Theoretical and numerical investigation of the size-dependent optical effects in metal nanoparticles

Alexander A. Govyadinov[*]

CIC nanoGUNE Consolider, Avenida de Tolosa 76, E-20018 Guipuzcoa, Spain

George Y. Panasyuk[*]

Propulsion Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433, USA

John C. Schotland

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109, USA

Vadim A. Markel

Departments of Radiology and Bioengineering and the Graduate Group in Applied Mathematics and Computational Science,
University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

We further develop the theory of quantum finite-size effects in metallic nanoparticles, which was originally formulated by F. Hache, D. Ricard, and C. Flytzanis [J. Opt. Soc. Am. B **3**, 1647 (1986)] and (in a somewhat corrected form) by S. G. Rautian [Sov. Phys. JETP **85**, 451 (1997)]. These references consider a metal nanoparticle as a degenerate Fermi gas of conduction electrons in an infinitely high spherical potential well. This model (referred to as the HRFR model below) yields mathematical expressions for the linear and the third-order nonlinear polarizabilities of a nanoparticle in terms of infinite nested series. These series have not been evaluated numerically so far and, in the case of nonlinear polarizability, they cannot be evaluated with the use of conventional computers due to the high computational complexity involved. Rautian has derived a set of remarkable analytical approximations to the series but direct numerical verification of Rautian's approximate formulas remained a formidable challenge. In this work, we derive an expression for the third-order nonlinear polarizability, which is exact within the HRFR model but amenable to numerical implementation. We then evaluate the expressions obtained by us numerically for both linear and nonlinear polarizabilities. We investigate the limits of applicability of Rautian's approximations and find that they are surprisingly accurate in a wide range of physical parameters. We also discuss the limits of small frequencies (comparable to or below the Drude relaxation constant) and of large particle sizes (the bulk limit) and show that these limits are problematic for the HRFR model, irrespective of any additional approximations used. Finally, we compare the HRFR model to the purely classical theory of nonlinear polarization of metal nanoparticles developed by us earlier [G. Y. Panasyuk, J. C. Schotland, and V. A. Markel, Phys. Rev. Lett. **100**, 47402 (2008)].

I. INTRODUCTION

Metal nanoparticles have received an extraordinary amount of attention recently because of their ability to greatly enhance local fields. The enhancement is attributed to the excitation of surface plasmons and it has a variety of applications in photovoltaics,[1] sensing,[2] and surface-enhanced Raman scattering.[3–5] Currently, nanoparticles of very small sizes, up to a few nanometers, are customarily used in experiments. The theoretical description of the optical properties of such nanoparticles is most frequently based on the macroscopic electrodynamics. At least, this is typical in the field of plasmonics. However, macroscopic electrodynamics cannot capture certain effects of finite size. Hache, Ricard, and Flytzanis[6] and Rautian (in a somewhat modified form)[7] have developed an elaborate theory of quantum finite-size effects in metal nanoparticles. In Refs. 6,7, a nanoparticle was modeled as degenerate Fermi gas confined in an infinite potential well of spherical shape (below, the HRFR model). Despite being fairly simple, the HRFR model results in very complicated formulas, which cannot be evaluated numerically even with the aid of modern computers. For example, the expression for the third-order nonlinear polarizability involves a twelve-

fold nested summation. Rautian has reduced the number of summations from twelve to eight by performing summation over the magnetic sublevels analytically; he then obtained a number of remarkable approximations to the resulting eightfold summation.[7] However, these approximations have never been verified directly due to the overwhelming numerical complexity involved.

In this paper, we develop the analytical theory of Rautian a step further by reducing the number of nested summations involved from eight to five without making any additional approximations. This turns out to be sufficient to render the formulas amenable to direct numerical implementation. We then compare the results of numerical evaluation of the fivefold series derived by us to the results which follow from Rautian's approximate formulas, and discuss various physical limits, including the limits of low frequency and large particle size.

In Sec. II we review Rautian's theory. Here we use somewhat simplified notation and, in particular, avoid the use of irreducible spherical tensors and $6j$ symbols. In Sec. III, we develop the theory further by utilizing the orbital selection rules and reduce the nested summation involved in the definition of the third-order nonlinear polarizability from

eightfold to fivefold. In Sec. IV, we describe a simple method to relate the internal and applied fields, which is to the first order consistent with the approach proposed in Ref. 8, but is more mathematically rigorous. In Sec. V, the results of numerical computations are reported. A summary of obtained results and a discussion are contained in Sec. VI.

II. RAUTIAN'S THEORY

We start by reviewing Rautian's theory of quantum finite-size effects in conducting nanoparticles.[7] The physical system under consideration is a gas of $\mathcal{N}$ noninteracting electrons placed inside a spherical, infinitely deep potential well of radius a and subjected to a harmonically oscillating, spatially uniform electric field,

$$\mathbf{E}_i(t) = \mathbf{A}_i \exp(-i\omega t) + \text{c.c.} \quad (1)$$

Note that $\mathbf{E}_i$ is the electric field *inside* the nanoparticle. It will be related to the *external* (applied) field $\mathbf{E}_e$ in Sec. IV below.

Since the nanoparticle is assumed to be electrically small (that is, $a \ll \lambda = 2\pi c/\omega$), the electron-field interaction can be described in the dipole approximation by the time-dependent operator

$$V(t) = -\mathbf{d} \cdot \mathbf{E}(t), \quad (2)$$

where $\mathbf{d}$ is the dipole moment operator. Under the additional assumption that $\mathbf{E}_i(t)$ is linearly polarized with a purely real amplitude $\mathbf{A}_i = A_i\hat{\mathbf{z}}$, we can write

$$V(t) = G \exp(-i\omega t) + \text{c.c.}, \quad (3)$$

where

$$G = -\mathbf{d} \cdot \mathbf{A}_i. \quad (4)$$

Rautian made use of the interaction representation in which the wave function is expanded in the basis of the unperturbed Hamiltonian eigenstates. The single-electron unperturbed states are

$$\phi_{nlm}(\mathbf{r}) = \frac{1}{Z_{nl}} j_l(\xi_{nl}r/a) Y_{lm}(\hat{\mathbf{r}}), \quad (5)$$

where $j_l(x)$ are the spherical Bessel functions of the first kind and order l; ξ_{nl} is the nth positive root ($n = 1,2,\ldots$) of the equation $j_l(x) = 0$, $Y_{lm}(\hat{\mathbf{r}})$ are the spherical functions (viewed here as functions of the polar and azimuthal angles of the unit vector $\hat{\mathbf{r}} = \mathbf{r}/r$), and

$$Z_{nl} = \sqrt{\frac{a^3}{2}} j_{l+1}(\xi_{nl}) \quad (6)$$

are normalization factors. The energy eigenstates are labeled by the main quantum number n, the orbital number l, and the magnetic number m. The unperturbed energy levels are given by the formula

$$E_{nl} = E_0 \xi_{nl}^2, \quad (7)$$

where

$$E_0 = \frac{\hbar^2}{2m_e a^2} \quad (8)$$

and m_e is the electron mass.

In what follows, we use the composite indices μ, ν, η, and ζ to label the eigenstates. Each composite index corresponds to the triplet of quantum numbers (n,l,m). By convention, if $\mu = (n,l,m)$, then $\mu' = (n',l',m')$. The matrix elements of the z projection of the dipole moment operator are given by

$$\hat{\mathbf{z}} \cdot \mathbf{d}_{\mu\mu'} = ea\Delta_{\mu\mu'}, \quad (9)$$

where

$$\Delta_{\mu\mu'} = \delta_{mm'} R_{nl}^{n'l'}(b_{lm}\delta_{l-1,l'} + b_{l+1,m}\delta_{l+1,l'}), \quad (10)$$

$\delta_{ll'}$ are the Kronecker delta symbols, $l,l' \geqslant 0$, and

$$b_{lm} = \sqrt{\frac{l^2 - m^2}{4l^2 - 1}}, \quad R_{nl}^{n'l'} = \frac{4\xi_{nl}\xi_{n'l'}}{\left(\xi_{nl}^2 - \xi_{n'l'}^2\right)^2}. \quad (11)$$

Note that the diagonal elements of Δ are all equal to zero, as is the case for any system with a center of symmetry. Finally, the matrix elements of the operator G are given by

$$G_{\mu\mu'} = -ea\Delta_{\mu\mu'} A_i. \quad (12)$$

The density matrix of the system, ρ, can be written in the form $\rho_{\mu\nu}(t) = \tilde{\rho}_{\mu\nu}(t) \exp(i\omega_{\mu\nu}t)$, where $\omega_{\mu\nu} = (E_\mu - E_\nu)/\hbar$ are the transition frequencies and $\tilde{\rho}_{\mu\nu}(t)$ is the so-called slow-varying amplitude, which obeys the following master equation:[9]

$$\left(\frac{\partial}{\partial t} + i\omega_{\mu\nu} + \Gamma_{\mu\nu}\right)\tilde{\rho}_{\mu\nu}$$
$$= \delta_{\mu\nu}\Gamma_{\mu\mu}N_\mu - \frac{i}{\hbar}\sum_\eta [V_{\mu\eta}(t)\tilde{\rho}_{\eta\nu} - \tilde{\rho}_{\mu\eta}V_{\eta\nu}(t)]. \quad (13)$$

Here N_μ are the equilibrium-state populations and $\Gamma_{\mu\nu}$ are phenomenological relaxation constants. Following Rautian, we assume that

$$\Gamma_{\mu\nu} = \Gamma_1\delta_{\mu\nu} + \Gamma_2(1 - \delta_{\mu\nu}). \quad (14)$$

Equation (14) is the least complex assumption on $\Gamma_{\mu\nu}$, which still distinguishes the relaxation rates for the diagonal and off-diagonal elements of the density matrix.

It can be seen that, for the case of zero external field, $\tilde{\rho}_{\mu\nu} = \delta_{\mu\nu}N_\mu$. The Fermi statistics is introduced at this point by writing

$$N_\mu = \frac{2}{\exp[(E_\mu - E_F)/(k_BT)] + 1}, \quad (15)$$

where E_F is the Fermi energy, k_B is Boltzmann's constant, T is the temperature, and the factor of two in the numerator accounts for the electron spin. Conservation of particles reads $\sum_\mu N_\mu = \mathcal{N}$. When $\mathcal{N} \gg 1$, the well-known analytical formula for the Fermi's energy,

$$E_F = E_0(3\pi^2)^{2/3}\left(\frac{a}{\ell}\right)^2 = (3\pi^2)^{2/3}\frac{\hbar^2}{2m_e\ell^2}, \quad (16)$$

holds with a good accuracy. Here ℓ is the characteristic atomic scale, defined by the relation

$$\ell^3 = \Omega/\mathcal{N}, \quad (17)$$

where

$$\Omega = 4\pi a^3/3 \quad (18)$$

is the nanoparticle volume. Thus, ℓ^3 is the specific volume per conduction electron. We note that ℓ is, generally, different from the lattice constant h. Many metals of interest in plasmonics have an fcc lattice structure with four conduction electrons per unit cell. In this case $h = 4^{1/3}\ell$. For example, in silver, $\ell \approx 0.26$ nm, $h \approx 0.41$ nm, and $E_F \approx 5.51$ eV. At room temperature ($T = 300$ K), $k_B T \approx 0.026$ eV, so that $T = 0$ is a good approximation. In this case, $N_\mu = 2$ if $E_\mu \leqslant E_F$ and $N_\mu = 0$ otherwise. Most numerical results shown below have been obtained in this limit. However, to illustrate the effects of finite temperature, we have also performed some computations at $T = 300$ K. Finally, the Fermi velocity is given by the equation

$$v_F = \sqrt{\frac{2E_F}{m_e}} = (3\pi^2)^{1/3} \frac{\hbar}{m_e \ell}. \tag{19}$$

In silver, $v_F \approx 1.2 \times 10^8$ cm/sec and, correspondingly, $c/v_F \approx 250$. Electron velocities in excited states are expected to be no larger than a few times Fermi velocity, still much smaller than c. This justifies the use of nonrelativistic quantum mechanics.

The solution to (13) has the form of a Fourier series:

$$\tilde{\rho}_{\mu\nu}(t) = \sum_{s=-\infty}^{\infty} \tilde{\rho}_{\mu\nu}^{(s)} \exp(-is\omega t). \tag{20}$$

The expansion coefficients $\tilde{\rho}_{\mu\nu}^{(s)}$ obey the system of equations:

$$\tilde{\rho}_{\mu\nu}^{(s)} = \delta_{\mu\nu}\delta_{s0}N_\mu - \frac{\Lambda_{\mu\nu}^{(s)}(\omega)}{\hbar\omega} \sum_\eta \left[G_{\mu\eta}\left(\tilde{\rho}_{\eta\nu}^{(s-1)} + \tilde{\rho}_{\eta\nu}^{(s+1)}\right) \right.$$
$$\left. - G_{\eta\nu}\left(\tilde{\rho}_{\mu\eta}^{(s-1)} + \tilde{\rho}_{\mu\eta}^{(s+1)}\right)\right], \tag{21}$$

where

$$\Lambda_{\mu\nu}^{(s)}(\omega) = \frac{\omega}{\omega_{\mu\nu} - s\omega - i\Gamma_{\mu\nu}} \tag{22}$$

are Lorentzian spectral factors. The optical response of the nanoparticle is determined by the quantum-mechanical expectation of its total dipole moment, which is given by

$$\langle d(t)\rangle = ea \sum_{\mu\nu} \Delta_{\mu\nu} \tilde{\rho}_{\mu\nu}(t). \tag{23}$$

Upon substitution of (20) into (23), we obtain the expansion of $\langle d(t)\rangle$ into temporal Fourier harmonics. We now consider the optical response at the fundamental frequency ω, which describes degenerate nonlinear phenomena such as the four-wave mixing. Denoting the component of $\langle d(t)\rangle$, which oscillates at the frequency ω, by $\langle d_\omega(t)\rangle$, we can write

$$\langle d_\omega(t)\rangle = D \exp(-i\omega t) + \text{c.c.}, \tag{24}$$

where

$$D = ea \sum_{\mu\nu} \Delta_{\mu\nu} \tilde{\rho}_{\mu\nu}^{(1)}. \tag{25}$$

The coefficients $\tilde{\rho}_{\mu\nu}^{(1)}$ and the amplitude D in (24) can be expanded in powers of A_i. Namely, we can write

$$D = \Omega\chi(A_i)A_i, \tag{26}$$

where

$$\chi(A_i) = \chi_1 + \chi_3(A_i/A_{\text{at}})^2 + \chi_5(A_i/A_{\text{at}})^4 + \cdots. \tag{27}$$

Here we have introduced the characteristic atomic field

$$A_{\text{at}} = e/\ell^2 \tag{28}$$

and have used the assumption that A_i is real valued; in the more general case, the expansion contains the terms of the form $\chi_3|A_i/A_{\text{at}}|^2$, etc. Note that the definition of χ_3 in (26) and (27) is somewhat unconventional. The nonlinear susceptibility $\chi^{(3)}$, as defined in standard expositions of the subject,[10] has the dimensionality of the inverse square of the electric field, that is, of $\text{cm}^2/\text{statvolt}^2 = \text{cm}^3/\text{erg}$ in the Gaussian system of units. Here we find it more expedient to define dimensionless coefficients χ_1, χ_3, χ_5, etc., and expand the dipole moment amplitude D in powers of the dimensionless variable A_i/A_{at}.

The first two coefficients in the expansion (27), χ_1 and χ_3, have been computed by Rautian explicitly and are given by the following series:

$$\chi_1 = \frac{(ea)^2}{(\hbar\omega)\Omega} \sum_{\mu\nu} N_{\mu\nu} \Lambda_{\mu\nu}^{(1)} \Delta_{\mu\nu} \Delta_{\nu\mu}, \tag{29a}$$

$$\chi_3 = \frac{(ea)^4 A_{\text{at}}^2}{(\hbar\omega)^3\Omega} \sum_{\mu\nu\eta\zeta} B_{\mu\nu}^{\zeta\eta} \Delta_{\mu\zeta} \Delta_{\zeta\eta} \Delta_{\eta\nu} \Delta_{\nu\mu}, \tag{29b}$$

where

$$B_{\mu\nu}^{\zeta\eta} = \Lambda_{\mu\nu}^{(1)} N_{\zeta\mu}\left[\Lambda_{\mu\eta}^{(0)}\left(\Lambda_{\mu\zeta}^{(1)} + \Lambda_{\mu\zeta}^{(-1)}\right) + \Lambda_{\mu\eta}^{(2)}\Lambda_{\mu\zeta}^{(1)}\right]$$
$$+ \Lambda_{\mu\nu}^{(1)} N_{\nu\eta}\left[\Lambda_{\zeta\nu}^{(0)}\left(\Lambda_{\eta\nu}^{(1)} + \Lambda_{\eta\nu}^{(-1)}\right) + \Lambda_{\zeta\nu}^{(2)}\Lambda_{\eta\nu}^{(1)}\right]$$
$$- \Lambda_{\mu\nu}^{(1)} N_{\eta\zeta}\left(\Lambda_{\zeta\eta}^{(1)} + \Lambda_{\zeta\eta}^{(-1)}\right)\left(\Lambda_{\zeta\nu}^{(0)} + \Lambda_{\mu\eta}^{(0)}\right)$$
$$- \Lambda_{\mu\nu}^{(1)} N_{\eta\zeta}\Lambda_{\zeta\eta}^{(1)}\left(\Lambda_{\zeta\nu}^{(2)} + \Lambda_{\mu\eta}^{(2)}\right). \tag{30}$$

Here $N_{\mu\nu} = N_\mu - N_\nu$. Note that all quantities inside the summation symbols are dimensionless and so are the factors in front of the summation symbols.

The expression (29b) involves a staggering 12-fold summation (recall that each composite index μ, ν, η, and ζ consists of three integer indices). Rautian used the mathematical formalism of irreducible spherical tensors and $6j$ symbols to perform summation over magnetic sublevels analytically and to reduce the expression to an 8-fold summation. However, this approach does not make use of the orbital selection rules, which are explicit in (10). In Sec. III, we will use the orbital selection rules to analytically reduce (29b) to a 5-fold summation. The resultant formula is amenable to direct numerical implementation, as will be illustrated in Sec. V.

Having performed the summation over the magnetic sublevels, Rautian has evaluated the resulting series by exploiting the following two approximations:

(1) Adopt the two-level approximation.[6] In this approximation, only the terms with $\mu = \eta$ and $\nu = \zeta$ are retained in the right-hand side of (29b).

(2) Assume that there are two dominant contributions to the series (29). The off-resonant (Drude) contribution is obtained by keeping only the terms with $\omega_{\mu\nu} \ll \omega$ in the Lorentzian factors $\Lambda_{\mu\nu}^{(s)}(\omega)$. The resonant contribution is obtained by keeping only the terms with $\omega_{\mu\nu} \approx \omega$. Each contribution is then evaluated separately by replacing summation with integration.

Using the same approximations, we have reproduced Rautian's analytical results. For χ_1, we obtained

$$\chi_1 = -\frac{1}{4\pi} \frac{\omega_p^2}{\omega + i\Gamma_2} \left[\frac{F_1}{\omega + i\Gamma_2} - i g_1 \frac{v_F/a}{\omega^2} \right], \quad (31)$$

where

$$\omega_p = \sqrt{\frac{4\pi e^2}{m_e \ell^3}} \quad (32)$$

is the plasma frequency and F_1, g_1 are dimensionless real-valued functions, which weakly depend on the parameters of the problem and are of the order of unity. More specifically, F_1 is very close to unity for all reasonable particle sizes and approaches unity asymptotically when $a \to \infty$ (we have verified this numerically). In what follows, we assume that $F_1 = 1$. The function g_1 depends most profoundly on the ratio $\kappa = \hbar\omega/E_F$. We can write, approximately,

$$g_1 \approx \frac{1}{\kappa} \int_{1-\kappa}^{1} x^{3/2}(x+\kappa)^{1/2}dx, \quad \kappa = \frac{\hbar\omega}{E_F}. \quad (33)$$

An analytical expression for this integral and a plot are given in the Appendix.

Equation (31) is equivalent to combining Eqs. (3.16) and (3.23) of Ref. 7. On physical grounds, one can argue that these expressions are applicable only if $\Gamma_2/\omega \ll 1$. Indeed, in the classical Drude model, we have

$$\chi_1^{\text{Drude}} = -\frac{1}{4\pi} \frac{\omega_p^2}{\omega(\omega + i\gamma)}, \quad (34)$$

where γ is a relaxation constant. We expect the Drude model to be accurate in the limit $a \to \infty$, when the second term in the square brackets in (31) vanishes. Thus, (31) has an incorrect low-frequency asymptote. We argue that the asymptote is incorrect because the HRFR model disregards the Hartree interaction potential. This will be discussed in more detail in Sec. IV below. At this point, we assume that $\Gamma_2/\omega \ll 1$ and expand (31) in Γ_2/ω, neglect the correction to the real part of the resulting expression, and obtain

$$\chi_1 \approx -\frac{1}{4\pi} \left(\frac{\omega_p}{\omega} \right)^2 \left[1 - i \frac{2\Gamma_2 + g_1 v_F/a}{\omega} \right]. \quad (35)$$

This expression corresponds to Eq. (3.28) of Ref. 7. Note that neglecting the correction to the real part but retaining the correction to the imaginary part in the above equation is mathematically justified because the terms $2\Gamma_2$ and v_F/a can be of the same order of magnitude, as we will see below.

Comparing (35) to a similar expansion of (34), we conclude that the size-dependent relaxation constant γ is given by

$$\gamma \approx \gamma_\infty + g_1 \frac{v_F}{a}, \quad \gamma_\infty = 2\Gamma_2, \quad (36)$$

where γ_∞ is the relaxation constant in bulk. It can be seen that the ratio v_F/a plays the role of the collision frequency. The analytical result (36) is widely known and frequently used; it will be confirmed by direct numerical evaluation of (29a) below.

For the third-order nonlinear susceptibility χ_3, we obtain, with the same accuracy as above,

$$\chi_3 = \frac{\Gamma_2}{\Gamma_1} \frac{\alpha^2}{10\pi^3} \left(\frac{a}{\ell} \right)^2 \left(\frac{\lambda_p}{\ell} \right)^2 \left(\frac{\omega_p}{\omega} \right)^4$$
$$\times \left[F_3 - i \left(F_3 \frac{\gamma_\infty}{\omega} + g_3 \frac{(v_F/a)^5}{\omega^3 \gamma_\infty^2} \right) \right]. \quad (37)$$

Here $\alpha = e^2/\hbar c \approx 1/137$ is the fine-structure constant, $\lambda_p = 2\pi c/\omega_p$ is the wavelength at the plasma frequency (≈ 138 nm in silver), and F_3, g_3 is another set of dimensionless real-valued functions of the order of unity. For realistic parameters, the function F_3 varies only slightly[7,8] between 0.30 and 0.33; we have taken $F_3 = 0.33$ in the numerical computations of Sec. V. The function g_3 can be approximated by the following integral:

$$g_3 \approx \frac{1}{\kappa} \int_{1-\kappa}^{1} x^{5/2}(x+\kappa)^{3/2}dx, \quad \kappa = \frac{\hbar\omega}{E_F}. \quad (38)$$

The approximate formula (38) applies only for $\hbar\omega < E_F$. However, we are interested in the spectral region $\omega \lesssim \omega_p$. In silver, $\hbar\omega_p \approx 8.98$ eV and $\hbar\omega_p/E_F \approx 1.63$. This leaves us with the spectral range $E_F/\hbar < \omega < \omega_p$ in which (38) is not applicable. The integral (38) can be evaluated analytically; the resulting expression and plot are given in the Appendix.

Expression (37) contains several dimensionless parameters. For a silver nanoparticle of the radius $a = 10$ nm,

$$\frac{\alpha^2}{10\pi^3} \left(\frac{a}{\ell} \right)^2 \left(\frac{\lambda_p}{\ell} \right)^2 \approx 71.6. $$

The ratio Γ_2/Γ_1 is more puzzling. While Γ_2 can be related to the Drude relaxation constant through (36), Γ_1 does not enter the analytical approximations (31) and (35) or the exact expression (29a). Therefore, Γ_1 cannot be directly related to any measurement of the linear optical response. It was previously suggested[8] that, based on the available experimental studies of nonequilibrium electron kinetics in silver,[11–13] $\Gamma_2/\Gamma_1 \sim 10$. This ratio will be employed below.

Another interesting question is the dependence of the results on the particle radius, a. It follows from the analytical approximation (35) that χ_1 approaches a well-defined "bulk" limit when $a \to \infty$. The characteristic length scale is v_F/γ_∞ (≈ 44 nm in silver). Of course, direct numerical evaluation of χ_1 according to (29a) is expected to reveal some dependence of χ_1 on a, which is not contained in the analytical approximation (35), and this fact will be demonstrated below in Sec. V B. However, it will also be demonstrated that (35) becomes very accurate in the spectral range of interest when $a \gtrsim 5$ nm. Thus, the HRFR model yields a result for χ_1, which is consistent with the macroscopic limit.

The situation is dramatically different in the case of the nonlinear susceptibility χ_3. It follows from (37) that $\chi_3 \xrightarrow{a \to \infty} O(a^2)$. Therefore, there is no "bulk" limit for χ_3. This is an unexpected result. While some studies suggest that a positive correlation between χ_3 and a in a limited range of a is consistent with experimental measurements,[8] we cannot expect this correlation to hold for arbitrarily large values of a, as this would, essentially, entail an infinite value of χ_3 in bulk. Such a prediction appears to be unphysical. Of course, Rautian's theory is not expected to apply to arbitrarily large

values of a because the interaction potential (2) is written in the dipole approximation and, moreover, it assumes that the electric field inside the nanoparticle is potential; that is, $\nabla \times \mathbf{E} = 0$ is a good approximation. Still, the absence of a "bulk" limit for χ_3 is troublesome. We, therefore, wish to understand whether the quadratic dependence of χ_3 on a is a property of the HRFR model itself or an artifact of the additional approximations made in deriving the analytical expression (37). More specifically, we can state the following two hypotheses:

(1) The quadratic dependence of χ_3 on a is an artifact of the approximations made in deriving the analytical expression (37) from (29b) [these approximations are listed explicitly between Eqs. (30) and (31)]. In this case, we can expect that direct evaluation of (29b) will not exhibit the quadratic growth.

(2) The quadratic dependence of χ_3 on a is a property of the HRFR model itself. In particular, the absence of a "bulk" limit for χ_3 can be caused by the following reasons: (i) The HRFR model neglects the retardation effects in large particles. (ii) The HRFR model does not account for the Hartree interaction potential. (In reality, interaction of the conduction electrons with the induced charge density may be important, especially, for computing nonlinear corrections.) (iii) The HRFR model makes use of a phenomenological boundary condition at the nanoparticle surface.

Verification of these hypotheses was previously hindered by the computational complexity of the problem. In what follows, we render Rautian's theory amenable to direct numerical validation. Then we show that the analytical approximation (37) is surprisingly good. Therefore, the second hypothesis must be correct.

III. RAUTIAN'S THEORY FURTHER DEVELOPED

It is possible to simplify (29b) without adopting any approximations. To this end, we deviate from Rautian's approach of using irreducible spherical tensors and $6j$ symbols. Instead, we directly substitute the expressions (9), (10), and (11) into (29b). We use the selection rules in (10) and the results

$$Z_l \equiv \sum_{m=-l}^{l} b_{lm}^4 = \frac{l(4l^2+1)}{15(4l^2-1)}, \tag{39a}$$

$$S_l \equiv \sum_{m=-l}^{l} b_{lm}^2 b_{l+1,m}^2 = \frac{2l(l+1)}{15(2l+1)} \tag{39b}$$

to evaluate summations over all magnetic quantum numbers and over all orbital quantum numbers but one. This leaves us with a fivefold summation over four main quantum numbers and one orbital quantum number. After some rearrangements, we arrive at the following expression:

$$\chi_3 = \frac{(ea)^4 A_{\mathrm{at}}^2}{(\hbar\omega)^3 \Omega} \sum_{l=1}^{\infty} (Z_l P_l + S_l Q_l), \tag{40}$$

where

$$P_l = \sum_{n_1,n_2,n_3,n_4} \left[B_{n_3,l-1,n_4,l}^{n_1,l,n_2,l-1} + B_{n_1,l,n_3,l-1}^{n_2,l-1,n_4,l} \right]$$

$$\times R_{n_3,l-1}^{n_1,l} R_{n_1,l}^{n_2,l-1} R_{n_2,l-1}^{n_4,l} R_{n_4,l}^{n_3,l-1}, \tag{41a}$$

$$Q_l = \sum_{n_1,n_2,n_3,n_4} \left[B_{n_3,l-1,n_4,l}^{n_1,l,n_2,l+1} + B_{n_1,l,n_3,l-1}^{n_2,l+1,n_4,l} \right.$$

$$+ B_{n_2,l+1,n_1,l}^{n_4,l,n_3,l-1} + \left. B_{n_4,l,n_2,l+1}^{n_3,l-1,n_1,l} \right]$$

$$\times R_{n_3,l-1}^{n_1,l} R_{n_1,l}^{n_2,l+1} R_{n_2,l+1}^{n_4,l} R_{n_4,l}^{n_3,l-1}. \tag{41b}$$

This expression is exact within the HRFR model. The two-level approximation corresponds to keeping only the first term in the brackets in (40) and, further, keeping only the terms with $n_2 = n_3$ and $n_1 = n_4$ in (41a).

IV. RELATING THE INTERNAL AND THE APPLIED FIELDS

In the HRFR model, electrons move in a given, spatially uniform internal field (1). In practice, one is interested in the optical response of the nanoparticle to the external (applied) field. We denote the amplitude of the external field by $\mathbf{A}_e = A_e \hat{\mathbf{z}}$. The two fields differ because of a charge density induced in the nanoparticle. The interaction of the conduction electrons with the induced charge density is described by the Hartree potential. However, rigorous introduction of the Hartree interaction into the HRFR model is problematic. Doing so would require the mathematical apparatus of density-functional theory. We can, however, apply here the classical concept of the depolarizing field, although this approach is less fundamental.

In the macroscopic theory, a sphere (either dielectric or conducting), when placed in a spatially uniform, quasistatic electric field of frequency ω and amplitude $\mathbf{A}_e$, is polarized and acquires a dipole moment of an amplitude $\mathbf{D}$. The electric field inside the sphere is also spatially uniform and has the amplitude $\mathbf{A}_i$. The induced charge accumulates at the sphere surface in a layer whose width is neglected. Under these conditions, $\mathbf{A}_i = \mathbf{A}_e - \mathbf{D}/a^3$. Note that a linear dependence between $\mathbf{D}$ and $\mathbf{A}_e$ is not assumed here. The form of the depolarizing field, $-\mathbf{D}/a^3$, follows only from the assumption of spatial uniformity of the internal field and from the usual boundary conditions at the sphere surface. Then the Hartree interaction can be taken into account as follows.

Let us introduce the dimensionless variables $x = A_i/A_{\mathrm{at}}$ and $y = A_e/A_{\mathrm{at}}$. Then we can expand D in both variables:

$$D = \Omega A_{\mathrm{at}}(\chi_1 x + \chi_3 x|x|^2 + \chi_5 x|x|^4 + \cdots), \tag{42a}$$

$$D = \Omega A_{\mathrm{at}}(\alpha_1 y + \alpha_3 y|y|^2 + \alpha_5 y|y|^4 + \cdots), \tag{42b}$$

where

$$x = y - \frac{D}{A_{\mathrm{at}} a^3} = y - \frac{4\pi}{3}(\alpha_1 y + \alpha_3 y|y|^2 + \cdots). \tag{43}$$

Here we have accounted for the fact that there can be a phase shift between the internal and the external fields; therefore, A_i and A_e cannot be real valued *simultaneously*. In the theory presented above, we assume that A_i is real valued, and this can always be guaranteed by appropriately choosing the time origin. In this case, A_e is expected to be complex.

The coefficients χ_k in (42a) can be found from Rautian's theory; our task is to find the coefficients α_k in (42b) given the constraint (43). To this end, we substitute (43) into (42a) and obtain a series in the variable y. We then require that the

coefficients in this series and in (42b) coincide. This yields an infinite set of equations for α_k, the first two of which read

$$\chi_1 \left(1 - \frac{4\pi}{3} \alpha_1 \right) = \alpha_1, \tag{44a}$$

$$\chi_3 \left(1 - \frac{4\pi}{3} \alpha_1 \right) \left| 1 - \frac{4\pi}{3} \alpha_1 \right|^2 - \chi_1 \frac{4\pi}{3} \alpha_3 = \alpha_3. \tag{44b}$$

It is convenient to introduce the linear field enhancement factor f_1 according to

$$f_1 = \frac{1}{1 + (4\pi/3)\chi_1} = \frac{3}{\epsilon_1 + 2}, \tag{45}$$

where $\epsilon_1 = 1 + 4\pi \chi_1$ is the linear dielectric permittivity. Then the solutions to (44) have the form

$$\alpha_1 = f_1 \chi_1, \quad \alpha_3 = f_1^2 |f_1|^2 \chi_3. \tag{46}$$

The factor f, which relates the external and internal field amplitudes according to $A_i = f A_e$, is then found from

$$f = \frac{x}{y} = 1 - \frac{4\pi}{3}(\alpha_1 + \alpha_3 |y|^2 + \cdots). \tag{47}$$

Using (46), we find that, to first order in $I/I_{\text{at}} \equiv |y|^2 = |A_e/A_{\text{at}}|^2$,

$$f = f_1 - \frac{4\pi}{3} f_1^2 |f_1|^2 \chi_3 \frac{I}{I_{\text{at}}}. \tag{48}$$

Here we have introduced the intensity of the incident beam, $I = (c/2\pi)|A_e|^2$, and the "atomic" intensity $I_{\text{at}} = (c/2\pi)|A_{\text{at}}|^2$.

Note that our approach to finding the field enhancement factor f is somewhat different from that adopted in Ref. 8, where the expansion (42a) has been truncated at the third order and the truncated expression was assumed to be exact. The results obtained in the two approaches coincide to first order in I/I_{at}. In Ref. 8, higher order corrections to this result have also been obtained. In our approach, these corrections depend on the higher order coefficients χ_5, χ_7, etc., which have not been computed by Rautian.

We finally note that the phenomenological accounting for the Hartree interaction described in this section, while necessary for comparison with the experiment, does not remove the two main difficulties of the HRFR model. Specifically, it does not fix the low-frequency limit for χ_1 and does not affect the $\propto a^2$ dependence of χ_3. Regarding the low-frequency limit, we note that $\lim_{\omega \to 0} f = 0$ and the internal field in the nanoparticle tends to zero in this limit. The induced macroscopic charge is localized at the sphere surface where the electric field jumps abruptly. In a more accurate microscopic picture, the width of this surface layer is finite and the electric field changes smoothly over the width of this layer. Unfortunately, the classical concept of depolarizing field cannot capture surface phenomena of this kind.

V. NUMERICAL RESULTS

A. Convergence

We have computed the Bessel function zeros ξ_{nl} using the method of bisection and achieved a numerical discrepancy of the equation $j_l(\xi_{nl}) = 0$ of less than 10^{-15} for all values of
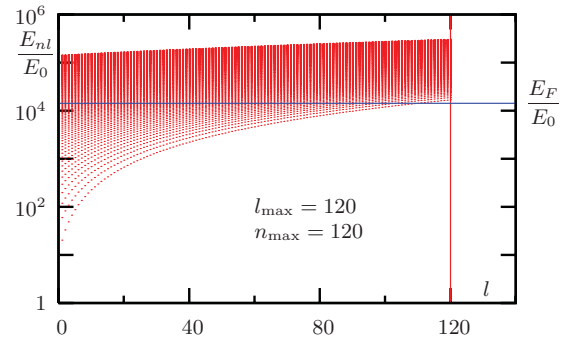

FIG. 1. (Color online) Energy eigenstates, which enter the summation according to (40) and (41), for $a = 10$ nm and $l_{\max} = n_{\max} = 120$. The horizontal blue line shows the Fermi energy. In this example, the total number of states below the Fermi surface is $\mathcal{N} \approx 2.4 \times 10^5$ (counting all degeneracies) and the total number of states shown in the figure is $2n_{\max}l_{\max}(l_{\max} + 2) \approx 3.5 \times 10^6$.

indices. Since the function $j_l(x)$ is approximately linear near its roots, we believe that we have computed ξ_{nl} with sufficiently high precision.

The summation over l in (40) was truncated so that $l \leqslant l_{\max}$ and the quadruple summation in (41) was truncated so that $n_1, n_2, n_3, n_4 \leqslant n_{\max}$. A typical set of energy levels used in the summation is shown in Fig. 1 for the case $a = 10$ nm, $l_{\max} = n_{\max} = 120$. Here the energy levels (normalized to E_0) are shown by dots and the horizontal axis corresponds to the orbital number, l. Referring to Fig. 1, we note that $l_{\max}$ has been chosen so that all states with $l \geqslant l_{\max}$ are above the Fermi surface. Since the electron transitions occur between two states with l and l' such that $l' = l \pm 1$, the factors $N_{\mu\nu}$ for any transition involving the states with $l \geqslant l_{\max}$ are zero (or exponentially small at finite temperatures). It can be seen that convergence with l is very fast—the contribution of the terms in (40) with $l \geqslant l_{\max}$ is either zero (at $T = 0$) or exponentially small.

The choice of $n_{\max}$ is a more subtle matter. Since there are no selection rules on n, transitions can occur between two states (one below and one above the Fermi surface) with very different values of n and, correspondingly, very different energies. However, transitions with energy gaps much larger than $\hbar\omega$ are suppressed by the Lorentzian factors (22). In most numerical examples, we have chosen $n_{\max}$ so as to account for, at least, all transitions with the energy gaps of $\Delta E \leqslant 3\hbar\omega$. Many (but not all) transitions with larger energy gaps were also accounted for. This approach yields a result with seven significant figures. However, it results in too many terms in the summation when $\hbar\omega \sim E_F$ and $a \geqslant 10$ nm. For these values of parameters, we have used a smaller $n_{\max}$ so as to account for, at least, all transitions with $\Delta E \leqslant \hbar\omega$. We estimate that the relative error incurred by this truncation is $\lesssim 10\%$.

B. Linear response

We begin by considering the linear susceptibility χ_1. In computations, we use the commonly accepted parameters for silver, $\hbar\omega_p = 8.98$ eV ($\lambda_p = 2\pi c/\omega_p = 138$ nm) and $\gamma_\infty/\omega_p = 0.002$. Here the relaxation constant Γ_2, which enters (29a), is determined from $\gamma_\infty = 2\Gamma_2$ [see (36)]. The frequencies used satisfy the condition $\gamma_\infty/\omega \ll 1$. More specifically,
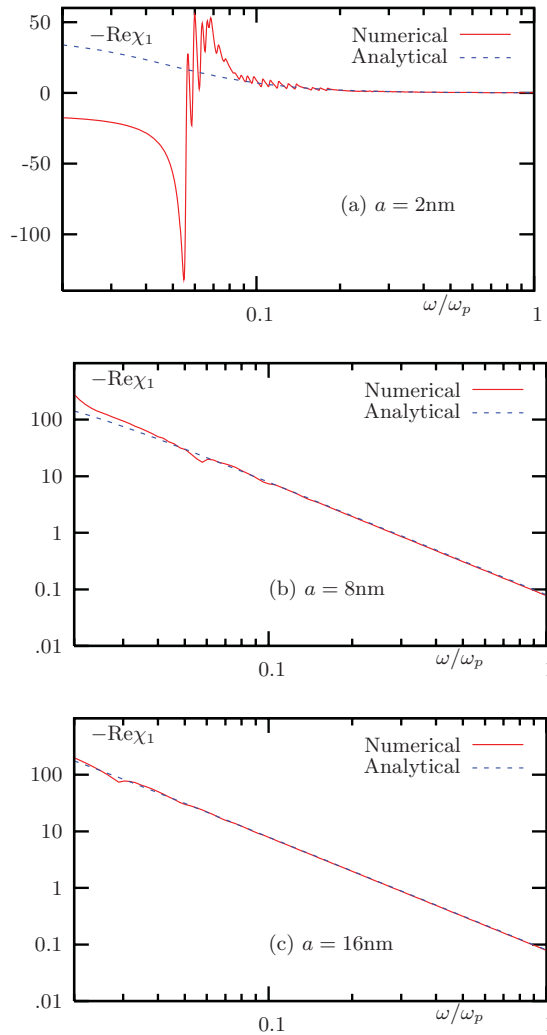
FIG. 2. (Color online) The quantity $-\mathrm{Re}\chi_1$ as a function of frequency for particles of different radius a, as labeled. Centered symbols correspond to direct numerical evaluation of (29a) and continuous curves show the Drude formula (34) in which the size-corrected relaxation constant γ (36) has been used.

FIG. 3. (Color online) The quantity $\mathcal{Z}$ defined in (49) as a function of frequency for different particle radiuses, as labeled. Results of direct numerical computation are compared to the size-corrected Drude relaxation constant γ [given by (36)] and to its bulk value γ_∞ (obtained in the limit $a \to \infty$).

the ratio ω/ω_p varies in the range $0.02 \leqslant \omega/\omega_p \leqslant 1$. We do not consider the frequencies above ω_p because silver exhibits strong interband absorption in that spectral range. Except when noted otherwise, all computations have been carried out at $T = 0$.

Figure 2 displays the quantity $-\mathrm{Re}\chi_1$ computed numerically by direct evaluation of (29a) and by the Drude formula (34) with the size-corrected relaxation constant γ (36). The factor g_1 in (36) has been computed using the analytical formula (A1). At sufficiently high frequencies, the Drude model predicts that $-4\pi\mathrm{Re}\chi_1 \approx (\omega_p/\omega)^2$ and this behavior is reproduced for all radiuses considered with good precision. However, at smaller frequencies, there are differences between the analytical approximation and the numerical results. These differences are especially apparent for $a = 2$ nm. In this case, the optical response of the sphere is, effectively, dielectric rather than metallic for $\omega \lesssim 0.06\omega_p$. A similar behavior has been observed at $a = 4$ nm (data not shown). The emergence
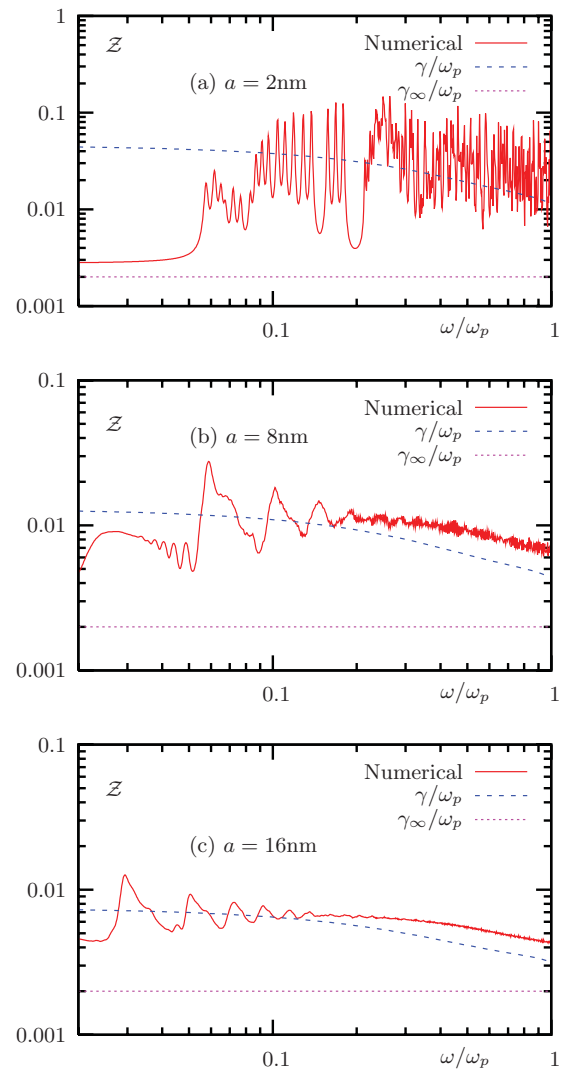
of a dielectric response in metal nanoparticles of sufficiently small size at sufficiently low frequencies has been overlooked in the past. It occurs due to discreteness of the energy states. Consider a particle with $a = 2$ nm at zero temperature. In this case, the lowest energy electronic transition, which is allowed by Fermi statistics (that is, a transition with $N_{\mu\nu} \neq 0$), occurs between the states $(n = 1, l = 18)$ and $(n' = 1, l' = 19)$. The corresponding transition frequency is $\omega_{\min} \approx 0.056\omega_p$. It can be seen from Fig. 2(a) that the particle becomes dielectric for $\omega \lesssim \omega_{\min}$.

We now turn to consideration of the relaxation phenomena. To this end, we plot in Fig. 3 the quantity

$$\mathcal{Z} = -\frac{1}{4\pi}\frac{\omega_p}{\omega}\mathrm{Im}\frac{1}{\chi_1} \qquad (49)$$

as a function of frequency. We note that $\mathcal{Z}$, as defined in (49), is positive for all passive materials and, in the Drude model, $\mathcal{Z} = \gamma/\omega_p$; here γ is size corrected. It can be

seen that the analytical formula (36) captures the relaxation phenomena in the nanoparticle surprisingly well. However, as in Fig. 2(a), the analytical approximation breaks down when $a = 2$ nm and $\omega \lesssim 0.06\omega_p$. A similar breakdown was observed for $a = 4$ nm (data not shown). For all other values of parameters, the numerically computed $\mathcal{Z}$ is reasonably close to the size-corrected value of γ/ω_p and exhibits the same overall behavior. The small systematic error at higher frequencies is, most likely, caused by the approximation (33) for g_1. It was, in fact, mentioned by Rautian that (33) is hardly accurate when $\hbar\omega \sim E_F$.

The fine structure visible in Figs. 3(a) and 3(b) is due to discreteness of electron states. The allowed transition frequencies can be "grouped," which results in the appearance of somewhat broader peaks, clearly seen in Figs. 3(b) and 3(c). While spectral signatures of discrete states in metal nanoparticles have been observed experimentally (including the effect of "grouping"),[14] the positions of individual spectral peaks should not be invested with too much significance. In any realistic system, these peaks will be smoothed out by particle polydispersity, by variations in shape, and by nonradiative relaxation and energy transfer to the surrounding medium.

The finite-size correction (36) to the Drude relaxation constant is widely known and used. However, the derivations of (36) have been, so far, either heuristic or relied on poorly controlled approximations. In Fig. 3, we have provided, to the best of our knowledge, the first direct, first-principle numerical verification of (36) and of its limits of applicability.

C. Nonlinear response

We next turn to the nonlinear susceptibility χ_3. The same parameters for silver as before will be used. In addition, the calculations require the relaxation constant Γ_1. As was mentioned above, the experimental value of Γ_1 cannot be inferred by observing the linear optical response. It was previously suggested[8] that $\Gamma_2/\Gamma_1 \approx 10$. This value will be used below.

In Fig. 4, we plot the absolute value of χ_3 as a function of the particle radius a for $\omega = 0.1\omega_p$ and for the Frohlich frequency $\omega = \omega_p/\sqrt{3} \approx 0.58\omega_p$, and compare the results of direct numerical evaluation of (40) to the analytical approximation (37). In the case $\omega = 0.1\omega_p$, the analytical approximation is very accurate for $a \gtrsim 8$ nm and gives the correct overall trend for $a \lesssim 4$ nm. A systematic discrepancy of unknown origin between the approximate and the numerical results is observed for 4 nm $< a <$ 8 nm. In the case $\omega = \omega_p/\sqrt{3}$, the analytical approximation gives the correct trend in the whole range of a considered. Note that, in the case $\omega = 0.1\omega_p$, the absolute value of χ_3 is dominated by Reχ_3 and Imχ_3 for large and small values of a, respectively. When $\omega = \omega_p/\sqrt{3}$, the real part of χ_3 is dominating for all values of a used in the figure.

Consider first particles with $a \lesssim 4$ nm. As expected, the discreteness of energy levels plays an important role in this case and results in a series of sharp maxima and minima of $|\chi_3(a)|$. As is shown in the inset of Fig. 4(a), the function $|\chi_3(a)|$ is discontinuous. These discontinuities are artifacts of the zero-temperature approximation. The introduction of a finite temperature ($T = 300$ K) removes the discontinuities (see the inset) but does not eliminate the fine structure of
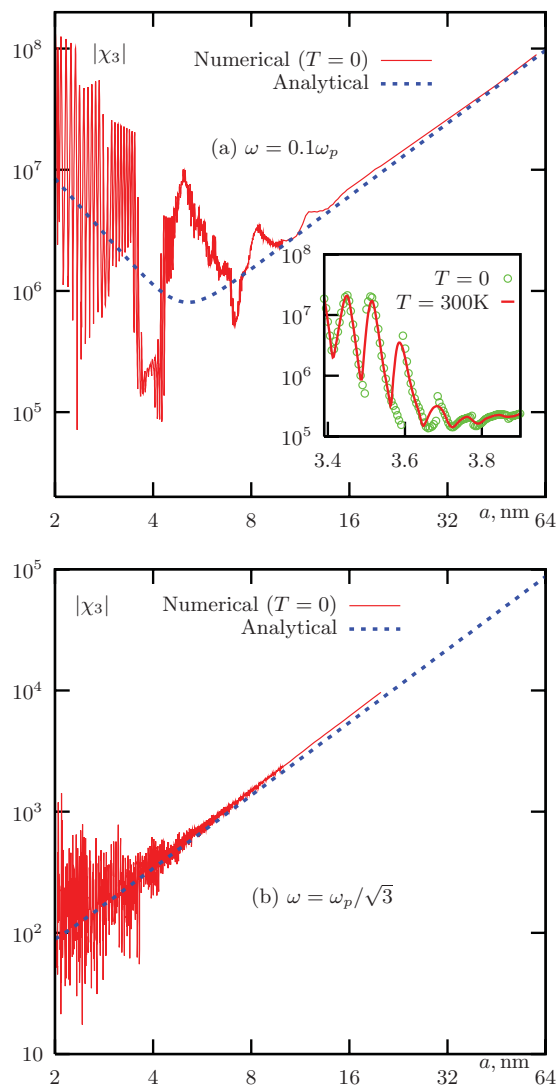

FIG. 4. (Color online) Absolute value of the nonlinear susceptibility, $|\chi_3|$, computed by direct evaluation of (40) and by analytical approximation (37) as a function of the particle radius for $\omega = 0.1\omega_p$ (a) and for $\omega = \omega_p/\sqrt{3} \approx 0.58\omega_p$ (b). Logarithmic scale is used on both axes. The inset in panel (a) shows a zoom of the plot for 3.4 nm $\leqslant a \leqslant$ 3.7 nm. In the inset, the results of evaluating (40) at $T = 0$ and at $T = 300$ K are shown.

the curve. Note, however, that the computations have been carried out with a very fine step in a, which is, arguably, unphysical: The parameter a in a real nanosphere can change only in quantized steps of the order of the lattice constant h ($\approx$ 0.41 nm in silver). Moreover, the fine structure of χ_3 is unlikely to be observable experimentally due to the unavoidable effects of particle polydispersity. Therefore, the general trend given by the analytical approximation (37) can be a more realistic estimate of χ_3 for $a \lesssim 4$ nm.

Next consider the large-a behavior. For $a \gtrsim 8$ nm, the analytical and the "exact" formulas yield results, which are scarcely distinguishable. In particular, the quadratic growth of χ_3 with a has been confirmed up to $a = 64$ nm in the case $\omega = 0.1\omega_p$—the largest radius for which numerical evaluation of (40) is still feasible. This confirms hypothesis 2 stated above, namely, that the quadratic growth of $\chi_3(a)$
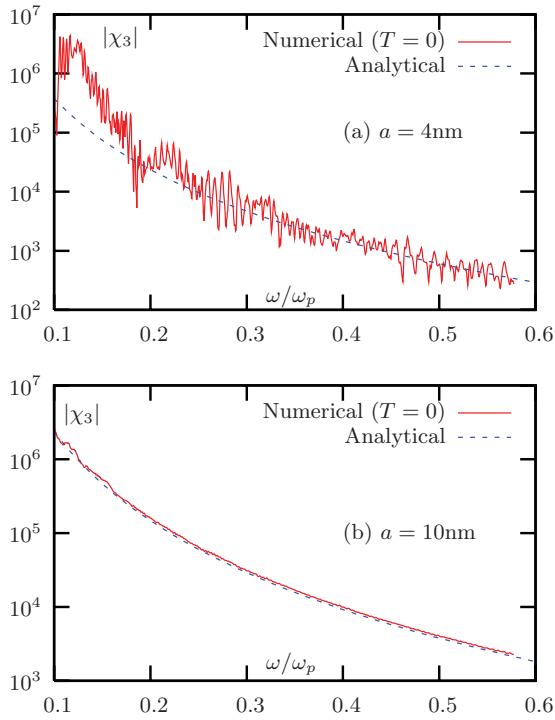
FIG. 5. (Color online) Absolute value of the nonlinear susceptibility, $|\chi_3|$, computed by direct evaluation of (40) at zero temperature and by analytical approximation (37) as a function of the frequency for $a = 4$ nm (a) and for $a = 10$ nm (b).

is a property of the HRFR model itself rather than of the additional approximations, which were made to derive the analytical results.

In Fig. 5, we study the dependence of $|\chi_3|$ on the frequency ω for fixed values of a. It can be seen that the accuracy of Rautian's approximation improves for larger particles and higher frequencies. At $a = 10$ nm, the approximation is nearly perfect in the full spectral range considered.

D. Magnitude of the nonlinear effect and comparison with the classical theory of electron confinement

In the previous subsection, we have plotted the coefficient χ_3, which appears in the expansion (27). The dimensionless parameter of this expansion, A_i/A_{at}, contains the amplitude of the internal electric field, A_i. However, it is the amplitude of the external (applied) field, A_e, which can be directly controlled in an experiment. The incident beam intensity is given by $I = (c/2\pi)|A_e|^2$. We can use the results of Sec. IV to write

$$D = \Omega A_e[\alpha_1 + \alpha_3(I/I_{\mathrm{at}}) + \cdots], \tag{50}$$

where α_1 and α_3 are related to χ_1 and χ_3 by (46) and $I_{\mathrm{at}} = (c/2\pi)A_{\mathrm{at}}^2$ is the characteristic "atomic" intensity. The quantity I_{at} can be expressed in terms of the fundamental physical constants and the material-specific parameter ℓ. In the case of silver, $I_{\mathrm{at}} \approx 2.3 \times 10^{14}$ W/cm^2. Obviously, intensities of such magnitude are not achievable in any experiment. However, the magnitude of the nonlinear correction can be amplified by the two important effects:[10] the effect of synchronism (constructive interference), which is not considered here, and

the effect of local field enhancement, which will be taken into account by using the expressions derived in Sec. IV.

We will also compare the expression (50) to the results obtained from the purely classical arguments.[15] In Ref. 15, we have argued that the surface charge in a polarized metal nanoparticle cannot be confined to an infinitely thin layer. When the width of this layer is not negligible (compared to the particle radius), a nonlinear correction to the particle polarizability is obtained. After some rearrangement of the formulas derived in this reference, we can express the amplitude D, defined analogously to (24), as

$$D = \Omega A_e[\beta_1 + \beta_3\sqrt{I/I_{\mathrm{at}}} + \cdots], \tag{51}$$

where

$$\beta_1 = \frac{3}{4\pi}\frac{\omega_p^2/3}{\omega_p^2/3 - \omega^2 - i\gamma\omega}, \tag{52a}$$

$$\beta_3 = -\frac{3}{\pi}\frac{\ell}{a}\beta_1|\beta_1|. \tag{52b}$$

We note that $\alpha_1 \xrightarrow{\omega/\gamma \to \infty} \beta_1$. That is, the linear polarizabilities of both theories are the same in the region of parameters where the theories are applicable. The classical theory, however, does not contain the low-frequency anomaly in the linear polarizability. On the other hand, relaxation is introduced in Ref. 15 through the phenomenological parameter γ whose dependence on a cannot be deduced theoretically. Below, we use the result of the quantum theory, namely, Eq. (36), for the relaxation constant γ (52a).

It can be seen that the classical and quantum expressions for D are quite different. The first nonvanishing nonlinear correction in (52b) is of the order of $\sqrt{I/I_{\mathrm{at}}}$ but contains an additional small parameter ℓ/a. Thus, the nonlinear correction depends differently on the incident intensity, frequency, and particle radius a in the two theories. Additionally, Rautian's theory contains the parameter Γ_1/Γ_2, which does not enter into the classical theory. These factors complicate a direct comparison of the two results. We will, therefore, focus on the trends for one particular value of the incident power, $I = 10$ kW/cm^2. One should bear in mind that the nonlinear corrections depend on the incident power differently in the two theories.

In Fig. 6, we plot the absolute value of the nonlinear correction to the particle polarizability normalized by its volume as a function of radius for the same values of frequency as were used in Fig. 4. We denote the quantity being plotted by $\mathscr{D}_{\mathrm{NL}}$, and

$$\mathscr{D}_{\mathrm{NL}} \equiv \begin{cases} \alpha_3(I/I_{\mathrm{at}}), & \text{in the "quantum" case,} \\ \beta_3\sqrt{I/I_{\mathrm{at}}}, & \text{in the "classical" case.} \end{cases} \tag{53}$$

The nonlinear effects should be observable in measurements with incoherent light if $|\mathscr{D}_{\mathrm{NL}}| \gtrsim 1$. If $|\mathscr{D}_{\mathrm{NL}}| \ll 1$, detection of the nonlinear effects requires coherent laser excitation and utilization of the effect of synchronism.

The parameters used in Fig. 6 are such that the approximate analytical formulas for χ_1 (34) and χ_3 (37) are fairly accurate, as was demonstrated above. Correspondingly, we have used these formulas to generate the curves, which are displayed in Fig. 6. To obtain the "quantum" curves, the following
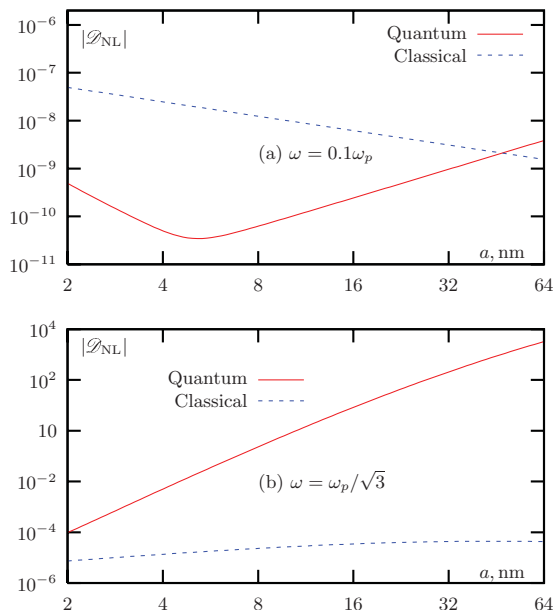
FIG. 6. (Color online) Absolute value of the nonlinear correction to the nanoparticle polarizability, $\mathscr{D}_{\mathrm{NL}}$, computed using Eq. (50) (the "quantum" curves) and Eq. (51) (the "classical" curves) for $\omega = 0.1\omega_p$ (a) and $\omega = \omega_p/\sqrt{3} \approx 0.58\omega_p$ (b). Here $\mathscr{D}_{\mathrm{NL}} = \alpha_3(I/I_{\mathrm{at}})$ (the "quantum" curves) and $\mathscr{D}_{\mathrm{NL}} = \beta_3\sqrt{I/I_{\mathrm{at}}}$ (the "classical" curves). The incident power is $I = 10$ kW/cm^2, $I/I_{\mathrm{at}} \approx 4.3 \times 10^{-11}$. To compute α_3, the internal field enhancement factor has been taken into account according to (46).

procedure has been followed. First, we have computed the function $\chi_3(a)$ according to (37) for each frequency considered. Then we have computed $\chi_1(a)$ according to (34) for the same frequencies. In Eq. (34), we have accounted for the dependence of the relaxation constant γ on a according to (36). The computed function $\chi_1(a)$ was used to compute the linear field enhancement factor $f_1(a)$ according to (45). Finally, we have used the functions $f_1(a)$ and $\chi_3(a)$ to compute $\alpha_3(a)$ according to (46). The result was multiplied by $I/I_{\mathrm{at}} \approx 4.3 \times 10^{-11}$. In the "classical" case, β_3 was computed according to (52), where the relaxation constant γ was size corrected according to (36).

We now discuss the curves shown in Fig. 6 in more detail. First, in the "quantum" case, $\mathscr{D}_{\mathrm{NL}}$ exhibits an unlimited growth with a when $a \to \infty$. In the classical case, this growth is suppressed. As can be seen, the "classical" $\mathscr{D}_{\mathrm{NL}}$ decreases with a in the case $\omega = 0.1\omega_p$ and seems to reach a finite limit in the case $\omega = \omega_p/\sqrt{3}$. In reality, however, the "classical" curve in Fig. 6(b) reaches a maximum at $a \approx 44$ nm and then slowly approaches zero (the range of radiuses, which is necessary to see this behavior clearly, is not shown in the figure). In the classical theory, the nonlinearity is an effect of the finite size, which vanishes in the limit $a \to \infty$.

Second, when $\omega = 0.1\omega_p$, the local-field enhancement factor in the "quantum" theory is $|f_1|^4 \sim 10^{-8}$. That is, the field is effectively screened in the interior of the nanoparticle. Correspondingly, the nonlinear effect is very weak. In the "classical" theory, the field enhancement factor is different; namely, it is $|\beta_1|^2 \sim 1$. This dramatic difference is explained by the fact that the classical theory considers the induced electron

density near the nanoparticle surface where the electric field is not entirely screened. At the Frohlich frequency, $\omega = \omega_p/\sqrt{3}$, the situation is quite different: We have

$$|f_1|^4 \sim \omega_p^4/9\gamma^4 \xrightarrow{a \to \infty} \omega_p^4/9\gamma_\infty^4 \approx 7 \times 10^9.$$

Correspondingly, the "quantum" nonlinear correction can become very large at the Frohlich frequency; this is illustrated in Fig. 6(b). This result is probably unphysical—one cannot expect that $\mathscr{D}_{\mathrm{NL}} \sim 10^3$ at the modest incident intensity of 10 kW/cm^2. The classical curve, however, is still bounded at the Frohlich frequency below 10^{-4}. We can conclude therefore that the local-field correction plays a disproportionate role in the quantum theory and that, if used unscrupulously, it can predict utterly unrealistic magnitudes of the nonlinear effect.

VI. SUMMARY OF FINDINGS AND DISCUSSION

In this paper, we have further developed the quantum theory of Refs. 6,7 (the HRFR model). The goal was to describe the frequency and size dependence of linear and nonlinear optical susceptibilities of spherical metal nanoparticles. We have used the HRFR model without modification but have managed to simplify the previously published expressions to a point where these expressions became amenable to direct numerical implementation. Then, we have computed the linear and nonlinear susceptibilities numerically for various frequencies and various particle sizes and compared the obtained results to Rautian's analytical approximations. Previously, numerical computations of this kind have been hindered by the overwhelming computational complexity of the problem. We have also compared the predictions of the quantum theory of size-dependent optical susceptibilities with the predictions of a purely classical theory of Ref. 15. The following findings can be reported:

(1) We have found that the approximate formulas derived by Rautian[7] are surprisingly accurate in a wide range of parameters despite the use of a number of approximations. In particular, we have, for the first time, verified from first principles the correctness of the widely used finite-size correction to the Drude relaxation constant (36).

(2) We have found that, for sufficiently small values of radius and frequency, Rautian's approximations break down due to the discreteness of electron energy levels. At sufficiently small frequencies, a silver particle with $a \lesssim 4$ nm in radius behaves as a dielectric. This effect is illustrated in Fig. 2(a) for $a = 2$ nm.

(3) We have found that phenomenologically accounting for the local-field correction (see Sec. IV for details) does not remove the two main difficulties, which are encountered in the HRFR model, namely, the incorrect small-ω asymptote for the linear susceptibility χ_1 and the absence of a "bulk" limit for the nonlinear susceptibility χ_3. It appears that obtaining the correct asymptotes requires the rigorous account for the Hartree interaction potential. It is also conceivable that obtaining the correct large-a asymptote requires accounting for the retardation effects. However, the classical theory of Ref. 15 is quasistatic but does not possess a large-a anomaly. This suggests that the main focus in further development of Rautian's theory should be on a more accurate inclusion of Hartree interaction.
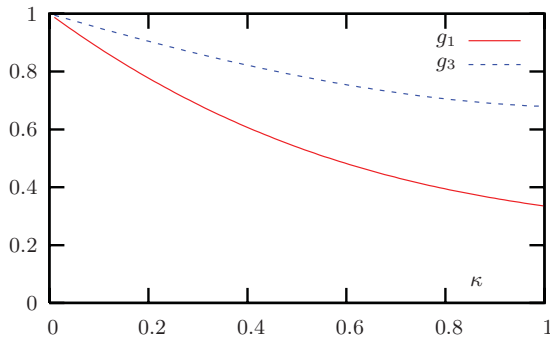
FIG. 7. (Color online) Functions $g_1(\kappa)$ and $g_3(\kappa)$.

One additional comment on the theory developed here is necessary. First, we have computed only a particular case of the nonlinear susceptibility $\chi^{(3)}(\omega; \omega_1, \omega_2, \omega_3)$. More specifically, the coefficient χ_3 defined in (27) is related to the latter quantity by $\chi_3 = A_{\text{at}}^{-2} \chi^{(3)}(\omega; \omega, -\omega, \omega)$. However, consideration of transient processes, generation of combination frequencies, and harmonics requires the knowledge of $\chi^{(3)}(\omega; \omega_1, \omega_2, \omega_3)$ as a function of all of its arguments. This is an important consideration. High incident intensities are usually obtained in short laser pulses. Moreover, many modern photonics applications such as waveguiding, etc., utilize short wave packets. Therefore, a proper description of optical nonlinearities in a transient process is very important. Generalizing the mathematical formalism described in this work to include three independent frequencies is not conceptually difficult, although it can lead to cumbersome calculations.

In summary, the HRFR model forms a perfect theoretical framework for studying optical nonlinearities and finite-size effects in nanoparticles. The only viable alternative to using this model is to resort to density functional theory (DFT). In a recent paper,[16] we have applied DFT to study the nonlinear electromagnetic response of metal nanofilms, but only at very low frequencies, well below plasmonic resonance of the system, and neglecting the relaxation phenomena. Higher frequencies, which are of interest in plasmonics, can be studied with the use of time-dependent DFT (TDDFT). Although TDDFT has been used successfully to compute linear response of nanoparticles,[17–19] and, in particular, to study the effects of surface adsorption of various molecules on the relaxation phenomena in metal,[20,21] the difficulties here are formidable. Most importantly, there is almost no hope of obtaining analytical approximations within DFT. It appears, therefore, that devising a way to include the Hartree interaction potential in the master equation (13) would be a useful and practically relevant development of the HRFR model and of Rautian's theory. Perhaps, some elements of DFT can be used to achieve this.

APPENDIX: FUNCTIONS $g_1(\kappa)$ AND $g_3(\kappa)$

The integrals (33) and (38) can be evaluated analytically with the following results:

$$\kappa g_1(\kappa) = -\frac{2}{9} + \frac{6\kappa}{7} - \frac{6\kappa^2}{5} + \frac{2\kappa^3}{3} + \frac{2\sqrt{1+\kappa}}{315}$$
$$\times (35 + 5\kappa - 6\kappa^2 + 8\kappa^3 - 16\kappa^4), \quad \text{(A1)}$$

$$640\kappa g_3(\kappa) = \{128 + \kappa(\kappa + 2)[88 + 5\kappa(3\kappa - 8)]\}\sqrt{1+\kappa}$$
$$- \{128 + \kappa(\kappa - 2)[168 - 5\kappa(3\kappa + 8)]\}\sqrt{1-\kappa}$$
$$+ 15\kappa^5 \ln \frac{1 + \sqrt{1-\kappa}}{1 + \sqrt{1+\kappa}}. \quad \text{(A2)}$$

The above expressions have been obtained from (33) and (38) without using any approximations. However, it should be kept in mind that (33) is valid for $-1 \leqslant \kappa$, while (38) is valid for $-1 \leqslant \kappa \leqslant 1$. Since $\kappa = \hbar\omega/E_F$, we are interested only in the region $\kappa > 0$. The functions $g_1(\kappa)$ and $g_3(\kappa)$ in the interval $0 \leqslant \kappa \leqslant 1$ are shown in Fig. 7.

*Formerly at the Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

[1] W. U. Huynh, J. J. Dittmer, and A. P. Alivisatos, Science **295**, 2425 (2002).

[2] Y. Sun and Y. Xia, Anal. Chem. **74**, 5297 (2002).

[3] K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz, J. Phys. Chem. B **107**, 668 (2003).

[4] B. Nikoobakht, J. Wang, and M. A. El-Sayed, Chem. Phys. Lett. **366**, 17 (2002).

[5] E. Hao and G. C. Schatz, J. Chem. Phys. **120**, 357 (2004).

[6] F. Hache, D. Ricard, and C. Flytzanis, J. Opt. Soc. Am. B **3**, 1647 (1986).

[7] S. G. Rautian, Sov. Phys. JETP **85**, 451 (1997).

[8] V. P. Drachev, A. K. Buin, H. Nakotte, and V. M. Shalaev, Nano Lett. **4**, 1535 (2004).

[9] S. G. Rautian and A. M. Shalagin, *Kinetic Problems of Non-linear Spectroscopy* (North-Holland, Amsterdam, 1991).

[10] R. W. Boyd, *Nonlinear Optics* (Academic Press, Boston, 1992).

[11] R. H. M. Groeneveld, R. Sprik, and A. Lagendijk, Phys. Rev. B **51**, 11433 (1995).

[12] N. Del Fatti, R. Bouffanais, F. Vallee, and C. Flytzanis, Phys. Rev. Lett. **81**, 922 (1998).

[13] J. Lehmann, M. Merschdorf, W. Pfeiffer, A. Thon, S. Voll, and G. Gerber, J. Chem. Phys. **112**, 5428 (2000).

[14] V. P. Drachev, E. N. Khaliullin, W. Kim, F. Alzoubi, S. G. Rautian, V. P. Safonov, R. L. Armstrong, and V. M. Shalaev, Phys. Rev. B **69**, 035318 (2004).

[15] G. Y. Panasyuk, J. C. Schotland, and V. A. Markel, Phys. Rev. Lett. **100**, 047402 (2008).

[16]G. Y. Panasyuk, J. C. Schotland, and V. A. Markel, Phys. Rev. B **84**, 155460 (2011).

[17]W. Ekardt, Phys. Rev. Lett. **52**, 1925 (1984).

[18]J. Lerme, B. Palpant, E. Cottancin, M. Pellarin, B. Prevel, J. L. Vialle, and M. Broyer, Phys. Rev. B **60**, 16151 (1999).

[19]I. Vasiliev, S. Ogut, and J. R. Chelikowsky, Phys. Rev. B **65**, 115416 (2002).

[20]G. Zhu, M. Mayy, M. Bahoura, B. A. Ritzo, H. V. Gavrilenko, V. I. Gavrilenko, and M. A. Noginov, Opt. Express **16**, 15576 (2008).

[21]A. V. Gavrilenko, C. S. McKinney, and V. I. Gavrilenko, Phys. Rev. B **82**, 155426 (2010).