_{02}x_{21}x_{10} = \frac{\hbar^2}{12m^3\omega_{10}^4} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0$. The SHG susceptibility can then be written in terms of the linear susceptibility for a single Lorentzian resonance, $\chi^{(1)}(\omega)$, as first shown in [23],

$$\begin{aligned}
\chi^{(2)}(\omega, \omega) &= -\frac{Ne^3}{2m^3} \left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0 \frac{1}{(\omega_{10}^2 - \omega^2)^2 (\omega_{10}^2 - 4\omega^2)} \\
&= -\frac{\left\langle \frac{\partial^3 V}{\partial x^3} \right\rangle_0}{2N^2 e^3} \chi^{(1)}(2\omega) \chi^{(1)}(\omega) \chi^{(1)}(\omega),
\end{aligned} \tag{27}$$

which provides an analytical expression for the empirical relation between nonlinear and linear susceptibilities known as Miller's rule [32] and also gives a specific definition for the proportionality constant (Miller's constant). The derived Miller's constant is consistent with findings from the anharmonic oscillator model, where the expected value of the third derivative of the potential at the ground level is substituted by the third derivative of the potential at the equilibrium position of a simple oscillator model [33]. We stress, however, that Miller's rule, which can be derived rigorously for a three-level system as shown here, can only provide approximate and qualitative predictions for more complex molecules and solid-state materials, since realistic materials involve more than three levels. Nevertheless, Eq. (27) is expected to be approximately valid far from the transition frequencies, for example in the low-frequency regime, where it predicts the same scaling law as the two-level model, $\chi^{(2)}(0, 0) \propto \omega_{10}^{-6}$. In particular, if we compare (25) with (27), the two-level system approximates well the predictions of the three-level system in the low-frequency limit $\omega \to 0$: $\left| \chi^{(2)}_{2\text{-level}} \middle/ \chi^{(2)}_{3\text{-level}} \right| = 4/5$. Figure 2 shows the absolute value of the susceptibilities calculated using Eqs. (25) and (27), both normalized in the same way. The three-level formula (27) predicts a stronger nonlinear response almost everywhere, especially near the double resonance at $\omega = \omega_{10}$, where the two-level model only predicts a single resonance. In general, the strongest SHG nonlinearity occurs in this doubly resonant scenario, with a clear trade-off between strength of nonlinearity and dispersion/bandwidth, which is similar to the linear case [15], but with this tradeoff explicitly dependent on the spatial derivatives of the potential function in addition to the electron density. Fig. 2(b) also shows that the two-level formula (25) exhibits a stronger nonlinear response in the high-frequency limit ($\omega \to \infty$). However, this result is physically incorrect, as it violates the expected, physical asymptotic behavior of the SHG susceptibility [5] and the relevant sum rules, as mentioned above.

One may also wonder whether the derived expressions, constrained by the derived sum rules, may allow establishing fundamental limits on the second-order susceptibilities of three-level systems, perhaps improving upon existing bounds [25–27], and accounting for the tradeoffs with dispersion and bandwidth mentioned above, as done in [15] for the linear susceptibility. However, as is clear from Eq. (27), bounding the nonlinear susceptibility would reduce to the challenging task of determining limitations to the expected value of the third derivative of the potential at the ground state. Moreover, it is hard in general to account for passivity constraints due to the non-positivity of the nonlinear oscillator strengths in the considered representation.
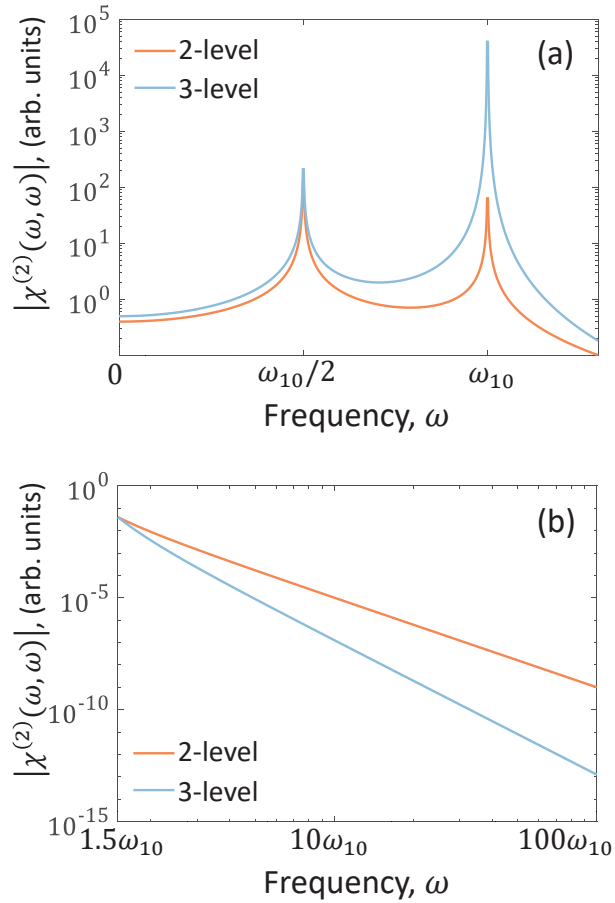
Fig. 2. Absolute value of the SHG susceptibility in logarithmic scale for two-level and three-level systems calculated using Eqs. (25) and (27), in the (a) low-frequency and near-resonance regime and (b) high-frequency regime. Both formulas predict resonances at $\omega = \omega_{10}$ and $\omega = \omega_{10}/2$ (although the two-level system does not predict a double-resonance). The three-level system results in stronger nonlinear response near-resonance and in the low-frequency limit, while the two-level system results in stronger nonlinear response in the high-frequency limit (such result is, however, physically incorrect, as discussed in the main text).

For example, in the case considered above, the nonlinear oscillator strengths can be explicitly found as $\alpha_1 = \frac{11}{72}\frac{e^3 N}{m^3 \omega_{10}^5}\left\langle \frac{\partial^3 V}{\partial x^3}\right\rangle_0$, $\alpha_2 = -\frac{4}{9}\frac{e^3 N}{m^3 \omega_{10}^5}\left\langle \frac{\partial^3 V}{\partial x^3}\right\rangle_0$ and $\alpha_3 = \frac{1}{24}\frac{e^3 N}{m^3 \omega_{10}^4}\left\langle \frac{\partial^3 V}{\partial x^3}\right\rangle_0$. Clearly, these oscillator strengths do not have the same sign and, therefore, the imaginary part of the SHG susceptibility may take both positive and negative values throughout the frequency spectrum, consistent with (2) and (3), without necessarily violating passivity.

We recently used a different approach to obtain fundamental limits to the resonant second/third-order susceptibilities of three/four-level systems, albeit without accounting for bandwidth/dispersion, using only the generalized TRK sum rules (17), which automatically encode passivity restrictions, but no bandwidth information [30]. Moreover, we note that the derivation of fundamental limits and scaling laws for higher-level systems will possibly require additional sum rules due to the increasing number of unknown nonlinear oscillator strengths. Specifically, such an analysis would necessitate sum rules that are orthogonal to the generalized TRK sum rules (17) and the

newly established SHG sum rule (16) (whereas (15) is dependent on the generalized TRK sum rules (17) as we showed in Section 4).

6. Discussion

In this paper, we have established a new "theoretical bridge" between causality-based integral sum rules for the second harmonic generation susceptibility and quantum-mechanics-based discrete sum rules that relate energy levels to transition dipole moments, the latter being valid for any linear and nonlinear optical materials. This connection is analogous to the well-established equivalence between the integral $f-$sum rule for the linear susceptibility and the Thomas-Reiche-Kuhn sum rule [5], with our theoretical results highlighting how this equivalence between causality-based and quantum-mechanics-based sum rules encompass also nonlinear optical phenomena.

The new set of sum rules presented here—derived both through a physically motivated discretization of the integral sum rules and from an independent quantum mechanical analysis—hold promise for validating measurements, theoretical models, and numerical simulations, as they provide strong constraints for the transition dipole moments and, hence, for the linear and nonlinear optical response of materials. Moreover, they may offer insights into the second-harmonic generation process across a range of scenarios, spanning from the characterization of molecular nonlinear susceptibilities and hyperpolarizabilities to the investigation of the optical responses of multi-quantum wells.

Looking ahead, we believe that an extension of our current analysis to third-order (and higher-order) harmonic generation susceptibilities is feasible, offering valuable insights into these important nonlinear optical processes and potentially unveiling new sum rules for the transition dipole moments. Moreover, an important question is whether the sum rules established in this work could be combined with other constraints to derive physically tight bounds on harmonic generation and, more generally, on the nonlinear optical response of systems with more than three levels (while in the linear case, the simplest causal system, namely, a single Lorentzian oscillator, saturating the total oscillator strength implied by the TRK sum rule, is optimal for maximizing the linear susceptibility [15], there is no guarantee that a three-level system maximizes the second-order susceptibility in all scenarios).

Another interesting question is whether the developed basis for the imaginary part of the harmonic generation susceptibility, connecting the coefficients resulting from the discretization of causality-based integral relations to nonlinear oscillator strengths, could be used to develop a causality-based electromagnetic scattering theory similar to [34] but focused on nonlinear harmonic generation processes. This approach could establish a scattering representation that inherently incorporates causality, and therefore bandwidth and dispersion properties, potentially revealing significant constraints in the mathematical structure and physical behavior of harmonically generated scattered fields.

Disclosures. The authors declare no conflicts of interest.

Data availability. The research data of this article are generated by the direct application of the derived equations. All information required to replicate the presented theoretical results is available in the main text.

References

1. J. S. Toll, "Causality and the dispersion relation: Logical foundations," Phys. Rev. **104**(6), 1760–1770 (1956).
2. H. M. Nussenzveig, *Causality and Dispersion Relations* (Academic Press, 1972).
3. A. Yariv, *Quantum Electronics* (John Wiley & Sons, 1972), 3rd ed.
4. R. W. Boyd, *Nonlinear Optics* (Academic Press, 2020), 4th ed.
5. V. Lucarini, J. J. Saarinen, K. E. Peiponen, *et al.*, *Kramers-Kronig Relations in Optical Materials Research* (Springer, 2005).

6. F. Bassani and S. Scandolo, "Dispersion relations and sum rules in nonlinear optics," Phys. Rev. B **44**(16), 8446–8453 (1991).

7. G. D. Boyd, A. Ashkin, J. M. Dziedzic, *et al.*, "Second-harmonic generation of light with double refraction," Phys. Rev. **137**(4A), A1305–A1320 (1965).

8. S. C. Kumar, G. K. Samanta, and M. Ebrahim-Zadeh, "High-power, single-frequency, continuous-wave second-harmonic-generation of ytterbium fiber laser in PPKTP and MgO:sPPLT," Opt. Express **17**(16), 13711–13726 (2009).

9. L. Moreaux, O. Sandre, and J. Mertz, "Membrane imaging by second-harmonic generation microscopy," J. Opt. Soc. Am. B **17**(10), 1685–1694 (2000).

10. E. Sidick, A. Knoesen, and A. Dienes, "Ultrashort-pulse second-harmonic generation. i. transform-limited fundamental pulses," J. Opt. Soc. Am. B **12**(9), 1704–1712 (1995).

11. E. Sidick, A. Dienes, and A. Knoesen, "Ultrashort-pulse second-harmonic generation. ii. non-transform-limited fundamental pulses," J. Opt. Soc. Am. B **12**(9), 1713–1722 (1995).

12. M. L. Ren, R. Agarwal, W. Liu, *et al.*, "Crystallographic characterization of ii–vi semiconducting nanostructures via optical second harmonic generation," Nano Lett. **15**(11), 7341–7346 (2015).

13. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer, 1957).

14. J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, 1972).

15. H. Shim, F. Monticone, and O. D. Miller, "Fundamental limits to the refractive index of transparent optical materials," Adv. Mater. **33**(43), 2103946 (2021).

16. V. Chernyak and S. Mukamel, "Generalized sum rules for optical nonlinearities of many-electron systems," J. Chem. Phys. **103**(17), 7640–7644 (1995).

17. S. Scandolo and F. Bassani, "Kramers-Kronig relations and sum rules for the second-harmonic susceptibility," Phys. Rev. B **51**(11), 6925–6927 (1995).

18. N. Bloembergen, *Nonlinear Optics* (W. A. Benjamin, Inc., 1965), 1st ed.

19. H. Kishida, T. Hasegawa, Y. Iwasa, *et al.*, "Dispersion relation in the third-order electric susceptibility for polysilane film," Phys. Rev. Lett. **70**(24), 3724–3727 (1993).

20. V. Lucarini and K.-E. Peiponen, "Verification of generalized Kramers-Kronig relations and sum rules on experimental data of third harmonic generation susceptibility on polymer," J. Chem. Phys. **119**(1), 620–627 (2003).

21. B. J. Orr and J. F. Ward, "Perturbation theory of the non-linear optical polarization of an isolated system," Mol. Phys. **20**(3), 513–526 (1971).

22. G. I. Stegeman and R. A. Stegeman, *Nonlinear Optics* (John Wiley & Sons, 2012).

23. S. Scandolo and F. Bassani, "Miller's rule and the static limit for second-harmonic generation," Phys. Rev. B **51**(11), 6928–6931 (1995).

24. V. Balakrishnan, *Mathematical Physics* (Springer, 2020).

25. M. G. Kuzyk, "Physical limits on electronic nonlinear molecular susceptibilities," Phys. Rev. Lett. **85**(6), 1218–1221 (2000).

26. M. G. Kuzyk, "Fundamental limits on third-order molecular susceptibilities," Opt. Lett. **25**(16), 1183–1185 (2000).

27. M. G. Kuzyk, "Quantum limits of the hyper-Rayleigh scattering susceptibility," IEEE J. Select. Topics Quantum Electron. **7**(5), 774–780 (2001).

28. M. G. Kuzyk, J. Perez-Moreno, and S. Shafei, "Sum rules and scaling in nonlinear optics," Phys. Rep. **529**(4), 297–398 (2013).

29. X. Hu, D. Xiao, S. Keinan, *et al.*, "Predicting the frequency dispersion of electronic hyperpolarizabilities on the basis of absorption data and Thomas-Kuhn sum rules," J. Phys. Chem. C **114**(5), 2349–2359 (2010).

30. H. Li, T. T. Koutserimpas, F. Monticone, *et al.*, "Approaching upper bounds to resonant nonlinear optical susceptibilities with inverse-designed quantum wells," ACS Photon. **2025**, 1 (2025).

31. S. Wang, "Generalization of the Thomas-Reiche-Kuhn and the Bethe sum rules," Phys. Rev. A **60**(1), 262–266 (1999).

32. R. C. Miller, "Optical second harmonic generation in piezoelectric crystals," Appl. Phys. Lett. **5**(1), 17–19 (1964).

33. F. Bassani and V. Lucarini, "General properties of optical harmonic generation from a simple oscillator model," Nuovo Cimento **20**(7-8), 1117–1125 (1998).

34. L. Zhang, F. Monticone, and O. D. Miller, "All electromagnetic scattering bodies are matrix-valued oscillators," Nat. Commun. **14**(1), 7724 (2023).