
This version is available at https://strathprints.strath.ac.uk/57291/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
Coherent control of plasmons in nanoparticles with nonlocal response

D. McArthur, B. Hourahine, F. Papoff

Department of Physics, SUPA, University of Strathclyde, 107 Rottenrow, Glasgow G4 0NG, UK

ARTICLE INFO

Article history:
Received 7 May 2016
Received in revised form 3 July 2016
Accepted 11 July 2016

Keywords:
Plasmonics
Nanoparticles
Nonlocality
Optical routing

ABSTRACT

We discuss a scheme for the coherent control of light and plasmons in nanoparticles that have nonlocal dielectric permittivity and contain nonlinear impurities or color centers. We consider particles which have a response to light that is strongly influenced by plasmons over a broad range of frequencies. Our coherent control method enables the reduction of absorption and/or suppression of scattering.

© 2016 The Authors. Published by Elsevier B.V. All rights reserved.

1. Introduction

Recent progress in nanophotonics and plasmonics has led to many important applications, including photovoltaics [1–3], sensors [4] and medicine [5]. Coherent control is a very topical subject in this area of research as it allows one to enhance the interaction of light with matter at the nanoscale: several groups have investigated nonlinear [6] and linear control based on pulse shaping [7–11], a combination of adaptive feedbacks and learning algorithms [12], as well as optimization of coupling through coherent absorption [13] and time reversal [14]. Coherent control of second-harmonic generation has been studied in nanowires [15,16] and nanospheres [17,18] while, in quantum optics, interference between fields was proposed as a way to suppress losses in a beam splitter [19] and has been recently applied to show control of light with light in linear plasmonic metamaterials [20]. These control methods have been applied only to systems with local responses. Particles with spatially nonlocal response behave very differently from particles with local response as they support irrotational charge density waves, such as plasmons, that do not radiate and can reach the central region of the particle over a large range of frequencies; on the contrary, particles with local responses support longitudinal modes only when the real part of the electric permittivity $\epsilon$ is null [21]. As a result, particles with nonlocal responses also exhibit a shift of the main resonance with respect to particles with local response for the same geometry and, in some metals, also have extra resonances at short wavelengths [22–27]. From the point of view of control, the main difference between media with local and nonlocal response is that in media with nonlocal response we can use light to control not only internal and scattered light, but also currents. In previous papers we have developed a coherent control theory for metallic nanospheres with diameters of at least 50 nm, for which nonlocal effects may be important only in a very thin layer at the boundary of the particles [17,18] where nonlinear processes take place. In this paper we investigate smaller nanoparticles in which the nonlinearity is due to an impurity, or color centers, inside the particle and for which nonlocal effects are important not only at the surface. We focus here on nanospheres because in this case the theory is fully analytical, but the approach we develop is based on the interference of fields at the surface of the particle and can be applied whenever longitudinal and transverse waves are both allowed, independently of the shape of the particle or the origin of the longitudinal waves. In particular, systems such as core-shell spherical particles, with diameters of 50 – 100 nm and an external layer with nonlocal response of a few nanometers, have similar properties to the spheres we consider here and interact more strongly with light. Consequently, these types of systems would be better from the point of view of applications. In this case, the control can be modeled similarly to the control method employed here by using the Mie theory for layered spheres [28,29]. However, depending on the materials used, there could be electron spill-out between the inner core and outer layer which would need to be included in the modeling of nonlocality [30].

We develop coherent control techniques which are extremely sensitive to phase variations and produce a reduction of the
absorption and variations of the scattered energy or of the amplitudes of the plasmons over several orders of magnitude. These unusual features enable applications such as detection of deeply sub wavelength changes in the position of the particle, reduction of dissipation, suppression of radiative losses, sensing of variations in the electric permittivity $\epsilon$ and magnetic permeability $\mu$ and optical routing.

2. Including nonlocality in Maxwell's equations

When one of the characteristic dimensions of the particle/structure is of the order of the electron free path, the free current is governed by a nonlocal equation that admits longitudinal waves. In the hydrodynamical model [22–27], the nonlocal response is modeled semi-classically by considering the free charges in the metal as a fluid governed by the linearized Navier–Stokes equation and with a pressure term that has a quantum origin and is proportional to the Fermi velocity. The interaction of the particle with light is then given by Maxwell’s equations combined with the linearized Navier–Stokes equations [31],

\[
\nabla \times \mathbf{E} = -\mu_0 \mathbf{J} + \varepsilon_0 \varepsilon_r \frac{\partial \mathbf{E}}{\partial t},
\]

\[
\nabla \times \mathbf{H} = \mathbf{J},
\]

\[
\left(\alpha_0 + \beta^2 \nabla \cdot \right) \mathbf{P} = \epsilon_0 \alpha_0^2 \mathbf{E},
\]

where $\mathbf{E}$, $\mathbf{H}$ are the electric and magnetic fields, $\mathbf{P}$ is the polarization due to the free current density $\mathbf{J}$, $\alpha_0$ is the electric permeability due to the bound charges, $\gamma_f$ is the damping factor due to collisions of the free charges, $\omega_F$ is the plasma frequency of the material and $\beta^2 = (3/5)\omega_F^2$, with $v_F$ the Fermi velocity. The nonlocal term $\nabla \times \mathbf{P}$ in (3) affects the interaction of the particle with light in two ways. First, the electric field contains both a transversal part $\mathbf{E}_t(\nabla \cdot \mathbf{E}_t = 0)$ and a longitudinal part $\mathbf{E}_l(\nabla \times \mathbf{E}_l = 0)$, each with its own dispersion relation. The longitudinal waves are expanded in terms of the longitudinal solutions of the Helmholtz equation and are associated with charge density waves, such as plasmons, but not to radiation as $\mathbf{E}_l$ is decoupled from time-dependent magnetic fields. Secondly, (3) also modifies the interaction with light through an additional boundary condition that is necessary to determine $\mathbf{P}$. This boundary condition is considered together with the usual continuity of the tangent components of $\mathbf{E}$ and $\mathbf{H}$ [32]. In media that do not support a surface density of free charges, the component of the free current density normal to the boundary of the particle is continuous [26,27]. At a dielectric-metal interface, this condition implies that the normal component of $\mathbf{P}$ in the metal has to vanish at the boundary, as dielectrics do not support free currents. Using the integral version of the divergence and an infinitesimal pillbox on the right-hand side of (2) shows that the normal component of $\epsilon_0 \mathbf{E} + \mathbf{P}$ is also continuous at the boundary. Therefore the continuity of the normal component of $\mathbf{P}$ is equivalent to the continuity of the normal component of $\epsilon_0 \mathbf{E}$, which provides the additional boundary condition,

\[
\hat{n} \epsilon_s \mathbf{E} = \hat{n} \epsilon_b (\mathbf{E} + \mathbf{E}^s),
\]

where $\epsilon_s^b$, $\epsilon_b^s$ are the permittivity due to bound charges of the external and internal media respectively, and $\mathbf{E}^s$, $\mathbf{E}^b$ are the internal, scattering and incident fields. From (2) we have $\nabla \cdot \mathbf{P} = -\varepsilon_0 \mathbf{E}$, which can be used to find the dispersion relation of the longitudinal waves and to recast the Maxwell equations in terms of the longitudinal and transverse electric fields; we then use the boundary conditions to determine the amplitudes of the longitudinal and transverse waves. The additional boundary condition in (4) and the presence of $\mathbf{E}_b$ lead to modified Mie coefficients [33] for the sphere.

3. Mode structure

At the heart of our theory is the Stratton-Chu representation theorem that allows one to express any internal and scattered fields of any smooth (possibly inhomogeneous) particle in terms of integral operators acting on the electromagnetic fields at the surface of the particles [34]. In practice this means that the response of a particle to light generated by impressed driving sources (which are either internal or external) can be determined by expanding the internal and the scattered fields in terms solutions of Maxwell’s equations for the internal and the external media that can approximate any field incident to the surface of the particle from the inside or the outside with arbitrary precision [35]. By defining surface fields with the electric and magnetic component parallel to the surface of the particle as $\mathbf{f} = \left[\mathbf{f} \times \hat{n}, -\mathbf{f} \times (\hat{n} \times \mathbf{H})\right]^T$, where $\mathbf{T}$ means transpose, $\hat{n}$ is the unit normal vector to the surface, and the scalar product of two surface fields is the overlap integral $\mathbf{f}_1 \cdot \mathbf{f}_2 = \int \mathbf{f}_1 \cdot \mathbf{f}_2 \, ds$, where the index $i$ labels the components of an arbitrary system of coordinates, the coefficients of the Mie modes of a sphere with local response are determined by projecting the incident fields on these modes using analytical formulae based on the scalar product defined on the surface [36]. These formulae apply also to particles whose modes can be found only numerically [35] and are very useful to determine the phase and amplitude of coherent light sources in order to modify the linear [36] and nonlinear [17,18] response of nanoparticles.

For particles with nonlocal response, one has also to include a complete set of longitudinal modes of the electric field corresponding to the plasmons; the coefficients for transverse and longitudinal modes can be calculated by fulfilling the continuity of the transverse component of the electric and magnetic fields as well as the boundary condition in (4) on the normal component of the electric field. To take into account (4), we need to also include the normal part of the electric field in the definition of the surface fields so that they now have five components,

\[
\mathbf{f} = \left[\hat{n} \epsilon_b \mathbf{E}^s, \mathbf{f}_1 \times (\hat{n} \times \mathbf{E}), -\mathbf{f}_2 \times (\hat{n} \times \mathbf{H})\right]^T.
\]

where $\epsilon_b^s$ is used for scattered and external fields, and $\epsilon_b^s$ is used for the internal field.\(^2\) The scalar product between surface fields is modified accordingly, and is given by the sum of the overlap integrals of these five components. As a consequence of the spherical symmetry, only modes with the same value of $l$ (orbital angular momentum) and $m$ (angular momentum along z) can couple. For each value of $l$ and $m$, using the angular dependence, one can group the modes into two sets: the set of transverse electric modes—as in a sphere with local response—and a set of two internal modes (the transverse magnetic and the longitudinal mode) and the transverse magnetic scattering mode. In a sphere the analysis can be limited to sets of modes with the same $l$ and $m$ because two modes with different $l$ or $m$ are orthogonal. For non

---

\(^1\) Note that in [22] the authors claim that it is not physically possible to distinguish between free and bound charges and, therefore, that it should be the normal component of the total current which is continuous at the surface. Using a similar approach as the one above, this assumption is equivalent to the continuity of the normal component of the electric field [23]. There is very limited difference in the numerical results given by these two additional boundary conditions and no experimental evidence to support one over the other.

\(^2\) Using the additional boundary condition described in [22,23], the first component of (5) should be replaced by $\hat{n} \mathbf{E}$.
spherical particles Mie modes can still be used to expand internal and scattered fields, but modes with different \( l \) and \( m \) can be coupled to one another. However, the Mie modes can be recombined to form principal modes that, for any smooth mode, have the same properties of Mie modes for a sphere [35]. Starting from sets of internal and scattering modes complete on the surface of the particle, one can first orthogonalize the modes then rearrange them so that the internal and scattering modes are coupled pairwise on the surface, as the Mie modes. Furthermore, for particles of the dimensions considered in this paper, only modes with \( l = 1 \) need to be considered.

In order to control the response of the nanosphere, we have to change the relative amplitudes of the modes that most affect the overall response of the particle, which, for nanoparticles of the size considered here, means transverse magnetic modes with \( l = 1 \). To describe effects that are most easily observed experimentally, we concentrate here on the control of two out of the three dominant modes using two coherent light sources with surface fields labeled \( f_1, f_2 \) that are functions of the position on the surface. The real amplitudes and phases of these fields are encoded in their complex amplitudes \( a_1, a_2 \). As the third mode with the same \( l \) and \( m \) is not controlled, the choice of which pair of modes to control depends on whether we want to affect the light or the current. The amplitude of the scattering mode, \( s \), and of an internal mode \( l \), are given by:

\[
\begin{bmatrix}
a_1 \\
-a_1 \\
-f_1 \\
-f_2 \\
-f_1 \\
-f_2 \\
-f_1 \\
-f_2 \\
\end{bmatrix} =
\begin{bmatrix}
f_1 & f_2 \\
-s f_1 & -s f_2 \\
-s f_1 & -s f_2 \\
-s f_1 & -s f_2 \\
-s f_1 & -s f_2 \\
-s f_1 & -s f_2 \\
-s f_1 & -s f_2 \\
-s f_1 & -s f_2 \\
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
a_1 \\
a_2 \\
-a_1 \\
-a_2 \\
-a_1 \\
-a_2 \\
\end{bmatrix}
\]

(6)

where \( i(s) \) is the biorthogonal mode that is orthogonal to all modes other than \( i(s) \), see the Appendix for explicit formulae. We can find any pair of mode amplitudes, \( a_1, a_2 \), by changing the amplitudes of the incident fields as long as the determinant of the matrix in (6) is invertible, i.e. as long as the condition \( (f_1 f_2 - f_2 f_1) \neq 0 \) is fulfilled. Similarly, one could use the second internal mode, \( f_2 \), instead of the scattering mode, replacing \( s \) with \( f_2 \) with \(-a_2\) with \( a_2 \), where the minus sign originates from the boundary condition and is due to the fact that internal and scattering modes are defined on opposite sides of the surface. For a sphere [36,17], the matrix in (6) is invertible only if the functional dependence of \( f_1 \) on the position on the surface is different from that of \( f_2 \); therefore we have to use two sources of light whose fields are expanded in terms of different types of spherical vector solutions of the Maxwell equations. Physically, this means that we can use a laser and either a converging or diverging spherical wave, but not two lasers or two converging or diverging waves. Note that using a standard external source and both a converging and a diverging wave with a given \( l \) and \( m \), it is possible to control the amplitude of three modes with the same \( l \) and \( m \). For instance, one could suppress both internal modes, eliminating absorption for that \( l \) and \( m \).

Converging spherical waves are solutions of the Maxwell’s equations that can in principle be realized using a phased array of radiation sources and sinks (which are the time reversal of sources and absorb incoming waves [14]): this is practically very challenging at the nanoscale, so we will concentrate in the following on diverging, or outgoing, waves. Dipole radiation can be excited by an impurity or color center of size \(< 1\) nm (an order of magnitude smaller than the size of the sphere so that it does not affect significantly the bulk properties of the sphere) located anywhere inside the sphere through any nonlinear up- or down-conversion process of a field at a different frequency [37]. The outgoing dipole radiation has total angular momentum \( l = 1 \) with respect to the center of the impurity, and can be decomposed in terms of spherical waves with \( l = 0, 1, 2 \) and \( m = 0, \pm 1 \) [38], where the value of \( m \) to consider depends on the specific nonlinear process used.

Using the translation theorem for spherical waves [32], we see that these waves radiated from the dipole can be recast as a superposition of spherical waves emerging from the center of the sphere, but with angular momentum with respect to the center of the sphere that can be larger than 2. However, for spheres of the dimensions considered here, the response to waves with \( l > 1 \) is extremely weak and can be neglected. Here we use modes with indexes \( l = 1, m = 0 \), which corresponds to assuming that one of the two waves is generated through a second harmonic process by an impurity at the center of the sphere [39], but the principles of control are independent of \( m \). A diagram of the proposed scheme is shown in Fig. 1. Modes with \( l = 1 \) would still dominate the response of the particle with an off-center impurity, but one may have to consider also modes with \( m = ± 1, 1 \), including as many external control beams as \( m \)-values one wishes to control. From a theoretical point of view, the matrix in (6) will include two extra modes for each additional \( m \)-value [36]. In principle an undefined position would be detectable by measuring the relative amplitudes of the \( m = −1, 0, 1 \) channels, which would have to be controlled separately. An estimate of the relative amplitudes with respect to the centered impurity is provided by the translation theorem for vector spherical wave functions and is at most of the order of the ratio between the distance of the impurity from the center of the

Fig. 1. Diagram of proposed control scheme. The response of a spherical metallic nanoparticle is controlled by varying the amplitude and relative phase between an external source which is regular over the surface of the particle, \( f_1 \), and a spherical wave, \( f_2 \), originating from a very small impurity at the center of the particle, pumped at a frequency sufficiently separated from that of the control source.
particle and the wavelength, so extremely small for the particles considered here.

4. Results and discussion

We consider in the following aluminum and gold spheres primarily of 5 nm radius producing resonances in the ultraviolet region, as these particle should be easily available in experiments. These particles exhibit a blue-shift of the main resonance with respect to particles with local response for the same geometry but do not show extra resonances at short wavelengths. At the end of the section we include results for smaller particles to demonstrate that the control methods work in principle even for a smaller particle, $r=1.5$ nm, which does exhibit these additional resonances, as well as for an intermediate particle, $r=3$ nm. Particles of this scale have been previously studied theoretically [27] as they are above the length scale where core plasmons are significant. However, it is difficult to realize smooth spheres with these dimensions which contain very small impurities or color centers; for this reason we focus more on the 5 nm radius case.

The dielectric functions $\epsilon(\omega)$ are given by a Lorentz–Drude model [40,41], but our calculations also produce similar results with different dispersion relations from the literature for these materials. As the dimensions of the particles considered are smaller than, or comparable to, the mean free path of the conduction electrons in the materials we include an additional size dependent damping term in the dielectric function to account for the increase in surface collisions [42]. To further demonstrate the general applicability of the proposed coherent control method, we include results obtained by applying the recently developed generalized nonlocal optical response theory (GNOR) [43], which explains the damping and broadening of plasmon resonances via electron diffusion. For metals, the diffusion constant is $D = v_F^2 \tau_s$, where the relaxation time associated to scattering processes can be approximated for spherical particles as $\tau_s \sim 4/|\omega|\tau$. If this constant is instead treated as a free parameter, we observe that the amplitudes of the modes vary exponentially with $\tau_s$ while their phases barely shift and the suppression achieved through the control is unaffected.

In Fig. 2(a) we compare the extinction cross sections for an aluminum particle of this size, both with and without nonlocal effects, and show in Fig. 2(b) that the particle response is dominated by absorption. In view of experimental realization, it is important to consider that the effects we show in the following can be qualitatively observed even when knowing the effective parameters of the particles with an uncertainty of $10 – 20\%$, as is often the case [44]. For instance, as long as $|a_0 f_1|$ and $|a_0 f_2|$ are of the same order, we can control the total scattered energy and introduce sharp spectral features in the scattering spectrum [17] simply by scanning the relative phase between $f_1$ and $f_2$. In this example, $f_1$ is a linearly polarized plane wave incident at $\pi/2$ with respect to the $z$ axis and $f_2$ is an outgoing spherical wave with $l = 1$, $m = 0$; which represents the dominant term resulting from the selection rules for second harmonic generation by an impurity, excited by a pump propagating along the $z$ axis, where the amplitude $a_2$ is normalized to unity. The amplitude, $a_1$, is found by expanding the external control wave in terms of spherical harmonics [32]. In practical applications the efficiency of the nonlinear process generating $f_2$ would always be less than unity and consequently the amplitude of the control wave would always be significantly smaller than that of the pump. We note that the size of the particles considered is smaller than the skin depth, so that a large fraction of the pump can reach the impurity. For sake of simplicity, we assume the impurity is at the center of the sphere, but we point out that moving the impurity off center would produce terms with different which do not change significantly the response of spheres of the dimensions considered here.

Using Mie modes we can control the relative amplitudes of the internal modes and, as a consequence, the distribution of currents and the absorption of energy due to the two internal modes in a given $l$, $m$ channel. In Fig. 3 we switch off the longitudinal mode with $l = 1$, $m = 0$, at a frequency far from the resonance in Fig. 2, and show the effect of the control upon the power absorbed by the particle as a function of the relative phase between $f_1$ and $f_2$. This effect could be best observed experimentally by looking at the extinction cross section, as demonstrated in Fig. 3(c). As the local model has no analog to the longitudinal mode, we apply the parameters used to control this mode for the nonlocal system in order to investigate the effect of the presence of nonlocality upon the extinction. Far from the main resonance, Fig. 2(a) shows us that we should expect the two models to produce similar results. This is indeed what we observe, however there is a slightly greater reduction in the extinction calculated with the local model,
indicating that it is the effect of the control upon the TM modes present in both models, rather than the longitudinal mode, which dominates the observed response for this particle. The reduction in extinction happens at a phase different from the one where optimal control of the longitudinal mode is achieved because the extinction depends on all of the modes. Similarly, by minimizing the amplitude of the scattering mode we can introduce extremely sharp features in the scattering cross section in the direction orthogonal to both the pump and control.

In Fig. 4, we control the scattering mode for an aluminum sphere at the peak of the main resonance in Fig. 2(a) where we observe that the amplitudes of all three modes are reduced by several orders of magnitude, as their spatial coherence on the surface is very high near the resonance, shown in Fig. 4(a). By integrating the scattered energy over a cone in the far field, corresponding to the signal that would be detected by a wide angle detector with a numerical aperture $\text{NA} \approx 0.8$, we observe a reduction of over six orders of magnitude in the light scattered, shown in Fig. 4(b).

To further investigate the differences between a local and nonlocal response to control, in Fig. 6 we consider an aluminum...
particle of radius 3 nm for which the effects of nonlocality should be more pronounced. We employ the same method of control, for the same sets of modes as in the previous figures. Fig. 6(b) shows that the scattering for the local mode is not minimized as sharply as it is for the nonlocal mode for this particle. Although there are many orders of magnitude of difference between the minima of the two curves, it may be easier to detect experimentally the difference in the extinction signals which are of an order of magnitude at their minima, figure 6(c), as the extinction signals are greater than those of the scattering. Furthermore, similar results could be expected to be observed if the amplitude of the control wave was reduced by approx. -102%. We note also that strong suppression of the modes can still be achieved for variations of the frequency of up to approx. 10%.

Finally, in Fig. 7(a) we compare the extinction power for an aluminum sphere of radius 1.5 nm exhibiting local and nonlocal response to excitation by plane wave, where the dielectric function does not include a damping term in the dielectric function to account for surface collisions [42]. This figure is analogous to Fig. 7(b1) in [27], where they show that additional resonances appear in the nonlocal spectrum above the plasma frequency, for particles of these dimensions. Again, we show in Fig. 7(b)-(c) that we are able to control the response of these particles and also that for particles this small there is a clear difference between the local and nonlocal response. A particle with this type of control, if placed at the juncture between two chains of nanoparticles, could allow or interdit the passage of light at the frequency where the control is being operated, as even the surface field outside the particle can be made negligible. In this example, the external field would be the field emitted from the end of one chain near to the sphere. For this specific application, however, one would have to consider carefully the distance between particles: for sub-nanometer distances modifications to the nonlocal theory [30] and quantum effects such as spill-out charges [45] need to be included in the model. Furthermore, we have shown that there are differences in the extinction cross section due to the presence or absence of the longitudinal mode for all the particle sizes considered. For scattering, the difference is only detectable for the smaller particles, r=3 nm and r=1.5 nm. For these particles, variations on the extinction cross section are much easier to detect than variations on
the scattering cross section, so in principle one could use control techniques to identify the effect of nonlocality.

5. Conclusions

In conclusion, we have presented a scheme for the coherent control of scattered light and current in nanospheres with nonlocal response, identifying extreme sensitivity to variations of the relative phase of two coherent incident waves that can lead to novel applications. These include the suppression of radiative losses, sensing of variations in the electromagnetic properties of the host medium and optical routing. We note that when using light to control currents, the variation in absorption can be used to reduce dissipation, enhance the detection of nonlinear processes and lower amplification thresholds. The theory we have presented can be easily generalized to particles with other shapes: in most cases the Mie approach based on separation of variables does not apply, but the principal mode theory can be performed and it leads to the same type of equations used here to design control schemes.

Acknowledgments

D.M. is supported by the Engineering and Physical Sciences Research Council (EPSRC) (EP/K503174/1).

Appendix

We recall that for monochromatic fields with frequency $\omega$, the longitudinal waves are associated to charge density waves, or plasmons, but not to radiation, as $E_L$ is decoupled from time-dependent magnetic fields. These waves have a wave vector [26]

$$k_L = \left[ (\epsilon_{\infty}\omega + i\Gamma) - \omega_p^2 \right]^{1/2},$$

(A.1)

and are expanded in terms of the longitudinal solutions of the Helmholtz equation [23], where $\epsilon_{\infty}$ is the relative permittivity due to bound charges and $\Gamma$ is the damping constant of the material.

The electric and magnetic components of the Mie modes are the transverse solutions of the Helmholtz equation,

$$M_{lm}(r, \theta, \phi) = z_l(kr)\bar{m}_{lm}(\theta, \phi)$$

(A.2)

and the longitudinal solutions,

$$L_{lm}(r, \theta, \phi) = k^{-1}\bar{v}_l(kr)\bar{y}_{lm}(\theta, \phi) \sqrt{[l(l+1)]},$$

(A.4)

where $k$ is the wavenumber of the transverse waves, $\bar{m}_{lm}$ and $\bar{n}_{lm}$ are the vector spherical harmonics, $\bar{y}_{lm}$ are the scalar spherical harmonics [46], $z_l$ can be either the spherical Bessel function (internal field) or the Hankel function of the first kind (scattering field) of order $l$ and $j_l$ is the spherical Bessel function of order $l$.

For spheres with nonlocal dispersion, the normalized electric multipole Mie modes are

$$i_{lm}^e = \frac{\epsilon_{\infty}(L_{lm} \bar{m}_{lm})}{L_L^e},$$

(A.5)

$$i_{lm}^m = \frac{\left( N_{lm}^l - iC^l M_{lm}^l \right)}{R_L^m},$$

(A.6)

$$s_{lm} = \frac{\left( N_{lm}^s - iC^s M_{lm}^s \right)}{R_L^s},$$

(A.7)

where the bracketed notation defines the electric and magnetic components of the mode successively and $C^l = \sqrt{\epsilon}$, where $\epsilon$ is the permittivity of the transverse waves that includes also the free carriers contribution, $C^s = \sqrt{\epsilon}$ and the normalization factors and phase used above are [35],

$$R_L^e = \left| N_{lm}^e - iC^e M_{lm}^e \right|,$$

(A.8)

$$R_L^m = \left| N_{lm}^m - iC^m M_{lm}^m \right|,$$

(A.9)

$$\psi = -\text{Im} \left[ \log \left( \left( N_{lm}^s - iC^s M_{lm}^s \right) \left( L_{lm} \bar{m}_{lm} \right) \right) \right].$$

(A.10)

The biorthogonal modes are given by the formula

$$u_{lj} = u_l C_{lj}^{-1},$$

(A.11)

where $u_1 = s_{lm}$, $u_2 = i_{lm}$, $u_3 = i_{lm}$, $G^{-1}$ is the inverse of the (Gram) matrix with elements $C_{lj} = \left( u_l, u_j \right)$ and we sum over repeated indexes, where $k_{lm}^l$ is the longitudinal mode spatially correlated to $s_{lm}$ and $i_{lm}$. The magnetic multipole Mie modes can be found by swapping $\bar{M}$ and $\bar{N}$ in ((A.6)–(A.9)) and are not affected by the presence of nonlocality. (there are only longitudinal electric multipoles.)