



Paradox of the many-state catastrophe of fundamental limits and the three-state conjecture

Shoresh Shafei and Mark G. Kuzyk*

Department of Physics and Astronomy, Washington State University, Pullman, Washington 99164-2814, USA

(Received 4 June 2013; published 30 August 2013)

The calculation of the fundamental limits of nonlinear susceptibilities posits that when a quantum system has a nonlinear response at the fundamental limit, only three energy eigenstates contribute to the first and second hyperpolarizability. This is called the three-level ansatz and is the only unproven assumption in the theory of fundamental limits. In light of the observation that the measured nonlinear response of a majority of molecules falls far short of these limits, the three-level ansatz warrants closer scrutiny. All calculations that are based on direct solution of the Schrödinger equation, including systems of arbitrarily placed electrons and nuclei in external electromagnetic fields and interacting electrons in an arbitrary potential, whose configuration space is sampled using numerical optimization techniques, yield intrinsic hyperpolarizabilities less than 0.709 and intrinsic second hyperpolarizabilities less than 0.6. In this work, we show that relaxing the three-level ansatz and allowing an arbitrary number of states to contribute leads to divergence of the optimized intrinsic hyperpolarizability in the limit of an infinite number of states, what we call the many-state catastrophe. This is not surprising given that the divergent systems are most likely not derivable from the Schrödinger equation, yet obey the sum rules. The sum rules are the second ingredient in limit theory and apply also to systems with more general Hamiltonians. These “exotic Hamiltonians” may not model any real systems found in nature. Indeed, a class of transition moments and energies that come from the sum rules do not have a corresponding Hamiltonian that is expressible in differential form. In this work, we show that the three-level ansatz acts as a constraint that excludes many of the nonphysical Hamiltonians and prevents the intrinsic hyperpolarizability from diverging. We argue that this implies that the true fundamental limit is smaller than previously calculated. Since the three-level ansatz does not lead to the largest possible nonlinear response, contrary to its assertion, we propose the intriguing possibility that the three-level ansatz is true for any system that obeys the Schrödinger equation, yet this assertion may be unprovable.

DOI: [10.1103/PhysRevA.88.023863](https://doi.org/10.1103/PhysRevA.88.023863)

PACS number(s): 42.65.An, 33.15.Kr

I. INTRODUCTION

The fundamental limits of the first and second hyperpolarizabilities [1–4] can be used to define scale-invariant quantities [5,6] that have been used to test the theoretical basis of the nonlinear susceptibility as well as exploring their consequences as embodied in real quantum systems. Slepikov *et al.* used the theory of fundamental limits to understand scaling of the surprising nonlinear optical properties of polyene oligomers [7]. May *et al.* used the fundamental limits as an absolute standard for comparing molecules, independent of their size, to show the promise of small molecules [8,9]. Chen *et al.* showed that the nonlinear response of certain nano-engineered polymers gave a larger intrinsic nonlinear-optical response than what would be expected from the individual building blocks [10]. The first- (β) and second-order (γ) response of twisted π -electron chromophores were identified to constitute a new paradigm for enhanced electro-optic materials [11–13] based on the fact that they fell far into the gap between the best molecules ever measured and the fundamental limit [14]. The fundamental limits have also been used to define the concept of modulated conjugation in the bridge between the donor and acceptor ends as a new paradigm for enhancing the *intrinsic hyperpolarizability*, β_{int} [15,16], defined by

$$\beta_{\text{int}} = \frac{\beta}{\beta_{\text{max}}}, \quad (1)$$

where β_{max} is the fundamental limit [1–4] of the first hyperpolarizability and is calculated using the sum rules [17–19] and the three-level ansatz [1–4].

While the fundamental limits have been used as a metric in comparing molecules, provide insights in understanding the scaling behavior of molecular homologues, and act as a guide in developing new molecular paradigms, the assumptions underlying the limits have not been scrutinized, nor have their implications been fully studied. This paper revisits the assumptions and investigates their consequences to fundamental physics. The paper is organized into four parts. First, the derivations of the limits are reviewed and the assumptions fully interpreted. Next, the behavior predicted from the limits is compared with theoretical calculations and experimental measurements of the nonlinear susceptibility. The third section demonstrates how the theory of the fundamental limits changes when the assumptions are relaxed, and their implications discussed. Finally, the last section describes hints of potentially new fundamental physics that lies just beyond our horizon of understanding.

II. FUNDAMENTAL LIMITS

The fundamental limits are calculated in three straightforward steps. First, the sum-over-states (SOS) expressions of Orr and Ward [20] for the nonlinear susceptibilities are simplified using the generalized Thomas-Reiche-Kuhn sum rules, which relate the position operator matrix elements, $x_{ij} \equiv \langle i|x|j \rangle$, and the energies, E_i , to each other:

$$\sum_n \left(E_n - \frac{E_m + E_p}{2} \right) x_{mn} x_{np} = \frac{\hbar^2 N}{2m} \delta_{mp}, \quad (2)$$

*kuz@wsu.edu

where n , m , and p index the energy eigenstates, and N the number of electrons. Then, the three-level ansatz is applied under the assumption that a three-level model accurately describes any system with a nonlinear-optical response near the fundamental limit. This leads to an expression with only two adjustable parameters, the ratio between the energies of the two states, $E = E_{10}/E_{20}$, and the matrix element of the position operator, $x = x_{10}/x_{10}^{\max}$, where x_{10}^{\max} is the upper bound of x_{10} , and it is determined from sum rules. Finally, the expression for the nonlinear susceptibility is optimized under the assumption that x and E are independent. The process can be reversed, with the three-level ansatz being applied prior to the sum rules, as we outline below.

First, we reiterate the definition of the three-level ansatz (TLA): “When the hyperpolarizability of a quantum system is at its fundamental limit, only three states contribute to the nonlinear response.” The converse is not true; that is, a quantum system in which only three states contribute to the nonlinear response will not necessarily have a large nonlinear response and may in fact be far from the limit.

The three-level ansatz is motivated by the simpler calculation of the off-resonance polarizability, given by

$$\begin{aligned}\alpha &= 2 \sum_n \frac{e^2 |x_{n0}|^2}{E_{n0}} \\ &= 2 \sum_n \frac{e^2 |x_{n0}|^2 E_{n0}}{E_{n0}^2} \leq \frac{2e^2}{E_{10}^2} \sum_n |x_{n0}|^2 E_{n0}.\end{aligned}\quad (3)$$

Using Eq. (2) with $m = p = 0$,

$$\alpha \leq \frac{e\hbar^2}{m} \frac{N}{E_{10}^2} \equiv \alpha_{\max}.\quad (4)$$

An alternative method for getting Eq. (4) is to note that Eq. (3) is a sum of positive definite terms. If the full transition strength is placed in State 1, the sum rules predict that all other transition moments x_{n0} vanish. Thus, since the numerator is made maximum by placing all the transition strength between states 0 and 1, and this denominator is the smallest of all terms, then α is maximum in a two-level model.

Consequently, one might expect that the same may be true for β ; but two problems arise. First, since the sum over states expression does not have all positive-definite terms, it is not a simple task to show that placing all the transition strength in the first excited state is the best strategy. A more damning problem is that the sum rule for the two-level model with $m = 0$ and $p = 1$ yields

$$E_{10}x_{10}(x_{11} - x_{00}) = 0,\quad (5)$$

which does not allow both x_{10} and $x_{11} - x_{00}$ to be nonzero. Since in the two-level model, $\beta \propto |x_{10}|^2(x_{11} - x_{00})$, the implication is that $\beta = 0$ for a two-level system, clearly not a maximum. However, no such problems arise for the three-level model. Since using the minimum possible number of states makes the transition strengths of the nonvanishing terms large, and if the transition energies are small; that is, the lowest energy states are the dominant ones, then the three-level model will yield the largest nonlinear response. Optimization of this expression by varying the independent physical quantities is then assumed to lead to the fundamental limit.

Since this sum rule derivation of limits appears to hold for all measurements and all exact analytical calculations, the three-level approximation for an optimized system appears reasonable. Indeed, the highly successful two-level model, used in understanding the nonlinear response of large numbers of donor or acceptor molecules by hundreds of researchers [21], has not been questioned because of its success despite the fact that such an expression violates the sum rules. The two-level model may work well for molecules with a hyperpolarizability far from the fundamental limit, where many states contribute but two dominate the response, perhaps due to slight resonance enhancement. In contrast, when the two-level model is constrained to obey the sum rules and forcing only two states to contribute while suppressing all others, the two-level model becomes unphysical.

Applying the sum rules to a three-level model yields its own set of problems. One can show that the truncation of the sum rules is perfectly legitimate when applied to a few-state SOS expression of the hyperpolarizability in the manner used in calculating the limits [22]. The problem arises in the choice of sum rules used in simplifying the SOS expression. For example, in a three-level model, the largest state index is 2 and the sum rule corresponding to Eq. (2) with $m = p = 2$ (call this sum rule Σ_{22}) is nonsensical and yields the obviously nonsensical result that

$$-\sum_n E_{2n} |x_{2n}|^2 = \frac{\hbar^2 N}{2m},\quad (6)$$

that is, a negative number equals a positive one.

In calculating the limits, the sum rule equation Σ_{22} is ignored. Using the sum rules Σ_{00} , Σ_{10} , and Σ_{20} reduces the three-level approximation of the SOS expression to one with two adjustable parameter that are varied to find the maximum. The assumption here is that all nonlinear-optical susceptibilities near the limit can be approximated using this two-parameter model.

The above approach ignores sum rule Σ_{12} . One may argue that this choice is arbitrary, as did Champagne and Kirtman [23]. For example, why not also ignore sum rule Σ_{20} . With this choice, there would be no upper bound on the hyperpolarizability. In the end, the appropriate choice is the one that is consistent with the data. Potential optimization studies [24–26], nuclear placement [27,28], application of electromagnetic fields [28], and interactions between electrons in a potential well [29] all show maximum hyperpolarizabilities of $\beta_{\text{int}} \leq 0.709$ and $\gamma_{\text{int}} \leq 0.6$ [30]. Using fewer sum rule equations makes the upper bound infinite while additional constraints decreases the limit below these values. Thus, while not a rigorous proof, these observations show that the approximations used in calculating the limit are reasonable.

In summary, the calculations of the fundamental limits uses the three-level ansatz, which has not been rigorously derived, but appears to hold over a broad set of observations. Similarly, using the sum rules Σ_{00} , Σ_{10} , and Σ_{20} , and ignoring the others is somewhat arbitrary but again yields the correct results. These assumptions are scrutinized and their implications discussed below.

III. COMPARISON OF LIMITS WITH DATA

The TLA and the neglect of the pathological sum rules are the two untested assumptions of the theory of limits. Since they cannot be tested directly, and proof with analytical techniques has been unsuccessful, an indirect approach is to sample a large enough domain of the full configuration space of quantum systems to determine if any of the predictions of the theory are violated. Are there instances where a system beats the limits? Do all systems fall substantially short of the limit? How many states are typically involved near the limits? These and related questions can be answered using analytical calculations and experimental data.

Since the hyperpolarizability is not a scale-invariant quantity, we use the *intrinsic hyperpolarizability*, β_{int} , as an absolute standard for evaluating a quantum system that is independent of the quantum unit's size [5,14,24] and spans the range $[-1, 1]$ with $|\beta_{\text{int}}| = 1$ defining the fundamental limits. The factor of 30 gap between the largest nonlinearities observed (prior to 2007) by experiment and the fundamental limits and the fact that the hyperpolarizability of the clipped harmonic oscillator is less than a factor of 2 smaller than the limit suggested that new paradigms for making better molecules and artificial quantum systems would be required to make larger intrinsic hyperpolarizabilities practical [31]. More importantly, these studies established that the fundamental limits held for all known molecules and therefore represented the true upper bound.

To understand the nature of this gap, Tripathy *et al.* used linear spectroscopy, Raman spectroscopy, β values measured by Hyper-Rayleigh scattering, and Stark spectroscopy to determine which parameters are critical [31]. Included were tests of dilution effects due to vibronic states, investigations of unfavorable energy spacing of the molecule or atom, simplifications inherent in local field models, and an analysis of the effects of truncation of the sum rules. These studies concluded that the energy spectrum of real systems compared with the ideal is the most likely factor that keeps the hyperpolarizabilities of real molecules well below the limit [31].

To gain insights about the gap, the effects of several different parameters, including molecular geometry, external electromagnetic fields, and electron-electron interactions, have been investigated. The effect of molecular geometry on the hyperpolarizability is determined by varying the positions and magnitudes of charges in two dimensions and correlating them with dipolar charge asymmetry and the variations of angle between point charges in octupolar structures. It was shown that the best dipolar and octupole-like molecules have intrinsic hyperpolarizabilities near 0.7 [27].

Hamiltonians of the standard form, given by

$$H = \frac{p^2}{2m} + V(r), \quad (7)$$

are used to calculate the sum rules. However, more general Hamiltonians also obey the sum rules. For example, the first hyperpolarizability of molecules was studied in the presence of an external electromagnetic field modeled by a vector potential term added to the momentum, and also led to best intrinsic hyperpolarizabilities of about 0.7 even when the external field

TABLE I. The largest intrinsic hyperpolarizabilities obtained using various theoretical approaches.

Theoretical models	β_{int}
Charge asymmetry	<0.7
External electromagnetic field	$\simeq 0.7$
Electron interactions	0.709
Potential optimization	0.709
Monte Carlo simulation	$\simeq 1$

is comparable with internal molecular fields [28]. Thus, all real one-electron systems modeled appear to have an upper bound of 0.709.

More recent work addresses the role of electron interactions on β_{int} [29] and shows that while β for two noninteracting electrons are twice that of a single-electron system, since the fundamental limit for a two-electron system scales as $N^{3/2}$ [1], increasing the number of electrons that do not interact with each other reduces β_{int} . At best, the theory of fundamental limits predicts that interacting electrons will lead to $N^{3/2}$ scaling of the hyperpolarizability. Calculations of two interacting electrons confirm this prediction, so the best values of β_{int} for this system are again found to be about 0.7 [29]. There is no reason to believe that adding additional electrons will lead to a larger *intrinsic* nonlinear response.

Other studies aimed at investigating the gap have used numerical optimization techniques to find the potential energy functions that maximize the hyperpolarizability. In this approach, many classes of starting potentials (polynomial, trigonometric, etc.) are varied until the hyperpolarizabilities are optimized using a finite-difference method to calculate β using

$$\beta = \frac{1}{2} \left. \frac{\partial^2 p}{\partial E^2} \right|_{E=0}, \quad (8)$$

where p is dipole moment calculated in the presence of the applied electric field. These too find the largest β_{int} values to be approximately 0.71 [24,25].

Atherton *et al.* used a different approach to optimize a piecewise linear potential function of the form $V(x) = Ax_n + B_n$ for $n \in \{1, 2, 3, \dots, N-1\}$, with N as the number of segments [32]. They also find the universal value of optimized hyperpolarizability to be $\beta_{\text{int}} \approx 0.71$ [24,25,33]. The authors use a Hessian matrix [34] of β_{int} to find the relevant parameters defining the hyperpolarizability and find that (a) two parameters suffice to find the appropriate potential function and (b) increasing the number of parameters does not improve upon the value of β_{int} . A summary of the largest attainable β_{int} values are presented in Table I.

In the above approaches, the Schrödinger equation is solved directly for many Hamiltonians. All quantum systems with an intrinsic hyperpolarizability near the fundamental limit, independent of the underlying Hamiltonian, are found to share universal properties: the three-level ansatz is obeyed; the ratio of the second to first excited state energies is $E = E_{10}/E_{20} \simeq 0.49$; and $X = x_{01}/x_{01}^{\text{max}} \simeq 0.79$, where $x_{01}^{\text{max}} = \hbar/(2mE_{10})^{1/2}$ is the largest possible transition moment to the first excited state. It is intriguing that such a broad range of systems share the same properties, and the natural question is why. In addition

to the fundamental questions that arise, universal properties can be used as a guide for optimizing the hyperpolarizability.

A larger configuration space can be probed using the sum rules directly in lieu of the Schrödinger equation. In this approach, Monte Carlo simulations are used to generate a distribution of β_{int} values based on millions of random samplings of the energies and transition moments under the constraint that the eigenenergies, E_i , and transition moments, x_{ij} , obey the sum rules. The largest hyperpolarizabilities are found once again to be characterized by three dominant states; but the hyperpolarizabilities can approach arbitrarily closely to the limit [35,36]. Since the sum rules cover a broader domain than the Schrödinger equation, it is not surprising that values greater than given by the Schrödinger equation are found. However, none are larger than $\beta_{\text{int}} = 1$.

The present work seeks to investigate the domain not included in the Schrödinger equation but constrained by the sum rules. An ancillary goal is to revisit the validity of the calculations of the fundamental limits and their applicability to systems described by standard Hamiltonians such as those given by Eq. (7). Since the sum rules encompass a broader range of systems than those derivable from typical Hamiltonians [36], we also investigate nonstandard Hamiltonians that may lead to artificial materials with larger nonlinear response. Finally, we discuss the many-state catastrophe, and why the three-level ansatz may be appropriate. This approach allows us to contrast our new findings with previous studies that lead to new interpretations of the TLA.

IV. THE HYPERPOLARIZABILITY OF N -LEVEL MODEL

In this section, we introduce the sum rules and use them as a constraint to optimize β for a general four-level model. Subsequently we generalize those results to an N -level model and show that a highly degenerate energy spectrum appears to break the fundamental limit. In the infinite-state model, the limits are found to diverge. This leads to a reformulation of the fundamental limits based on the Schrödinger equation, which suggests a reduced limit.

A. Sum rules

Sum rules, given by Eq. (2), are the foundation of the theory of fundamental limits of the first- and second-order hyperpolarizabilities, β [1] and γ [2], respectively. They have been used to calculate the fundamental limits of the off-diagonal components of β as measured with hyper-Rayleigh scattering [37], limits of two photon absorption cross sections [38], and used to formulate the dipole-free SOS expression for β [39] and γ [40].

B. Four-level model

In this section we will first apply sum rules to a four-level model of the hyperpolarizability to a very specific state, namely, to what we call a fully degenerate state, and find that the hyperpolarizability can exceed the fundamental limit so the intrinsic hyperpolarizability exceeds unity. This calculation is then generalized to an N -level fully degenerate state using the same method, which shows a divergence of the fundamental limit as $N \rightarrow \infty$, which we call the many-state catastrophe.

Using the dipole-free (DF) SOS expression [39], which is a simplification derived using a subset of the sum rules, the electronic first hyperpolarizability for a four-level model in the off-resonant regime, $\beta \equiv \beta_{4L}^{\text{DF}}$, is given by

$$\begin{aligned} \beta_{4L}^{\text{DF}} &= \sum_{m \neq n}^{3'} \beta_{mn}^{\text{DF}} \\ &= \beta_{12} + \beta_{21} + \beta_{13} + \beta_{31} + \beta_{23} + \beta_{32} \\ &= -3e^3 \left[x_{01}x_{12}x_{20} \left(\frac{2}{E_{10}E_{20}} - \frac{2E_{10} - E_{20}}{E_{20}^3} \right. \right. \\ &\quad \left. \left. - \frac{2E_{20} - E_{10}}{E_{10}^3} \right) + x_{01}x_{13}x_{30} \left(\frac{2}{E_{10}E_{30}} \right. \right. \\ &\quad \left. \left. - \frac{2E_{10} - E_{30}}{E_{30}^3} - \frac{2E_{30} - E_{10}}{E_{10}^3} \right) + x_{02}x_{23}x_{30} \right. \\ &\quad \left. \times \left(\frac{2}{E_{20}E_{30}} - \frac{2E_{20} - E_{30}}{E_{30}^3} - \frac{2E_{30} - E_{20}}{E_{20}^3} \right) \right], \quad (9) \end{aligned}$$

where the prime on the sum indicates that the ground state (labeled by zero) is excluded from the summation and

$$\beta_{mn}^{\text{DF}} = -3e^3 x_{0m}x_{mn}x_{n0} \left(\frac{1}{E_{n0}E_{m0}} - \frac{2E_{n0} - E_{m0}}{E_{m0}^3} \right). \quad (10)$$

Since the sum rules relate transition moments and energies to each other, they can be used to reduce the number of parameters in a truncated SOS expression. It has been shown that this procedure can be applied in a way that avoids pathologies [22].

We begin by using the sum rules to eliminate x_{03} , x_{13} , and x_{23} , as follows. The sum rule $(m, p) = (0, 0)$ [see Eq. (2)] for a four-level model yields

$$E_{10}|x_{01}|^2 + E_{20}|x_{02}|^2 + E_{30}|x_{03}|^2 = \frac{\hbar^2}{2m}; \quad (11)$$

$(m, p) = (1, 1)$ gives

$$E_{01}|x_{01}|^2 + E_{21}|x_{12}|^2 + E_{31}|x_{13}|^2 = \frac{\hbar^2}{2m}; \quad (12)$$

and $(m, p) = (2, 2)$ yields

$$E_{02}|x_{02}|^2 + E_{12}|x_{12}|^2 + E_{32}|x_{23}|^2 = \frac{\hbar^2}{2m}. \quad (13)$$

Solving Eqs. (11)–(13) for x_{03} , x_{13} , and x_{23} yields

$$x_{03} = \pm \sqrt{F} x_{01}^{\text{max}} \left(1 - X^2 - \frac{Y^2}{E} \right)^{1/2}, \quad (14)$$

$$x_{13} = \pm \sqrt{\frac{F}{1-F}} x_{01}^{\text{max}} \left(1 + X^2 - \frac{1-E}{E} Z^2 \right)^{1/2}, \quad (15)$$

$$x_{23} = \pm \sqrt{\frac{F}{E-F}} x_{01}^{\text{max}} (E + Y^2 + (1-E)Z^2)^{1/2}, \quad (16)$$

where

$$\begin{aligned} E &\equiv \frac{E_{10}}{E_{20}}, & F &\equiv \frac{E_{10}}{E_{30}}, & X &\equiv \frac{x_{10}}{x_{01}^{\text{max}}}, \\ Y &\equiv \frac{x_{20}}{x_{01}^{\text{max}}}, & \text{and } Z &\equiv \frac{x_{12}}{x_{01}^{\text{max}}}. \end{aligned} \quad (17)$$

x_{01}^{\max} is the largest possible transition moment from the ground state to any other state, and it is given by

$$x_{01}^{\max} = \left(\frac{\hbar^2 N}{2mE_{10}} \right)^{1/2}. \quad (18)$$

Rewriting Eq. (9) in terms of E , F , X , Y , and Z gives

$$\beta_{4L}^{\text{DF}} = \beta_{3L}^{\max} \mathcal{L}(X, Y, Z, E, F), \quad (19)$$

$$\begin{aligned} \mathcal{L}(X, Y, Z, E, F) = & \left(\frac{3}{4} \right)^{3/4} \left(XYZ[2E + (1 - 2E)E^2 - 2/E + 1] + F \left(1 - X^2 - \frac{Y^2}{E} \right)^{1/2} \left\{ \frac{X}{(1 - F)^{1/2}} \right. \right. \\ & \times \left(1 + X^2 - \frac{1 - E}{E} Z^2 \right)^{1/2} [2F + (1 - 2F)F^2 + -2/F + 1] \\ & \left. \left. + \frac{Y}{(E - F)^{1/2}} [E + Y^2 + (1 - E)Z^2]^{1/2} \left[(E + F)^2 - \left(\frac{2F^3}{E} + \frac{2E^3}{F} \right) \right] \right\} \right). \end{aligned} \quad (21)$$

$\mathcal{L}(X, Y, Z, E, F)$ in fact is the intrinsic hyperpolarizability for a four-level model, i.e.,

$$\beta_{\text{int}}^{4L} \equiv \frac{\beta}{\beta_{3L}^{\max}} = \mathcal{L}(X, Y, Z, E, F). \quad (22)$$

The maximum of Eq. (22) corresponds to the fundamental limit calculated from a four-level ansatz.

Based on the theory of the fundamental limits and the three-level ansatz, all intrinsic hyperpolarizabilities should be in the range $[-1, 1]$. By numerically finding the maximum of the function given by Eq. (22), we find that the largest value of β_{int} results from the extreme case where States 1 and 2 are degenerate and the energy of State 3 approaches infinity [Fig. 1(a)] or

$$E \simeq 1 \quad \text{and} \quad F \rightarrow 0 \quad (23)$$

yielding

$$|\beta_{\text{int}}^{4L}|^{\max} \simeq 1.28, \quad (24)$$

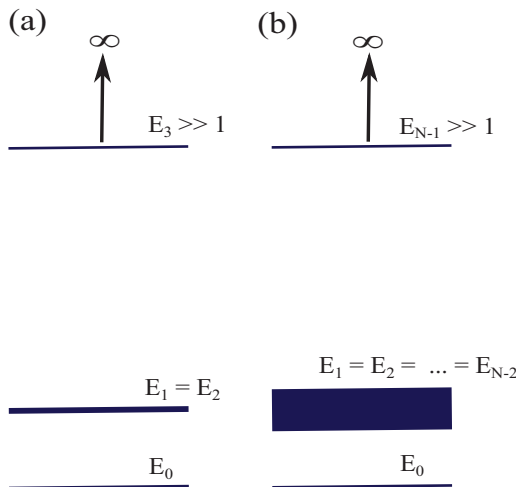


FIG. 1. (Color online) The ideal energy spectrum configuration that maximizes the hyperpolarizability of (a) four- and (b) N -level models.

where we have used the fundamental limit of the first hyperpolarizability obtained from the three-level model,

$$\beta_{3L}^{\max} = 3^{1/4} \left(\frac{e^3 \hbar^3}{m^3} \right) \left[\frac{N^{3/2}}{E_{10}^{7/2}} \right]. \quad (20)$$

$\mathcal{L}(X, Y, Z, E, F)$ is given by

which implies that

$$\beta_{4L}^{\max} = 1.28 \beta_{3L}^{\max}, \quad (25)$$

as shown in Fig. 2. Later we will generalize the result and discuss its pathologies.

V. N -LEVEL MODEL

In this section, we derive a general relationship for β_{int} for an arbitrary number of eigenstates that form a fully degenerate state. Starting from the dipole-free SOS expression [39]

$$\beta_{\text{int}} = \left(\frac{3}{4} \right)^{3/4} \sum_{i \neq j} \xi_{i0} \xi_{ij} \xi_{j0} \left(\frac{1}{e_i e_j} - \frac{2e_j - e_i}{e_i^3} \right), \quad (26)$$

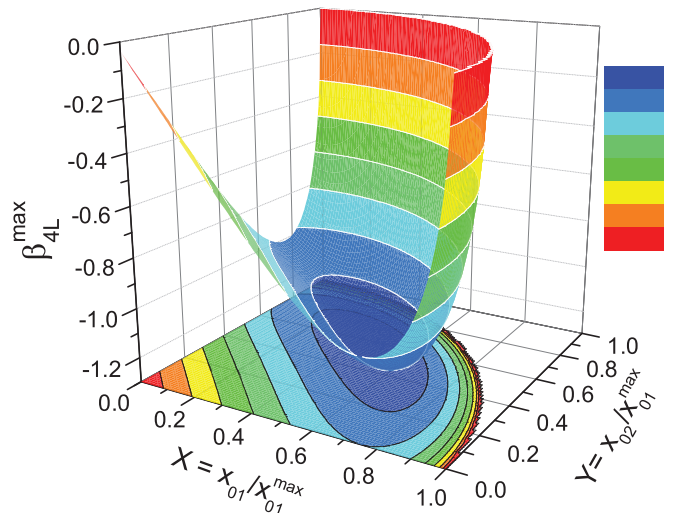


FIG. 2. (Color online) The numerical simulation for β_{int} of a four-level model. β_{int} exceeds the three-level maximum hyperpolarizability in the dark blue region.

where normalized transition moments and energies are defined by

$$\xi_{ij} = \frac{x_{ij}}{x_{01}^{\max}} \quad (27)$$

and

$$e_i = \frac{E_{i0}}{E_{10}}. \quad (28)$$

Monte Carlo calculations and empirical trial and error numerical simulations are used to find the maximum. Figure 1(b) shows the optimum, where states 1 to $(N - 2)$ are degenerate,

$$e_1 = e_2 = e_3 = \dots = e_{N-2}, \quad (29)$$

all the transition moments from the ground state to the $(N - 2)$ excited states are the same, given by

$$\xi_{01} \simeq \xi_{02} \simeq \xi_{03} \simeq \dots \simeq \xi_{0,N-2}, \quad (30)$$

and the energy of the highest-lying excited state approaches infinity,

$$e_{N-1} \gg 1. \quad (31)$$

The fully degenerate N -level state, defined by the above conditions as shown in Fig. 1(b), yields an intrinsic hyperpolarizability of

$$\begin{aligned} \beta_{\text{int}}^{\max} &= \beta_{1,N-1} + \beta_{2,N-2} + \dots + \beta_{N-2,N-1} + \beta_{N-1,N-2} \\ &= \left(\frac{3}{4}\right)^{3/4} \sum_i \xi_{0i} \xi_{i,N-1} \xi_{N-1,0} \\ &\quad \times \left(\frac{2}{e_i e_{N-1}} - \frac{2e_{N-1} - e_i}{e_i^3} - \frac{2e_i - e_{N-1}}{e_{N-1}^3} \right) \\ &= \left(\frac{3}{4}\right)^{3/4} (N-2) \xi_{10} \xi_{1,N-1} \xi_{N-1,0} \\ &\quad \times \left(\frac{2}{e_1 e_{N-1}} - \frac{2e_{N-1} - e_1}{e_1^3} - \frac{2e_1 - e_{N-1}}{e_{N-1}^3} \right) \\ &\simeq -2 \left(\frac{3}{4}\right)^{3/4} (N-2) e_{N-1} \xi_{01} \xi_{1,N-1} \xi_{N-1,0}, \end{aligned} \quad (32)$$

where we have neglected terms with $1/E_{N-1}$, which tends to zero when the highest excited state energy tends to infinity.

For further simplification, we use the sum rule $(m, p) = (0, 0)$,

$$e_{10} |\xi_{10}|^2 + e_{20} |\xi_{20}|^2 + \dots + e_{N-1,0} |\xi_{N-1,0}|^2 = 1, \quad (33)$$

whence

$$|\xi_{N-1,0}| = \sqrt{\frac{1 - (N-2) |\xi_{10}|^2}{e_{N-1,0}}}. \quad (34)$$

$(m, p) = (1, 1)$ gives

$$e_{01} |\xi_{10}|^2 + e_{21} |\xi_{21}|^2 + \dots + e_{N-1,1} |\xi_{N-1,1}|^2 = 1, \quad (35)$$

whence

$$|\xi_{N-1,1}| = \sqrt{\frac{1 + |\xi_{10}|^2}{e_{N-1,1}}} \simeq \sqrt{\frac{1 + |\xi_{10}|^2}{e_{N-1,0}}}. \quad (36)$$

Introducing Eqs. (34) and (36) into Eq. (32) we find

$$\begin{aligned} \beta_{\text{int}}^{\max} &= -2 \left(\frac{3}{4}\right)^{3/4} (N-2) \xi_{10} \\ &\quad \times \sqrt{1 - (N-3) |\xi_{01}|^2 - (N-2) |\xi_{01}|^4}. \end{aligned} \quad (37)$$

To optimize β with respect to ξ_{01} we first find ξ_{01}^{\max} by solving $\frac{\partial \beta}{\partial \xi_{01}} = 0$, leading to

$$\xi_{01}^{\max} = \left[\frac{\sqrt{(N-3)^2 + 3(N-2)} - (N-3)}{3(N-2)} \right]^{1/2}, \quad (38)$$

which holds for $N \geq 3$. Inserting Eq. (38) into (37) results in

$$\begin{aligned} \beta_{\text{int,SR}}^{\max} &= \left(\frac{1}{12}\right)^{1/4} \\ &\quad \times \left[(N-2) \frac{N^2 - 3 - (N-3) \sqrt{N(N-3) + 3}}{(N-3) + \sqrt{N(N-3) + 3}} \right]^{1/2}, \end{aligned} \quad (39)$$

where the subscript SR indicates the result is constrained by the sum rules. According to Eq. (39), for $N = 3$, $\beta_{\text{int}} = 1$ and for $N = 4$, $\beta_{\text{int}} = 1.28$ which is in agreement with the previous calculations.

The consequence of Eq. (39) is that β_{int} does not converge for an infinite number of states, but diverges as

$$\lim_{N \rightarrow \infty} \beta_{\text{int,SR}}^{\max} = \frac{3}{2} \left(\frac{1}{12}\right)^{1/4} \sqrt{N}. \quad (40)$$

Therefore, using only the sum rules without additional constraints leads to a many-state catastrophe for the hyperpolarizability of a quantum system. The same approach can be applied to the second hyperpolarizability, γ . The many-state catastrophe contradicts all studies based on direct solution of the Schrödinger equation, as well as experiment.

VI. DISCUSSION

It is not surprising that previous numerical simulations did not see this behavior given the extreme or unphysical conditions that are required. The many state catastrophe can be understood by considering the generalization from a three- to a four-level ansatz, which adds an additional degree of freedom, leading to two parameters, E and F . The number of degrees of freedom increases in proportion to the total number of states, making it possible to find a specific combination of parameters that yields an ever-larger nonlinear response. However, since only parameters in the neighborhood of one very specific set out of many possible configurations gives $\beta_{\text{int}} > 1$, even in millions of iterations in Monte Carlo simulations, the outliers are missed [35]. However, the very special set of parameters leading to Eq. (39) may not correspond to any *real* system, as we describe below. Given that analytical optimization of the hyperpolarizability using the Schrödinger equation always gives a result smaller than what is predicted by the three-level model, the three-level ansatz appears to shadow a fundamental principle that constrains the magnitude of the true nonlinear response.

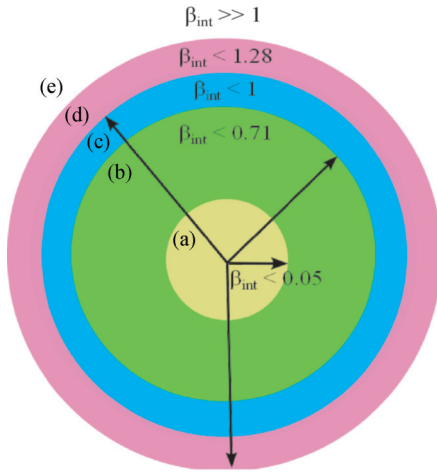


FIG. 3. (Color online) The range of attainable hyperpolarizabilities of (a) real systems such as atoms and molecules; (b) Hamiltonian-based approaches; (c) sum rule-based Monte Carlo simulations; (d) four-level model sum-rule-constrained optimization; and (e) sum rule-based $N > 4$ -level model optimization.

Figure 3 summarizes what we know about the space of all allowed transition moments and energies derivable from the sum rules. The observations are the following:

(1) All theoretical and numerical approaches that are based on exact solutions of the Schrödinger equation based on the standard Hamiltonian given by Eq. (7)—as well as generalizations that include spin, vector potentials, spin-orbit coupling, etc.—lead to $\beta_{\text{int}} \leq 0.71$.

(2) Monte Carlo simulations that are based on random sampling of transition moments and energies constrained to obey the sum rules generate $\beta_{\text{int}} < 1$.

(3) The fundamental limit calculations assert that when the maximum hyperpolarizability is attained, only three energy eigenstates contribute to the nonlinear response. This is observed for all computations using solutions of the Schrödinger equation when $\beta_{\text{int}} \approx 0.7$. However, in the case of highly degenerate states, we observe that $\beta_{N\text{-level}}^{\text{max}} > \beta_{N-1\text{-level}}^{\text{max}} > \dots > \beta_{3\text{-level}}^{\text{max}}$ and when the number of states, N , approaches infinity, the sum rules impose no limit on the hyperpolarizability unless a finite N -level ansatz is imposed.

Motivated by the above observations, we seek to answer the following questions:

(1) Why does the sum rules-based Monte Carlo approach lead to different results than direct solution of the Hamiltonian?

(2) Why have sum-rule-based Monte Carlo simulations not observed violations of the three-level ansatz, i.e., generate values with $\beta_{\text{int}} > 1$?

(3) If the intrinsic nonlinear-optical response of an optimized N -level system can in principle be infinite, is it possible that standard Hamiltonians may be fine-tuned to lead to ultralarge hyperpolarizabilities?

(4) Is there a well-defined limit for the largest attainable β of a quantum systems?

(5) How can we interpret the apparent success of the three-level ansatz and the large gap between the three-level-ansatz-based fundamental limit and most molecular systems?

The answer to these questions requires a detailed analysis of the theory of fundamental limits.

A. Potential issues with the Monte Carlo approach

The intrinsic first and second hyperpolarizabilities are functions of normalized transition moments, ξ_{ij} , and energies, e_i , given by Eqs. (27) and (28), respectively. In Monte Carlo simulations, these quantities are picked randomly under the constraint of the *diagonal* sum rules [35] and then using the dipole-free formulation [39,40], the hyperpolarizabilities are calculated. The wave functions are assumed real and nondegenerate, so the underlying Hamiltonians that describe these system, if they exist, must be invariant under time reversal [41].

There is potentially a pathology of this approach due to a subtle mixed-counting of states in the dipole-free expression, which uses the off-diagonal sum rules to remove the dipolar terms using the relationship

$$\sum_{m=1}^{\infty} \frac{\Delta x_{m0} |x_{m0}|^2}{E_{m0}} = - \sum_{m=1}^{\infty} \sum_{n \neq m} \frac{E_{nm} + E_{n0}}{E_{m0}^2} x_{0m} x_{mn} x_{n0}, \quad (41)$$

where $\Delta x_{m0} = x_{mm} - x_{00}$. Since Eq. (41) holds only when all infinite number of states in the n summation are included, truncating the sums in the DF SOS expression implicitly truncates Eq. (41), which leads to inaccuracies in the dipolar terms. Indeed, it is found that the standard SOS and DF SOS expressions can disagree with each other even when the hyperpolarizabilities are calculated from a Hamiltonian [24,25,33].

The issue can be described as follows. In the standard SOS expression, the hyperpolarizability comes from two groupings of terms: ones that are functions of the dipole moment difference $\delta x_{n0} \equiv x_{nm} - x_{00}$, and other terms that are not functions of δx_{n0} . When the standard SOS expression is evaluated using an N -state subspace, all position matrix elements x_{nm} are ignored if $n > N$ OR $m > N$. When the sum rules are used to express the dipolar terms to sums of nondipolar terms to get the DF form of the SOS expression according to Eq. (41), an N -state subspace truncates the sum, making the dipolar term potentially inaccurate. However, it is not clear if the inaccuracy is so large as to lead to hyperpolarizabilities that exceed unity.

Another potential issue originates in the Monte Carlo approach, which may allow for unphysical behavior due to the procedure of demanding agreement with truncated sum rules. For example, there may be combinations of matrix elements and energies that obey the sum rules, yet are not derivable from a Hamiltonian. As a case in point, it is possible to contrive a system with a finite number of states that exactly obeys the sum rules [35,36], but the nature of a system with a finite number of states renders it incompatible with being a solution of a Hamiltonian that depends on continuous functions.

However, under most circumstances, such finite state models that obey the sum rules exactly may accurately approximate real systems with an infinite number of states if the higher-energy eigenstates do not contribute substantially to the nonlinear optical response. In such cases, the finite-state Monte Carlo-constrained parameters may approximate

the conditions that allow finite-state models to be a good approximation to systems with an infinite number of states [22]. The fact that millions of Monte Carlo runs miss the configuration that leads to the multistate catastrophe suggests that perhaps the pathologies are rare and can be ignored.

It must be stressed that in finite state models, all the sum rules are not obeyed, since the (N, N) sum rule is self-contradictory and thus ignored [42]. Recall that in the Monte Carlo simulation, only the diagonal sum rules up to $(N - 1, N - 1)$ are used to avoid the problem. However, in the special case when all states are nearly degenerate and the highest-energy eigenstate tends to infinite energy, the system may be contrived in just the right way to violate the higher-level sum rules. Thus, the combination of taking an unphysical configuration of states, and then taking the limiting case of an infinite number of states leads to the divergence.

It may seem paradoxical that in the limit of an infinite number of states, where the sum rules become exact, the many-state catastrophe is observed. However, at issue is the method of how the limit is defined.

One can ask why the calculation of the fundamental limit, based on the three-level model, appears to work so well when comparing it with all hyperpolarizabilities derived from the Schrödinger equation. The origin of its success may lie in the fact that the three-level ansatz in effect sidesteps the infinite-state catastrophe by allowing only the minimum number of states. The only resulting problem, based on many solutions of the Schrödinger equation using many different approaches [25–29, 32], is that the calculation overestimates the observations by 30%. It is interesting that a condition that would yield the observed limit is a 2.2 level ansatz, *viz.*, Eq. (39). Since the two-level model is unphysical, the three-level model is a compromise that yields the best result.

B. Fundamental limits

Equation (39) represents the maximum attainable intrinsic hyperpolarizability that diverges as the number of states approaches infinity. This would suggest that for real systems with an infinite number of states, the sum rules impose no limit on the maximum hyperpolarizability. However, no real quantum system has been observed with this characteristic; all the experimental data and numerical calculations of β_{int} and γ_{int} [2] values fall well below the limit predicted by the three-level model; and the three-level ansatz is obeyed in the neighborhood of the limit for β while a four-level model applies in the neighborhood of the limit for γ .

The reason for this conflict may reside in the fact that sum rules permit energies and transition moments that are unphysical. The sum rules hold for any system of particles of mass m , provided that the Hamiltonian obeys

$$\langle p|[x, [H, x]]|q\rangle = \frac{\hbar^2}{m} \delta_{pq}, \quad (42)$$

where $|p\rangle$ and $|q\rangle$ are eigenstates of the Hamiltonian H . Hamiltonians that obey Eq. (42) include those of the form

$$H = f(p, x) + g(x) + h(p) + k(A, B, \dots), \quad (43)$$

where p and x are the momentum and position operator and A and B are any other operators, such as the angular momentum

or spin. It is straightforward to show that a particle in the presence of electromagnetic field obeys this form, that is,

$$H = \frac{[p - eA(x)/c]^2}{2m} + e\phi(x), \quad (44)$$

where $A(x)$ and $\phi(x)$ are the vector and scalar potentials, respectively, with $f(p, x) = (p - eA(x)/c)^2$, $g(x) = e\phi(x)$ and $h(p) = k(A, B) = 0$.

There are clearly Hamiltonians that obey Eq. (43) that are more general than Eq. (44). Perhaps one of these exotic Hamiltonians may break the $\beta_{\text{int}} < 0.71$ barrier. Or perhaps values larger than unity are possible. Further study is needed to assess these possibilities.

The approaches that are based on direct treatment of standard Hamiltonian are more realistic than sum-rule-based approaches that are not derivable from a Hamiltonian. The observation of the many-state catastrophe is most likely not an indictment on the sum rule calculations, but on the pathology of the highly degenerate spectrum. In reality, we propose that the maximum attainable hyperpolarizability is most likely given by

$$\beta_{\text{max}} = 0.71 \beta_{3L}^{\text{max}}. \quad (45)$$

VII. A CONJECTURE

The sum total of all measurements and calculations using Hamiltonians leads to two observations: (1) the largest intrinsic hyperpolarizability is less than 0.709, and (2) when a quantum system is found to have a hyperpolarizability in the vicinity of this maximum, only three states contribute. These two observations have never been proven, so it behooves us to ask whether or not these are indeed fundamental laws of physics, or a mere coincidence. The many-state catastrophe is an important part of the puzzle because it violates both these observations. This is not surprising nor a matter of concern given that wave functions that are solutions to standard Hamiltonians give a range of dipole matrix elements and energies that are a subset of the dipole matrix elements and energies that obey the sum rules.

When computing limits and other fundamental relationships involving the quantum origin of optical nonlinearity, it is important to not lose sight of the fact that the hyperpolarizabilities are coefficients in an expansion of the dipole operator in terms of the applied electric field. As such, these quantities have no meaning when a power series expansion is not possible. The key point is that the hyperpolarizability is a quantity that is derived from perturbation theory. Thus, any properties of a given system must be derived directly from the Hamiltonian of the quantum system including the contribution from the perturbation of the photon field. Auxiliary quantities, such as the sum rules, must clearly be obeyed; but, as shown in this paper, may lead to behavior that is nonphysical if the properties of the Hamiltonian are not also used to constrain the system.

In the calculation of the fundamental limits of the hyperpolarizability using the sum rules, an additional auxiliary condition is added, namely, the three-level ansatz. The argument for its use is as follows. One can show rigorously

and without approximation that a two-state model optimizes the polarizability, α , in which the oscillator strength is placed into one excited state, avoiding dilution effects that arise from spreading oscillator strength between many higher-energy states. However, as we saw in Sec. II, the sum rules show that a two-level quantum system must have $\beta = 0$ (β is *minimum* when α is maximum). The three-state model has the minimum number of states that leads to a physically reasonable result, and is therefore used based on the dilution argument.

While no rigorous proof has been found that a three-state model yields a maximum, the approach is justified by the heuristic argument that a concentration of oscillator strength in a small number of low-energy states yields large numerators and small denominators in the SOS expression, thus maximizing the hyperpolarizability. With more states, the oscillator strength is distributed over many states, thus diluting the nonlinearity. However, the many-state catastrophe arises from a very specific highly degenerate quantum system that spreads oscillator strength over many states of equal energy, thus preserving oscillator strength but without dilution because the degeneracy keeps all energy denominators small.

Before arguing that the highly degenerate spectrum is nonphysical, one other loose end needs to be addressed: the *a priori* assumption of the three-level ansatz and its potential role in being responsible for the observation that all quantum systems ever calculated obey it. There is clearly no causal connection between calculating the limits using TLA and calculating β for a specific Hamiltonian. The three-level ansatz is used only in the calculations of the fundamental limits, and not in the analysis of the hyperpolarizabilities that are calculated from Hamiltonians. Thus, the observation that the three-level ansatz holds in all quantum systems tested is not tied to its assumption in the calculation of the limits.

The calculation of the fundamental limits of the polarizability is made simple by the fact that each term in the SOS expression is positive definite, and of the form

$$e^2 \frac{x_{0i}^2}{E_{i0}}. \quad (46)$$

As we saw in Eqs. (3) and (4), placing all the transition strength into one term, and picking the term with the minimum energy maximizes α . All linear harmonic oscillators are at the fundamental limit with $\alpha_{\text{int}} = 1$. While a harmonic oscillator has many states, all the oscillator strength resides in the transition to the first excited state. Thus, the two-state model holds exactly. β also is an exact two-level model for a linear harmonic oscillator, and as predicted by the sum rules, $\beta = 0$. γ also vanishes, as do all orders of nonlinearity by virtue of the fact that the linear harmonic oscillator is the prototypical purely linear system.

Using only the sum rules, the limits of β are made difficult to calculate. A typical term in the sum is of the form

$$e^3 \frac{x_{0i} \bar{x}_{ij} x_{j0}}{E_{i0} E_{j0}}, \quad (47)$$

where $\bar{x} = x - x_{00}$. Since each term of the form given by Eq. (47) is of indeterminate sign, the limit cannot be determined. The many-state catastrophe shows that using only the sum rules and not truncating the SOS expression leads to a diverging result, i.e., that there is no limit. This runs counter to

the evidence. Clearly, the correct approach is to apply the sum rules, and rather than truncating the SOS expression, to use an additional constraint that is determined from the general form of the Hamiltonian. All attempts to find such an auxiliary condition have failed.

The state of affairs can be summarized as follows. When applying the three-level ansatz to the SOS expression, it is found to be optimized for a specific value of energy ratio $E = E_{10}/E_{20}$ and $X = x_{10}/x_{10}^{\text{MAX}} \approx 0.76$. The value of β obtained in this way is found to be an upper bound for all calculations that evaluate a specific Hamiltonian when the errors resulting from the numerical approximations that are used to evaluate the wave functions, dipole matrix, energy eigenvalues, and hyperpolarizability are small (less than $\approx 1\%$). In other words, all accurate calculations support these conclusions. The best systems reach $\beta_{\text{int}} = 0.7089$. In the vicinity of this maximum, three states dominate the second hyperpolarizability, β_{xxx} . These suggest that the limits calculated and the three-level ansatz may be generally true.

The fact that the many-state catastrophe leads to hyperpolarizabilities that are much higher than experimentally observed and calculated values suggests that the combination of the sum rules and the three-level ansatz together yield a result that is near the true fundamental limit for hyperpolarizabilities that come from a standard Hamiltonian.

To analyze the many-state catastrophe in the vicinity of the highly degenerate energy spectrum, the degeneracy can be lifted in a smooth way by redefining the spectrum according to

$$e_n = e_1 + \epsilon(n - 1) = 1 + \epsilon(n - 1), \quad (48)$$

which holds for the nearly degenerate excited states, or for $n \neq 0$, the ground state, and $n \neq N - 1$, the highest excited state. ϵ is the splitting parameter that separates the $N - 2$ degenerate states into evenly spaced energy levels. Equation (48) can be inverted and solved for $E (=1/e_2)$, yielding

$$E = \frac{1}{1 + \epsilon}. \quad (49)$$

The nearly degenerate system is studied as follows. For each ϵ , 10 000 transition moments are randomly sampled, and the largest value plotted in Fig. 4 for a four- through 10-state model. The vertical lines show where $\beta_{\text{int}} = 1$, so to the right of these lines, the hyperpolarizability is below the limit. The subscript in the E_{NL} label represents the number of states used in the calculation of E . Note that this plot is approximate due to fluctuations associated with using a finite number of random samplings.

For the four-state model, $\beta_{\text{int}} < 1$ for $\epsilon < 1.11$, which corresponds to $E = 0.47$. With more states, β_{int} falls below unity for smaller values of ϵ . This is interesting in light of the fact that the best hyperpolarizabilities all share the universal value of $E \approx 0.49$, a value between that of a harmonic oscillator, which has a polarizability at the limit, and the crossing point of $E = 0.47$ for the four-level model.

The three-level ansatz can be cast in the form

$$\beta_{\text{int}} = f(E)G(X), \quad (50)$$

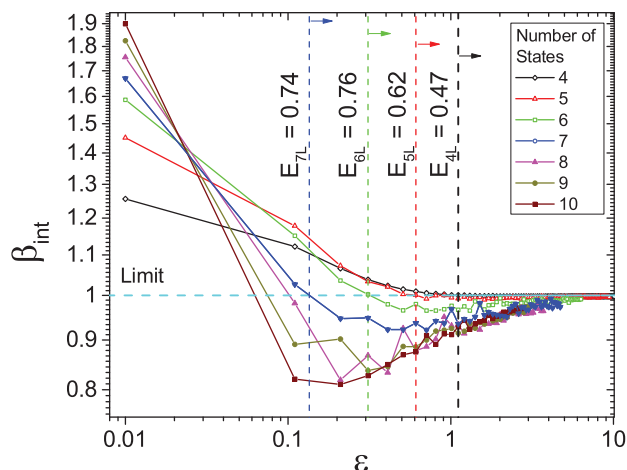


FIG. 4. (Color online) β_{int} as a function of energy spacing parameter ϵ .

where $E = E_{10}/E_{20}$ and $X = x_{10}/x_{10}^{\text{MAX}}$. Note that this expression is postulated to hold only near the fundamental limit and is observed to hold in the vicinity of a local maximum of β_{int} when close to the limit. The fundamental limit is found for $f(0) = 1$ and $G(\sqrt[3]{3} \approx 0.76)$. In potential optimization studies, the class of optimized potentials all share the universal properties of $\beta_{\text{int}} = 0.7089$, $E = 0.49$, and $X = 0.79$ [25]; so X is near the value needed to be at the limits. As such, the hyperpolarizability appears to be limited by the nature of the energy spectrum, as suggested in studies of various organic molecules [31,43].

In the calculations of the fundamental limit, X and E are assumed to be independent and are thus separately optimized. Since the universal value of X is near optimum, we focus on E . For a harmonic oscillator, $E = 1/2$ and $f(1/2) = 1/\sqrt{2} = 0.707$, a value tantalizingly close to the universal value of β_{int} . This suggests that perhaps X and E are not independent, so the best values of the intrinsic hyperpolarizability are constrained to have an energy spectrum that is similar to the harmonic oscillator when the transition is nearly optimized.

The sum rules (which are more general than what one obtains from solving the standard Schrödinger equation), the three-level ansatz (which is not generally true based on the exception found using the many-state catastrophe), and the assumption that E and X are independent (unproven) yields a calculated fundamental limit that is within 30% of the maximum value observed in many classes of optimized potentials. Furthermore, universal values are found that add credence to the ideas of absolute limits of scaled hyperpolarizabilities; and, the best measured molecules, as shown in Fig. 5, scale parallel to the limit lines, another indication that there is some substance behind the results.

We propose that the observations that the state vectors associated with the many-state catastrophe live in the realm obeyed by the sum rules but beyond solutions of the Schrödinger equation. Thus, in calculating limits of real systems, we must exclude those cases. Figure 4 illustrates the demarcation between the allowed (below $\beta_{\text{int}} = 1$) and disallowed. A four-level model demands that the degeneracy parameter be $\epsilon \geq 1.11$ yielding $E \leq 0.47$; a five-level model

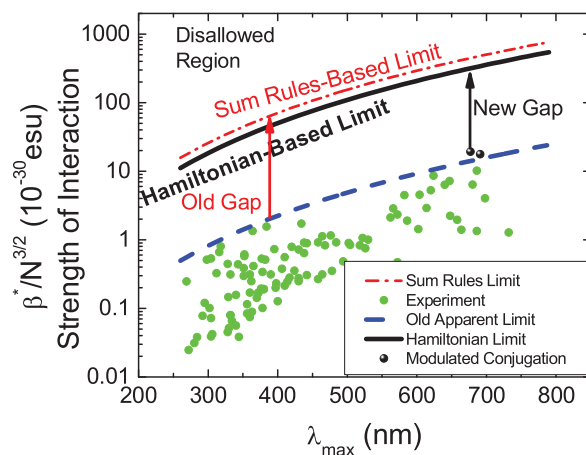


FIG. 5. (Color online) A survey of measured and calculated values of β . The red dashed-dot line is the sum rules-based limit assuming the TLA. The solid dark line is the slightly lower limit suggest by calculations using standard Hamiltonians. The green dots are experimental values reported prior to 2007. The old apparent limit, dashed blue line, was defined as the largest β values achieved at the time. The gulf between apparent limit and the sum rules-based FL is known as the gap.

demands that $\epsilon \geq 0.61$ yielding $E \leq 0.62$, etc., for a system of equally spaced intermediate states. Clearly, there are many other possible types of energy spectra, but this example illustrates how one can define a condition that sets limits on the allowed spectra for quantum systems that obey the standard Schrödinger equation.

Given these results, we propose the following conjectures:

(1) At the fundamental limit, only three states contribute to the hyperpolarizability and four states contribute to the second hyperpolarizability.

(2) $\beta_{\text{int}} \leq 0.7089$ and $\gamma_{\text{int}} \leq 0.6$ for any system derivable from a standard Hamiltonian.

Though we have not discussed γ in this paper, we have added the observed behavior for the second hyperpolarizability [30].

The SOS expression for β has terms of the form given by Eq. (47), which are indeterminate in sign, making the calculations of the limit impossible without the use of an auxiliary condition. Without placing a constraint on the type of potentials that are allowable, the limit is not calculable. As such, it is possible that these conjectures are true but unprovable.

As a corollary to our conjectures, we propose that they are true but unprovable. The fact that the conjecture has not been proven is not evidence that it is unprovable. However, without an additional condition, the conjectures can not be proven. As such, we must wait until such a condition is found.

The diagram in Fig. 6 summarizes the assumptions of limit theory, its consequences, and observations. When the sum rules alone are used to reduce the number of parameters, the intrinsic hyperpolarizability can be infinite in the neighborhood of the highly degenerate energy spectrum. However, when the three-level ansatz is applied, a finite limit results, and all quantum systems are observed to obey the limit, though the largest values observed from standard Hamiltonians is 0.7089. When

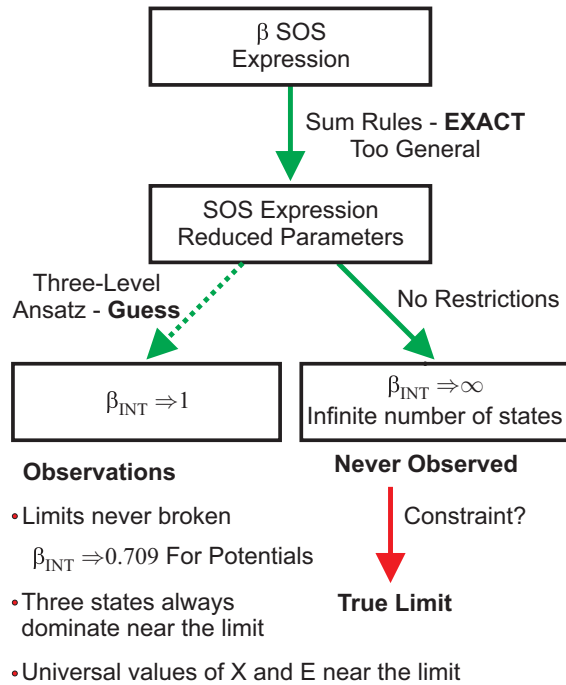


FIG. 6. (Color online) Flow chart summarizing the results.

the hyperpolarizability is at the limit, three states are always observed to dominate.

It is not possible to prove that a three-level system yields the limit because the many state catastrophe shows that an infinite number of states can contribute and there is no limit. However, enforcing the three-level ansatz leads to a finite limit for the hyperpolarizability that is consistent with all observations. Thus, the three-level ansatz cannot be proven because the many-state catastrophe is a counterexample. If the three-level ansatz is assumed to be true, it leads to the correct result. As such, the three-level ansatz is not deduced from other principles, but results from induction. In this sense, it is unprovable but true based on all known observations.

A proof of the three-level ansatz would require an additional condition that relates the energies and position matrix elements that has the effect of excluding all systems that are not derivable from a mechanical Hamiltonian yet selects every possible Hamiltonian that describes a *real* quantum system. Such a condition may be complex, and perhaps impossible to identify.

The three-level ansatz is to a good level of approximation a simple condition that seems to select for systems derivable from real Hamiltonians that have hyperpolarizabilities near the fundamental limit.

VIII. CONCLUSION

The many-state catastrophe is an observation that calls into question the assumptions used in calculating the fundamental limit. When many states are arranged into a highly degenerate energy spectrum, the hyperpolarizability diverges in the limit of an infinite number of states. While the divergence can be avoided by arguing that real systems that obey a standard Hamiltonian cannot have such a spectrum, the fact that this outlier both obeys the sum rules and contradicts the assumptions in the calculation of the limits demands that the assumptions be revisited.

By smoothly splitting the degeneracy using an energy spacing parameter, the fundamental limit theory is found to hold when $E > 0.47$, where the many-state catastrophe is avoided. The fact that the observed universal value for Hamiltonians expressed in terms of a potential energy function near the largest observed upper bound of the hyperpolarizability ($E \approx 0.49$) meets this condition supports the assertion that systems representable by a potential energy function have a restricted energy spectrum of this form.

Based on these observations, we propose a conjecture that the three-level ansatz is a fundamental law and that the true fundamental limit is given by $\beta_{\text{int}} = 0.7089$. We also posit that while true, the conjectures may not be provable due to the difficulty (or impossibility) of defining an auxiliary condition that standard Hamiltonians must obey while excluding more general systems.

The three-level ansatz appears to be the correct auxiliary condition needed to calculate the true fundamental limit to within 30% of the correct value. Reconciling the three-level ansatz with the many-state catastrophe and the overestimation of the true limit are open research problems under study.

ACKNOWLEDGMENTS

We would like to thank the National Science Foundation (ECCS-1128076) for generously supporting this work and Sean Mossman for his useful suggestions.

- [1] M. G. Kuzyk, *Phys. Rev. Lett.* **85**, 1218 (2000).
- [2] M. G. Kuzyk, *Opt. Lett.* **25**, 1183 (2000).
- [3] M. G. Kuzyk, *Phys. Rev. Lett.* **90**, 039902(E) (2003).
- [4] M. G. Kuzyk, *Opt. Lett.* **28**, 135 (2003).
- [5] M. G. Kuzyk, *Nonlin. Opt. Quant. Opt.* **40**, 1 (2010).
- [6] M. G. Kuzyk, J. Perez-Moreno, and S. Shafei, *Phys. Rep.* **529**, 297 (2013).
- [7] A. D. Slepukov, F. A. Hegmann, S. Eisler, E. Elliot, and R. R. Tykewinski, *J. Chem. Phys.* **120**, 6807 (2004).
- [8] J. C. May, J. H. Lim, I. Biaggio, N. N. P. Moonen, T. Michinobu, and F. Diederich, *Opt. Lett.* **30**, 3057 (2005).
- [9] J. C. May, I. Biaggio, F. Bures, and F. Diederich, *App. Phys. Lett.* **90**, 251106 (2007).
- [10] Q. Y. Chen, L. Kuang, Z. Y. Wang, and E. H. Sargent, *Nano. Lett.* **4**, 1673 (2004).
- [11] H. Kang, A. Facchetti, P. Zhu, H. Jiang, Y. Yang, E. Cariati, S. Righetto, R. Ugo, C. Zuccaccia, A. Macchioni, C. L. Stern, Z. Liu, S. T. Ho, and T. J. Marks, *Angew. Chem. Int. Ed.* **44**, 7922 (2005).
- [12] H. Kang, A. Facchetti, H. Jiang, E. Cariati, S. Righetto, R. Ugo, C. Zuccaccia, A. Macchioni, C. L. Stern, Z. F. Liu, S. T. Ho, E. C. Brown, M. A. Ratner, and T. J. Marks, *J. Am. Chem. Soc.* **129**, 3267 (2007).
- [13] G. S. He, J. Zhu, A. Baev, M. Samoć, D. L. Frattarelli, N. Watanabe, A. Facchetti, H. Ågren, T. J. Marks, and P. N. Prasad, *J. Am. Chem. Soc.* **133**, 6675 (2011).

- [14] J. Zhou and M. G. Kuzyk, *J. Phys. Chem. C* **112**, 7978 (2008).
- [15] J. Pérez-Moreno, Y. Zhao, K. Clays, and M. G. Kuzyk, *Opt. Lett.* **32**, 59 (2007).
- [16] J. Pérez-Moreno, Y. Zhao, K. Clays, M. G. Kuzyk, Y. Shen, L. Qiu, J. Hao, and K. Guo, *J. Am. Chem. Soc.* **131**, 5084 (2009).
- [17] W. Thomas, *Naturwissenschaften* **13**, 627 (1925).
- [18] F. Reiche and W. Thomas, *Z. Phys.* **34**, 879 (1925).
- [19] W. Kuhn, *Z. Phys. A: Hadrons Nuclei* **33**, 408 (1925).
- [20] B. J. Orr and J. F. Ward, *Mol. Phys.* **20**, 513 (1971).
- [21] M. G. Kuzyk and C. W. Dirk, *Characterization Techniques and Tabulations for Organic Nonlinear Optical Materials* (Marcel Dekker, New York, 1998).
- [22] M. G. Kuzyk, *J. Chem Phys.* **125**, 154108 (2006).
- [23] B. Champagne and B. Kirtman, *Phys. Rev. Lett.* **95**, 109401 (2005).
- [24] J. Zhou, M. G. Kuzyk, and D. S. Watkins, *Opt. Lett.* **31**, 2891 (2006).
- [25] J. Zhou, U. B. Szafruga, D. S. Watkins, and M. G. Kuzyk, *Phys. Rev. A* **76**, 053831 (2007).
- [26] U. B. Szafruga, M. G. Kuzyk, and D. S. Watkins, *J. Nonl. Opt. Phys. Mat.* **19**, 379 (2010).
- [27] M. G. Kuzyk and D. S. Watkins, *J. Chem Phys.* **124**, 244104 (2006).
- [28] D. S. Watkins and M. G. Kuzyk, *J. Chem. Phys.* **131**, 064110 (2009).
- [29] D. S. Watkins and M. G. Kuzyk, *J. Chem. Phys.* **134**, 094109 (2011).
- [30] D. Watkins and M. Kuzyk, *J. Opt. Soc. Am. B* **29**, 1661 (2012).
- [31] K. Tripathy, J. Perez-Moreno, M. G. Kuzyk, B. J. Coe, K. Clays, and A. M. Kelley, *J. Chem. Phys.* **121**, 7932 (2004).
- [32] T. Atherton, J. Lesnefsky, G. Wiggers, and R. Petschek, *J. Opt. Soc. Am. B* **29**, 513 (2012).
- [33] J. Zhou, M. G. Kuzyk, and D. S. Watkins, *Opt. Lett.* **32**, 944 (2007).
- [34] W. Press, B. Flannery, S. Teukolsky, W. Vetterling, *et al. Numerical Recipes*, Vol. 3 (Cambridge University Press, Cambridge, 2007).
- [35] M. C. Kuzyk and M. G. Kuzyk, *J. Opt. Soc. Am. B.* **25**, 103 (2008).
- [36] S. Shafei, M. C. Kuzyk, and M. G. Kuzyk, *J. Opt. Soc. Am. B* **27**, 1849 (2010).
- [37] M. G. Kuzyk, *IEEE J. Sel. Topics Quant. Elec.* **7**, 774 (2001).
- [38] M. G. Kuzyk, *J. Chem Phys.* **119**, 8327 (2003).
- [39] M. G. Kuzyk, *Phys. Rev. A* **72**, 053819 (2005).
- [40] J. Pérez-Moreno, K. Clays, and M. G. Kuzyk, *J. Chem. Phys.* **128**, 084109 (2008).
- [41] J. J. Sakurai, in *Modern Quantum Mechanics*, rev. ed., edited by S. F. Tuan (Addison Wesley Longman, New York, 1994).
- [42] M. G. Kuzyk, *J. Nonlin. Opt. Phys. Mat.* **15**, 77 (2006).
- [43] K. Tripathy, J. Perez-Moreno, M. G. Kuzyk, B. J. Coe, K. Clays, and A. M. Kelley, *J. Chem. Phys.* **125**, 9905 (2006).