A general framework of canonical quasinormal mode analysis for extreme nano-optics

Qiang Zhou, Pu Zhang, and Xue-Wen Chen

School of Physics and Wuhan National Laboratory for Optoelectronics,
Huazhong University of Science and Technology,
Luoyu Road 1037, Wuhan 430074, People’s Republic of China

Optical phenomena associated with extremely localized field should be understood with considerations of nonlocal and quantum effects, which pose a hurdle to conceptualize the physics with a picture of eigenmodes. Here we first propose a generalized Lorentz model to describe general nonlocal media under linear mean-field approximation and formulate source-free Maxwell’s equations as a linear eigenvalue problem to define the quasinormal modes. Then we introduce an orthonormalization scheme for the modes and establish a canonical quasinormal mode framework for general nonlocal media. Explicit formalisms for metals described by quantum hydrodynamic model and polar dielectrics with nonlocal response are exemplified. The framework enables for the first time direct modal analysis of mode transition in the quantum tunneling regime and provides physical insights beyond usual far-field spectroscopic analysis. Applied to nonlocal polar dielectrics, the framework also unveils the important roles of longitudinal phonon polaritons in optical response.

Introduction.— The research field of nano-optics has flourished along the development of concepts and techniques to shrink light to scales well below diffraction limit [1–5]. The subwavelength confinement primarily results from plasmonic oscillations in metal nanostructures [1, 2] or phononic oscillations in polar dielectric nanomaterials [6, 7]. Rapid advances in nanotechnologies have allowed sculpting structural morphology at nanometer and even subnanometer scales [8–12], pushing field concentration towards extreme [13]. Recent experimental demonstrations suggest optical field could even be confined to cubic-nanometer volumes [14–17]. Along this line of research, the field of nano-optics enters a new regime, where classical local treatment of nanomaterials becomes invalid. As an example, metallic gaps of a few nanometers or smaller exhibit significant nonclassical effects from electron nonlocality [9, 18], spill-over at metal surfaces [19, 20], Landau damping [21, 22] and quantum tunneling [23, 24]. To describe these effects for plasmonic nanostructures (≈100 nm) with tiny subfeatures by affordable computational resources, researchers have developed various effective models at different levels of approximation [19–21, 25–29]. For polar dielectric nanostructures, nonlocal responses also have been treated to properly describe the nanoscale physics [30–32].

While the system optical response could be computed by discretizing Maxwell’s equations in media with a suitable model, the interpretation of the physics and characterization of light-matter interaction properties are not straightforward. If the governing eigenmodes of the system were known, then system response could often be conceptualized. Thus the ability to obtain and orthonormalize the eigenmodes is essential to promote the development of extreme nano-optics. Both the openness and dissipative nature of the system call for a quasinormal mode (QNM) analysis, which has been established for classical local media [33–41]. There have been few attempts to extend the QNM analysis to include electron nonlocality under a hydrodynamic treatment [42, 43] and recently to incorporate a quantum hydrodynamic model (QHDM) to study extremely localized modes [44]. However, none of the extensions have attempted to construct an orthonormalization scheme for the QNMs, which is challenging since the aforementioned quantum effects have to be properly treated in the relation. Orthonormalization is crucial as it enables constructing system response through modal contributions for an arbitrary excitation and a direct evaluation of mode volume of each mode to facilitate quantum-optical studies [39–41, 45, 46].

Here we propose a generalized Lorentz model to describe general nonlocal media and formulate source-free Maxwell’s equations in the media as a linear eigenvalue problem (LEVP) to canonically define the QNMs. Then we introduce a general scheme to orthonormalize the eigenmodes and consequently establish a framework of canonical QNM analysis. Taking QHDM for metal and nonlocal polar medium for dielectric as examples, we present explicit QNM formalisms in corresponding media. We reveal the mode evolution process of a plasmonic dimer in the quantum tunneling regime where classical local treatment fails completely [21, 47] and show a mode transition occurs at a smaller gap than expected from extinction spectra. Moreover, we employ QNM analysis to interpret the optical responses of a silicon carbide (SiC) nanosphere influenced by longitudinal phonon modes. The examples cover a broad range of situations in extreme nano-optics.

Generalized Lorentz model and LEVP formulation.— Classical local optical responses of materials can often be described by the Drude model or Lorentz oscillator model of charged particles with single or multiple resonances [48]. Nonlocal responses of materials conspicuously manifest via longitudinal density waves of the constituent charged particles with a characteristic length much shorter than vacuum wavelength [49]. Inspired by
the nature of nonlocality as interaction induced by charge density gradient, we propose a generalized Lorentz model (GLM) to incorporate various nonlocal responses under weak excitation

$$\rho \ddot{X} + \hat{\Gamma}_x \dot{X} + \hat{\Theta}_x X = (q_p/m_p)\rho E, \quad (1)$$

where $X$ is the relative displacement and the restoring and damping force constants have become operators containing spatial gradients. $\rho$ is the stationary charge number density normalized by the average number density $n_u$ and $q_p/m_p$ is the charge-mass ratio. A nonuniform $\rho$ covers situations of electron spill-over [27], dynamic carrier control [50] and gradient-alloyed semiconductors [51].

The material response couples with Maxwell’s equations through electric polarization $P = n_u \rho q_p X$. Assuming a time convention of $\exp(-i\omega t)$ with $\omega$ being a complex frequency, the equation for $P$ in frequency domain reads

$$i\epsilon_0 \omega_p^2 E - i\rho \hat{\Theta} P - i\rho \hat{\Gamma} J = \tilde{\omega}J, \quad (2)$$

with the polarization current $J = -i\omega P$, modified restoring and damping force operators of $\hat{\Theta} = -\partial^{-1} \hat{\Theta}_x \rho^{-1}$ and $\hat{\Gamma} = -\partial^{-1} \hat{\Gamma}_x \rho^{-1}$. $\omega_p = [n_u q_p^2/(m_e \epsilon_0)]^{1/2}$ resembles the plasma frequency. Equations (1-2) establish a Lorentz-type operator description of spatially dispersive media at optical frequencies, in parallel to the usual electric susceptibility descriptions of $\chi(\omega, r, r')$ and $\chi(\omega, k)$ in real-space and wavevector-space forms respectively [52]. The GLM is basic since it is only based on the generic nature of material nonlocality and Newton’s second law. As listed in Table I and detailed in SM [53], our GLM accommodates at least the following nonlocal models, i.e., hard-wall hydrodynamic model [54], generalized nonlocal optical response [55], QHDM [19–21] and nonlinear polar dielectrics [32]. The local response approximation is a special case.

With the materials described by the GLM, the source-free Maxwell’s equations can be formulated as an LEVP

$$\mathcal{H} \Phi = \left[ \begin{array}{ccc} \frac{1}{\epsilon_0 \tau} & -\frac{i}{\epsilon_0 \tau} \nabla \times & 0 \\ \frac{i}{\mu_0 \tau} \nabla \times & \frac{1}{\mu_0 \tau} & 0 \\ 0 & 0 & \frac{i}{\epsilon_0 \tau} \end{array} \right] \Phi = \tilde{\omega} \Phi \quad (3)$$

Here $\Phi = [E, H, P, J]^T$ and $\tilde{\omega}$ are the eigenvector and eigenfrequency, respectively. $\epsilon_0 \tau$ is the non-resonant background permittivity. In open space, the radiation boundary condition ($E(r) \propto r^{-1} e^{i \omega r/c}$ as $r \rightarrow \infty$) should be imposed. Equation (3) is formally similar to the auxiliary-field formulations developed for normal modes [52] and extended to treat losses for QNM analysis [37].

The original contribution here lies in the promotion of constants to operators $\hat{\Theta}, \hat{\Gamma}$ and the use of a nonuniform $\rho$ to describe general nonlocal responses and to develop a corresponding canonical QNM theory.

| Model | $\rho$ | $\hat{\Theta}$ | $\hat{\Gamma}$ | B.C. |
|-------|-------|---------------|---------------|------|
| LRA   | $1$   | $\omega_i^2$  | $\gamma$      | $n \cdot P$ |
| HDM [54] | $1$   | $-\beta^2 \nabla (\nabla \cdot)$ | $\gamma$ | $n \cdot P$ |
| GNOR [55] | $1$   | $-(\beta^2 + \gamma D) \nabla (\nabla \cdot)$ | $\gamma - D \nabla (\nabla \cdot)$ | $n \cdot P$ |
| QHDM  | $\rho(r)$ | $-\hat{\Pi} \cdot \Sigma_d$ | $\gamma/\rho - \Sigma_1$ | $/Polar diel.$ | $1$ |
|       |       |               |               | $\gamma_d$ |

TABLE I. List of nonlocal models formulated in GLM [53].

**General framework of canonical QNM analysis.**— A canonical QNM theory should include a scheme to orthonormalize the modes. For normal modes of a Hermitian system, the normalization is carried out through an integration representing the field energy [56, 57]. For QNMs, the exponential divergence of the mode field in the far field causes a difficulty for normalization. With continuous efforts [33, 36, 37, 39, 58, 59], the orthonormalization scheme for QNMs in classical local media has been established by solving the divergence problem through a bilinear form and a complex coordinate mapping technique [36, 37]. However, the orthonormalization of QNM in nonlocal media remains an untouched challenging problem. By inspecting the orthonormalization formulas of Eq. (13) in Ref. [50] for normal modes and of Eq. (4) in Ref. [37] for QNMs, one sees each term in the expression for QNMs also has a close link to the mode energy. Energy-wise, for general nonlocal media, the difficulties are the proper treatments of various internal interaction energies of charged particles [19, 20, 32] and Landau damping as a type of interaction energy [21, 55]. Moreover, the nonuniform $\rho$ induces another barrier since this means a position-dependent plasma frequency and interaction energies. Our GLM formulation enables us to get around all these hurdles. Inspired by the close link to the field energy, we start with Poynting theorem [60] by evaluating the difference between the powers input by two current sources $I_v = \int V \cdot d r [i(J_{s,2} \cdot E_1 - E_2 \cdot iJ_{s,1})]$, where $E_i$ is the response field of source $J_{s,i}$ with $i = 1, 2$. This leads to the expression for the electromagnetic energy in general nonlocal media [53]. Next we switch to the bilinear version (remove complex conjugate operators), e.g. $I_v \rightarrow I_v = \int V \cdot d r [i(J_{s,2} \cdot E_1 - E_2 \cdot iJ_{s,1})]$. By replacing the current sources with the fields, the bilinear form of Poynting theorem can be arranged to [53]

$$\left(\omega_1 - \omega_2\right) \int_V \left(\frac{\epsilon_0 \epsilon_2 E_2 \cdot E_1 - \mu_0 H_2 \cdot H_1 + P_2 \cdot \hat{\Theta} P_1 - J_2 \cdot J_1}{\epsilon_0 \omega_2^2} - \frac{J_1 \cdot J_2}{\epsilon_0 \omega_2^2} \right) \left[\frac{\epsilon_0 \epsilon_2 E_2 \cdot E_1}{\epsilon_0 \omega^2} + i \omega \left(P_2 \cdot \hat{\Theta} P_1 - P_1 \cdot \hat{\Theta} P_2\right)\right]$$

$$- i \int_V \nabla \cdot (E_1 \times H_2 - E_2 \times H_1) = I_v. \quad (4)$$

Notice that the GLM formulation actually enables us to establish an unconjugated Lorentz reciprocity theorem
for general nonlocal media. We identify that \( \hat{\Theta} \) and \( \bar{\Gamma} \) should be real and symmetric under transposition from system energy considerations [53], such as the positive definiteness of eigen-energies and the exponential decay caused by coupling to a continuum. The symmetry condition also implicates certain requirements on the boundary conditions for \( \mathbf{P} \) across material interface [32, 53, 54]. Such symmetry condition is indeed satisfied for all the situations of interest [53]. Consequently, the second integral of Eq. (4) vanishes. The last integral of Eq. (4) is normally converted to a surface integral at infinity and effectively brought to zero by using the perfectly matched layers [36]. However, a nontrivial surface integral may arise when unconventional electromagnetic boundary conditions are applied, e.g., in Feibelman’s d-parameter model [26, 28, 29, 61]. Although it could be treated by following the general spirit of our orthonormalization procedure [62], for the sake of clarity, here we restrict to volumetric media responses and thus drop the term. Then the left hand side of Eq. (4) is left only with the first integral. Now considering the case that \( \mathbf{E}_1 \) and \( \mathbf{E}_2 \) are two sets of source-free \( (\mathbf{J}_{s_1} = 0 \quad \text{and} \quad \mathbf{J}_{s_2} = 0) \) eigenmode fields, the right hand side of Eq. (4) becomes zero [36]. Thus Eq. (4) directly leads to the orthonormal relation \( (\bar{\omega}_1 - \bar{\omega}_2)((\hat{\Phi}_2, \hat{\Phi}_1))_M = 0 \), where the bilinear form \( (\hat{\Psi}_2, \hat{\Psi}_1)_M \) is defined with

\[
\mathcal{M} = \text{diag} \left\{ \varepsilon_0 \varepsilon_{\infty}, -\mu_0, (\varepsilon_0 \omega_p^2)^{-1} \bar{\Theta}, (\varepsilon_0 \omega_p^2 \rho)^{-1} \right\}.
\]

The normalization factors of QNMs in general nonlocal media are immediately obtained as \( \mathcal{N}_m^2 = ((\hat{\Phi}_m, \hat{\Phi}_m))_M \). \( \mathcal{M} \) serves as a mapping operator such that the basis \( \{\hat{\Phi}_m\} \) and its dual basis \( \{\check{\Phi}_m\} \) form a biorthogonal system. The completeness of the system is discussed in SM [53]. The GLM formulation and orthonormal relation constitute the core of the general framework of canonical QNM analysis and empower an analytical description of optical responses with numerically calculated QNMs. Analytical formulas can be derived by applying the orthonormal relation in parallel to the classical local theory [37, 40, 53]. Linear responses can be expanded as \( \hat{\Psi} = \sum_m \alpha_m(\omega) \check{\Phi}_m \) with coefficients \( \alpha_m(\omega) = (\bar{\omega}_m - \omega)^{-1}((\hat{\Phi}_m, \mathcal{S})_M) \), where \( \mathcal{S} \) is the excitation source. The complex position-dependent mode volume reads \( V_m = 1/(2\varepsilon_0 \alpha_m^2(\mathbf{E}_m(\mathbf{r}_d) \cdot \mathbf{u}_d))^2 \), which is evaluated for a dipolar emitter at \( \mathbf{r}_d \) along a unit vector \( \mathbf{u}_d \) in the medium with a refractive index of \( n_d \).

**General framework applied to metals and dielectrics.**— In the following, we take QHDM as the most sophisticated nonlocal model for metal and nonlocal polar medium for dielectric as two archetypical examples to explicitly work out their QNM theories. When metallic nanostructures are excited by light in the linear nonlocal response regime, the electron gas experiences additionally a pressure force arising from its internal energy [20] and a viscoelastic force associated with Landau damping [21]. The two forces can be formulated as operators acting on the induced polarization \( \mathbf{P} \) and current \( \mathbf{J} \). Then the governing equation of QHDM can be written as [53]

\[
i \rho \varepsilon_0 \omega_p^2 \hat{\mathbf{E}} - i \rho (-\bar{\Pi} - \bar{\Sigma}_2) \hat{\mathbf{P}} - i \rho (\gamma/\rho - \bar{\Sigma}_1) \hat{\mathbf{J}} = \bar{\omega} \hat{\mathbf{J}}. \tag{6}
\]

\( \gamma \) is the phenomenological damping rate. Comparison of Eq. (6) with Eq. (2) immediately reveals \( \hat{\Theta} = -\bar{\Pi} - \bar{\Sigma}_2 \) and \( \bar{\Gamma} = \gamma/\rho - \bar{\Sigma}_1 \). The constituting operators are

\[
\bar{\Pi} = \nabla \bar{K}_1(\nabla \cdot \cdot ) - \nabla \nabla \cdot K_2 \nabla (\nabla \cdot \cdot ), \tag{7a}
\]

\[
(\bar{\Sigma}_1)_{kk} = \bar{D}_j \left[ \eta^k_j (\delta_{kk} + \delta_{kk}) + \xi^k_j \delta_{kk} \right] \bar{D}_j, \tag{7b}
\]

\[
(\bar{\Sigma}_2)_{kk} = \bar{D}_j \left[ \eta^k_j (\delta_{kk} + \delta_{kk}) + \xi^k_j \delta_{kk} \right] \bar{D}_j, \tag{7c}
\]

where \( \bar{D} = \nabla - \rho^{-1} \nabla \rho \), and \( \bar{D}_d = \nabla + \rho^{-1} \nabla \rho \). \( K_{1,2} \) and \( \eta^k_j, \mu^k_j, \xi^k_j, \zeta^k_j \) are functions of \( \rho \) [53]. The operators comply with the requirement of being real symmetric. The canonical QNM theory for QHDM naturally follows from the general framework. A direct evaluation of \( ((\hat{\Phi}_m, \hat{\Phi}_m))_M \) yields the normalization factor

\[
\mathcal{N}_m^2 = \int d\mathbf{r} 2\varepsilon_0 \varepsilon_{\infty} \mathbf{E}_m(\mathbf{E}_m - \int d\mathbf{r} [i\gamma \bar{\omega}_m/(\varepsilon_0 \omega_p^2 \rho)] \mathbf{P}_m(\mathbf{P}_m - \bar{\omega} \mathbf{P}_m) = (1/\varepsilon_0 \omega_p^2) \int d\mathbf{r} \mathbf{P}_m((2\bar{\Pi} + 2\bar{\Sigma}_2 - \varepsilon_0 \omega_p^2 \bar{\Sigma}_1)\mathbf{P}_m. \tag{8}
\]

In the expression, various nonlocal interactions are clearly manifested. The second integral accounts for the effect of nonuniform \( \rho \). The second line encapsulates various nonlocal responses, such as the contributions of electron pressure (\( \bar{\Pi} \)), elastic (\( \bar{\Sigma}_2 \)) and viscous (\( \bar{\Sigma}_1 \)) effects.

In polar dielectrics, nonlocality originates from ion interactions through longitudinal phonons. By inspecting the dynamic equation for lattice vibrations proposed in Ref. [32], it amounts to Eq. (2) with \( \rho = 1, \bar{\Gamma} = \gamma_d \) and

\[
\hat{\Theta} = \omega_T^2 + \beta_T^2 \nabla \nabla \cdot \beta_T^2 \nabla \times (\nabla \times). \tag{9}
\]

\( \omega_T \) and \( \beta_T \) (\( \beta_L \)) are the transverse optical phonon frequency and velocity of the transverse (longitudinal) phonons, respectively. The last two terms in Eq. (9) follow from the divergence of the stress tensor \( \mathbf{S} \), i.e., \( \nabla \cdot \mathbf{S} = (\hat{\Theta} - \omega_T^2) \mathbf{P} \). The nonlocal force operator is real symmetric with a required boundary condition \( \mathbf{n} \cdot \mathbf{S} \mathbf{P} = 0 \) [32]. Therefore our general framework directly leads to a QNM theory for nonlocal polar dielectrics. The detailed derivations and numerical implementation of the two exemplary theories are provided in SM [53].

**Mode transition in quantum tunneling regime.**— A metallic nanosphere dimer with an ultrathin gap is an excellent platform for studying intriguing quantum plasmonic phenomena such as quantum electron tunneling [23–25, 63–65]. Considerable experimental and theoretical endeavors focus on the plasmon mode transition behavior as the gap gradually closes, but have not reached a
complete consensus. A plausible reason is that the resonance information is indirectly retrieved from the far-field extinction or scattering spectra which strongly depend on the illumination and detection schemes. Our QNM theory provides an unprecedented opportunity to directly uncover the path of mode evolution in the quantum tunneling regime. Here we employ the QHDM-based QNM theory for the investigation since QHDMs with recent developments [19–21] can self-consistently describe various nonlocal and quantum effects for nanospheres down to 1 nm diameter [21] and to quantum tunneling regime where classical local treatment fails completely [21, 47]. As shown in Fig. 1(a), we study a gold nanosphere (50 nm diameter) dimer and plot the extinction spectrum evolution map obtained from full numerical QHDM-based simulations [53]. The involved bonding plasmon (BP) and charge transfer plasmon (CTP) resonances are indicated. The first CTP is beyond 1 µm and not shown here. The map of extinction spectra in Fig. 1(a) is in excellent qualitative agreement with the experimental spectra in Fig. 2 of Ref. [23]. As shown in more detail in SM [53], good quantitative agreements are found by modeling the same geometry and illumination scheme as in Ref. [23]. Then we perform QNM analysis to identify the eigenmodes of the system. As a benchmark for the orthonormal relation, we reconstruct the extinction spectrum for \( d = 0.3 \) nm. Figure 1(b) displays a perfect agreement between the full simulation and reconstruction with the contributions of 30 QNMs.

Previous far-field spectroscopic studies [23] recognize the onset of quantum tunneling at the gap of 0.3 nm, which seems to be confirmed by our extinction spectra in Fig. 1(a) (green dashed line). However, based on our QNM analysis, the mode evolution paths on the complex wavelength plane in Fig. 1(c) indicate that the mode transition occurs at a smaller gap of 0.2 nm, coincident with the kink on the BP1-CTP2 dashed trace in Fig. 1(a). It is better evidenced by the modal current profiles in Fig. 1(d). On the threshold, current at the gap center emerges and concomitantly the mode order changes. The mode has the smallest volume \( V_{\text{min}} \) as shown in an inset of Fig. 1(c) [53]. Despite the accompanying lowest quality factor, we emphasize that the optical responses under far-field excitation can still be conceptualized with the mode. The extinction spectrum around the resonance (680 nm) are dominated by the BP1/CTP2 mode. The reconstruction of the spectrum in terms of modal contributions at \( d = 0.2 \) nm, and the mode profile evolution for high-order modes are provided in SM [53].

**QNM analysis of a nonlocal polar dielectric nanoparticle.**— The optical responses of polar dielectrics are classically characterized with the Fröhlich resonance. As the particle size shrinks to nanometer scale, nonlocal effects owing to longitudinal phonons become significant [32, 66]. Here we perform QNM analysis for a SiC nanosphere (10 nm diameter) and unveil how the nonlocal responses are dictated by longitudinal phonon polaritons. Firstly the QNMs of the nanosphere are arranged according to the complex eigenfrequency as in Fig. 2(a). Consistent with the dispersion of longitudinal phonons [67], the eigenfrequencies decrease with the mode order. Meanwhile the imaginary parts are essentially constant \( \Im\{\omega_m\} = -\gamma_\alpha/2 \), which implies the electric fields of these QNMs are confined inside the nanosphere and largely longitudinal (left panel of Fig. 2(d)). Their magnetic fields could be dipolar (D), quadrupolar (Q), octupolar (O) and etc (right panel of Fig. 2(d)).

Assuming a plane wave illumination, the extinction spectra are calculated for both nonlocal and local responses. As shown in Fig. 2(b), the nonlocal corrections introduce extra resonances and can be perfectly reconstructed with modal contributions from 11 QNMs, confirming the validity of our orthonormal relation for nonlocal polar dielectrics. For the far-field plane wave excitation, only the dipolar modes are excited. By placing an electric dipole 5 nm away from the nanosphere, a wealth
FIG. 2. (a) QNM spectrum for a SiC nanosphere (diameter 10 nm; \( \gamma_d = 4 \text{ cm}^{-1} \)). (b) Extinction spectra of the nanosphere for a plane wave. (c) Radiation enhancement spectra of an electric dipole placed 5 nm away from and perpendicular to the nanosphere surface. Results of local, nonlocal full-wave simulations and QNM reconstructions are plotted in dash-balck traces, gray shadings and red traces, respectively. The pinkish shading indicates the region without phonon polariton modes. (d) A gallery of modal electric and magnetic field profiles. See SM [33] for the details of calculations.

Conclusion.— We have formalized a theoretical framework of QNM analysis for general nonlocal media under linear mean-field approximation. The proposed GLM incorporates various kinds of specific nonlocal and quantum effects into a concise form and leads to a canonical definition of QNMs and to the construction of an orthonormalization scheme. The exemplary embodiments for QHDM and nonlocal polar dielectrics are explicitly shown. We applied the QHDM-based QNM analysis to reveal that the intrinsic bonding-to-tunneling mode transition in the quantum tunneling regime occurs at a smaller gap than inferred from far-field spectroscopic studies. Our work greatly facilitates analytical formulations of the electromagnetic responses in general nonlocal media and expands the application scenarios of QNM analysis for e.g. Raman spectroscopy [14], photon emissions from tunneling devices [53, 68] and molecular junctions [69], and single photon emission from nanocavities [70]. Therefore we hope it constitutes a valuable asset for nano-optics.

We acknowledge financial support from the National Natural Science Foundation of China (Grant Number 9215011 and 11874166). We thank Wei Yan for useful discussions.
