Chapter 7
Ultrafast and Nonlinear Plasmon Dynamics

Markus B. Raschke, Samuel Berweger and Joanna M. Atkin

Abstract The interaction of light with a metal mediated by surface plasmon polaritons provides for sub-diffraction limited optical confinement and control. While the relationship of the linear plasmon response to the underlying elementary electronic excitations of the metal is well understood in general, the corresponding ultrafast and nonlinear plasmon interactions could provide further enhanced functionalities. However, while the ultrafast and nonlinear optics of metals is an advanced field, the understanding of the related plasmonic properties is less developed. Here we discuss ultrafast and nonlinear wave-mixing properties of metals and metallic nanostructures in terms of the elementary optical interactions related to electronic band structure, plasmon resonances, and geometric selection rules. These properties form the fundamental basis of the nonlinear plasmonic light-matter interaction. The understanding of these fundamental properties, together with the ability to measure and control the typically fast femtosecond intrinsic and extrinsic dephasing times, is important for the development of applications such as enhanced nano-imaging, coherent control of individual quantum systems, strong light-matter interaction and extreme nonlinear optics, and nano-photonic devices.

Keywords Nonlinear optics · Metal optics · Plasmonics · Ultrafast dynamics

7.1 Electronic Excitation at Metal Surfaces: Surface Modes

7.1.1 Introduction

Optical excitations of electrons at metal surfaces play an important role in a wide range of fundamental and applied science applications. The elementary electronic

M. B. Raschke (✉) · S. Berweger · J. M. Atkin
Department of Physics, Department of Chemistry, and JILA, University of Colorado,
Boulder, CO 80302, USA
e-mail: markus.raschke@colorado.edu

T. V. Shahbazyan and M. I. Stockman (eds.), *Plasmonics: Theory and Applications*,
Challenges and Advances in Computational Chemistry and Physics 15,
DOI: 10.1007/978-94-007-7805-4_7, © Springer Science+Business Media Dordrecht 2013

excitations determine carrier and thermal transport, and surface photo-chemistry including photocatalysis, with distinct characteristics for planar and nanostructured metal surfaces. Collective excitations in the form of surface plasmon polaritons (SPPs) at the boundaries of metallic media permit the tailoring of optical fields for surface-enhanced spectroscopy and sub-wavelength resolution microscopy, and have attracted wide attention for their potential for the design of new micro- and nano-scale photonic devices. The near-field and optical antenna properties of surface plasmon polaritons, in particular, may even open the door for qualitatively new optical physics in the near-field. This includes new ways to control the light-matter interaction in quantum systems, negative-index and related meta-materials, or new nonlinear optical phenomena.

The possibilities and fundamental limitations associated with several of these proposed ideas are linked to the fundamental properties of surface plasmon polaritons. Excellent books have been devoted to their linear optical properties [1–3]. Here we will discuss in particular the ultrafast and nonlinear optical properties of surface plasmon polaritons, and how they relate to the elementary electronic properties of metals that ultimately determine the radiative and non-radiative evolution of the plasmon excitation.

We start with a brief overview of the basic properties of surface waves and plasmon polaritons and the relevant underlying physics. We then focus on the ultrafast and nonlinear behaviour, which just as for linear SPPs is a convolution of the *intrinsic* elementary electronic properties of metals with *extrinsic* size and geometry-dependent structural resonances. Ultrafast and nonlinear optics involving SPPs are particularly sensitive to the combination and relative roles of intrinsic and extrinsic effects. The goal here is to provide a microscopic discussion of the dynamic processes of SPPs and the parameters that govern their spectral, spatial, and temporal characteristics linked to the ultrafast electron dynamics of metals.

The nonlinear SPP response is of interest for the generation of optical frequency components by parametric generation or nonlinear wave-mixing, optical saturation and gain, and strong light-matter interaction, taking advantage of the nonlinearity of the medium in combination with resonant or non-resonant field enhancement and optical antenna effects. The study of the ultrafast dynamics of SPPs also opens new experimental approaches for controlling the light-matter interaction and super-resolution microscopy simultaneously on femtosecond time and nanometer length scales as determined by the elementary processes in (homogeneous and heterogeneous) media.

7.1.2 Linear Optical Polarization

We first discuss some basics of the light-matter interaction to define the relevant nomenclature. The induced optical polarization of a (nonmagnetic) medium subject to an incident electromagnetic field is given by

$$\boldsymbol{P}(\omega) = \varepsilon_0 \chi(\omega) \boldsymbol{E}(\omega) \tag{7.1}$$

in the frequency domain, where $\chi(\omega)$ is the dielectric susceptibility, with the frequency-dependence in general arising from material resonances. $\chi(\omega)$ is related to the relative dielectric permittivity by $\varepsilon(\omega) = 1 + \chi(\omega)$, and the complex index of refraction $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega) = \sqrt{\varepsilon(\omega)}$. Both $\chi(\omega)$ and $\varepsilon(\omega)$ are tensor properties, but we will initially consider the medium to be isotropic.

Alternatively, the optical response can be described by an induced electrical current $\boldsymbol{j}(\omega)$ as

$$\boldsymbol{j}(\omega) = \sigma(\omega)\boldsymbol{E}(\omega) \tag{7.2}$$

with electrical conductivity $\sigma(\omega)$. The relationship between the typically complex $\sigma(\omega) = \sigma_1(\omega) + i\sigma_2(\omega)$ and $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$ is given by

$$\sigma(\omega) = -i\varepsilon_0\omega[\varepsilon(\omega) - 1]. \tag{7.3}$$

7.1.3 Time Domain Description

The standard frequency domain description of linear optics and the induced polarization from above has an equivalent time domain formalism. In this case the optical polarization at time t and location $\boldsymbol{r}$ is the result of the interaction of the optical electic field $\boldsymbol{E}(\boldsymbol{r}, t)$ with the medium at earlier times and possibly different locations (non-local response):

$$P(\boldsymbol{r}, t) = \varepsilon_0 \int_{-\infty}^{\infty}\int R(\boldsymbol{r}, \boldsymbol{r}', t')E(\boldsymbol{r}', t - t')d^3\boldsymbol{r}'dt' \tag{7.4}$$

where $R(\boldsymbol{r}, \boldsymbol{r}', t')$ is the *response function* of the system. The response function encodes the memory of the system, with causality dictating that for $t < 0$, $R(\boldsymbol{r}, \boldsymbol{r}', t) = 0$. Additionally, time invariance means that the dynamical response of the system is unchanged by a time offset. For most situations discussed in this chapter, we can neglect the spatial dependence of the response function.[1] The linear susceptibility in the frequency domain can then be derived from the response function as

$$\chi(\omega) = \int_{-\infty}^{\infty} R(t')e^{i\omega t'}dt'. \tag{7.5}$$

[1] In general the dielectric function is wave vector dependent, $\varepsilon(\boldsymbol{k}, \omega)$. However, for the regime discussed here, we can apply the local approximation $\varepsilon(\boldsymbol{k} = 0, \omega) = \varepsilon(\omega)$. Non-local effects and the associated spatial dispersion become significant for $\omega = v_F k$, where v_F is the Fermi velocity. This corresponds to $k > 1\ \text{nm}^{-1}$, i.e. structure sizes of a few nanometers at optical frequencies [4]. Note, however, that this effect is different from the finite-size effect, which can also alter the dielectric function when structure sizes become less than the characteristic scattering length or the onset of quantum confinement.

The well known Kramers-Kronig relationship relates real and imaginary parts of the susceptibility to each other.

A medium is considered instantaneously responding when the excitation is far off-resonant, meaning that the polarization at $t = t_0$ depends only on the electric field at that point in time. Resonant interactions are associated with memory effects, and relaxation processes following the excitation.

The usual Fourier transform relationships hold between the time and frequency domains and the two descriptions are complementary. While the frequency-domain description is typically employed for monochromatic optical interactions, the time-domain provides a more convenient way of analyzing problems where the excitation is induced by a short optical transient. The time domain analysis is therefore useful in particular for the dynamical properties of SPPs.

7.1.4 Electronic Properties of Metals

The spectral and temporal characteristics of SPPs for a metal are ultimately determined by the intrinsic electronic structure of the supporting metal. Most typical metals have hybridized *sp* bands that are parabolic to first order, with a density of states (DOS) that varies weakly within the range of a few eV above and below the Fermi level. *sp* bands resemble the free electron behavior typical for an *s*-metal [4]. In the absence of other electronic states or for sufficiently small photon energies, the optical response of the metal is determined by indirect *intraband* excitations.

Transition metals are characterized by an occupied *d*-band a few eV below the Fermi level, and weak dispersion with high DOS. Figure 7.1 shows a schematic of the

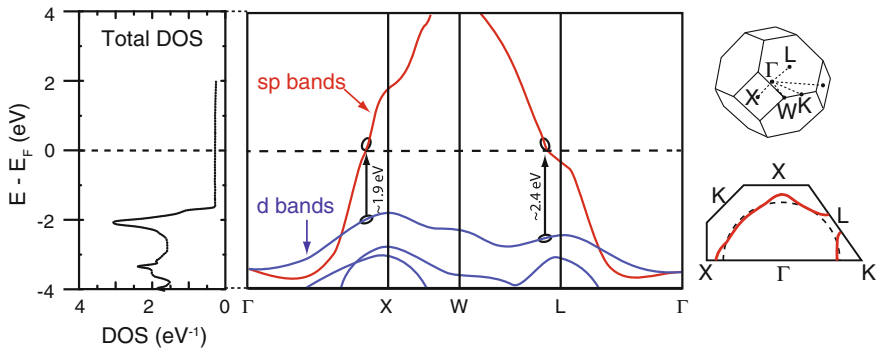

Fig. 7.1 Schematics of the calculated density of states, band structure, and Fermi surface of gold. Gold displays free-electron behavior for low photon energies. The absorption and color of the metal arise primarily due to the interband transition, from the occupied *d* band to the unoccupied *sp* band above the Fermi level. The onset is at ~1.9 eV near the X-point, which leads to a long tail in the experimentally observed absorption spectrum, and approximately 2.4 eV for the L-point, producing a sharp transition. The resulting dielectric function spectral response is shown in Fig. 7.2. DOS and band structure adapted from Ref. [5], and Fermi surface based on Ref. [6]

band structure for Au ([Xe] $4f^{14}5d^{10}6s$) as an example, near the high-symmetry X and Γ points. The dominant contributions to the *interband d-sp* transition are shown, with onset at ~1.9 eV and sharp rise at 2.4 eV.

The topology of the Fermi surface resembles the free electron sphere within the first Brillouin zone, except for the $\langle 111 \rangle$ direction (L neck). Near the Fermi level E_F the optical absorption is weak due to the absence of direct transitions, but is allowed for finite ω if translational symmetry is broken. This is the case for electrons with momenta $2\pi/l$, where l is the electron scattering length. With increasing wavelength this gives rise to an increase in absorption. The effect of the d band on the optical properties is discussed further in Sect. 7.1.6 after introducing the free electron response.

7.1.5 Drude-Sommerfeld Model

Classically the motion of carriers in a metal can be described as ballistic under the assumption of negligible Coulomb interaction. This is the so-called *free electron response* of metals, described by the Drude-Sommerfeld model [7]. Damping, which gives rise to ohmic resistance, can be introduced via the assumption of inelastic and instantaneous collisions with unspecified scattering centers. This leaves as the only key parameter the time τ_D between collision events, defining a relaxation rate $\Gamma = 1/\tau_D$. The equation of motion describing that *relaxor* behavior then corresponds to that of a damped harmonic oscillator without a restoring force term, giving rise to an apparent resonance at $\omega = 0\,\mathrm{s}^{-1}$. This is the Drude peak, describing the increasing absorption with decreasing frequency as mentioned above.

It is instructive to first consider the ideal Drude response without damping. From the equation of motion of free carriers of density n subject only to a driving external optical field $E(t)$, the induced optical polarization is given by

$$P(t) = -nex(t) = -\frac{ne^2}{m\omega^2}E(t), \tag{7.6}$$

with $x(t)$ the separation of electrons from the ions under the influence of the driving field, and m and e the electron mass and charge, respectively. Based on that expression the dielectric function of the charge plasma can be derived as

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2}, \quad \text{with } \omega_p = \sqrt{\frac{ne^2}{\varepsilon_0 m}} \tag{7.7}$$

termed the volume plasma frequency. This dielectric function is purely real, reflecting the absence of an energy dissipating term. The conductivity $\sigma(\omega)$ is then purely imaginary, indicating a 90° phase shift between an applied field and the induced

current. This reflects the hypothetical picture of a current that will persist infinitely long after a field is no longer applied.

Including damping in the form of scattering to describe the relaxation of the electron momentum with rate $\Gamma = 1/\tau_D$, the resulting dielectric function takes the form

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\omega\Gamma}, \tag{7.8}$$

that gives rise to an imaginary component of $\varepsilon(\omega)$, which describes the ohmic resistance.

At low frequencies $\omega \ll 1/\tau_D$, in the so called *Hagen-Rubens* regime, the polarization (current) is in phase with the driving field, hence real and purely dissipative. The conductivity is mostly real and frequency independent, and for $\omega \to 0$ converges to $\sigma_{DC} = ne^2\tau_D/m$. This conductivity is also used to describe radio frequency antenna resonance behavior. At intermediate frequencies, with the optical cycle period becoming comparable to τ_D at mid-infrared frequencies, the imaginary conductivity $\mathrm{Im}(\sigma(\omega))$ peaks at $\omega = 1/\tau_D$ and is equal to the real part $\mathrm{Re}(\sigma(\omega))$. Here, a phase lag appears between the applied field and current response due to the inertia of the electrons. Above $\omega = 1/\tau_D$ (into the near-IR) is the *relaxation* regime, where the response is characterized by decreasing real and imaginary parts with $\mathrm{Re}(\sigma(\omega))$ remaining larger than $\mathrm{Im}(\sigma(\omega))$, leading to large ohmic loss and phase lag, and consequently high damping of SPPs.

7.1.6 Interband Transition and Hybridization

Despite the fact that the electrons obey quantum statistics, the Drude model provides a satisfactory description for the observed dielectric function over a wide energy range well below the interband transitions.[2] However, as a purely phenomenological model it does not provide any physical insight into the damping mechanism, and requires modification for frequencies in the visible and near-IR as the optical frequencies approach d-band resonances.

Table 7.1 summarizes typical Drude and other parameters for Cu, Ag, and Au as representative free electron d-metals. From the Fermi velocity v_F and τ_D an effective electron mean free path $l = v_F\tau_D$ can be estimated between successive scattering events. To account for electron correlation effects in a heuristic manner, a correction to the electron rest mass via an effective mass m^* can be introduced.

The contribution of the positive ion cores to the dielectric function, which is not included in the Drude model, can be corrected for through an empirical, largely frequency independent term ε_∞, with typical values between 1 and 10 depending

[2] A quantum mechanical treatment of the electromagnetic response is provided in the form of the Kubo model. It is based on the fluctuation-dissipation theorem, and an interaction Hamiltonian to describe the interaction of the electromagnetic field with the charge carriers [8].

Table 7.1 Free carrier density n, plasma frequency ω_p, Drude relaxation time τ_D, effective mass m^*, correction term ε_∞, Fermi velocity v_f, band edge E_g, skin depth at 1 eV δ, and mean free path l for Cu, Ag, and Au. n and v_F from Ashcroft and Mermin [4], τ_D, ε_∞, and band edge from fitting to NIR-vis data in Johnson and Christy [9], and ω_p calculated from Eq. 7.9.

	n [cm^{-3}]	$\hbar\omega_p$ (eV)	$\tau_D = 1/\Gamma_0$ (fs)	v_F (nm/fs)	E_g (eV)	l (nm)	m^*/m	ε_∞	δ (nm)
Cu	8.47×10^{22}	8.85	6.9 ± 0.7	1.57	~2.4	~11	1.49	1.6	24
Ag	5.86×10^{22}	9.17	31 ± 12	1.39	~3.8	~43	0.96	3.7	22
Au	5.90×10^{22}	9.07	9.3 ± 0.9 to 14 ± 3	1.40	~2.15	~13	0.99	9.84	24

For Au, τ_D = 14 fs extracted from true Drude free electron behavior [10], with (9.3 ± 0.9) fs reflecting modifications in behavior due to polarization of core electrons at shorter wavelengths [9]. (Values for 300 K.)

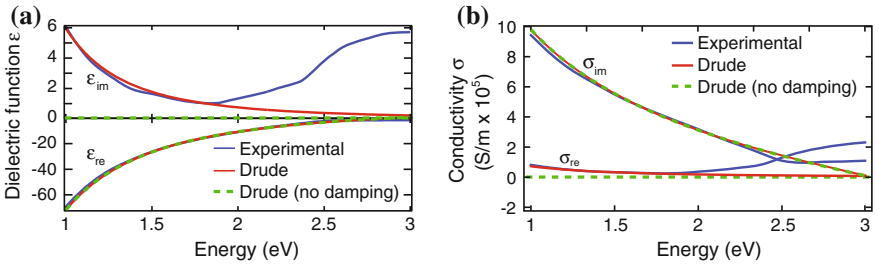

Fig. 7.2 Dielectric function $\varepsilon(\omega)$ for Au for ideal Drude behavior with (*red*) and without damping (*green*), in the near-IR to visible spectrum (**a**). The experimentally measured dielectric function for Au [9] (*blue*) shows the deviations at high frequencies, due to the contributions from interband transitions. **b** shows the corresponding conductivities $\mathrm{Im}(\sigma)$ and $\mathrm{Re}(\sigma)$

on, e.g., the degree of sp electron hybridization with ion core states (see Table 7.1). The dielectric function $\varepsilon(\omega)$ then takes the form

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\omega\Gamma}, \quad \text{with plasma frequency } \omega_p = \sqrt{\frac{ne^2}{\varepsilon_0 m^*}}. \tag{7.9}$$

The theoretical Drude behavior of Au for the parameters from Table 7.1, with and without damping Γ, in comparison with experimentally measured values [9, 10], are shown for $\varepsilon(\omega)$ and $\sigma(\omega)$ at visible frequencies in Fig. 7.2a, b, respectively. The Drude model provides a good fit to the data for energies below ~2 eV, but diverges above that energy due to the onset of $sp - d$ interband transitions.

The absorption spectrum of d-electron metals is characterized by the direct *interband transition* from d to sp bands (Fig. 7.1), with the absorption proceeding largely from the top of the d-band due to its high DOS. The excitation of free carriers via intraband sp band absorption is weak in comparison, since it requires additional momentum scattering, primarily through phonon scattering, but also scattering with impurities, defects, the surface, or other electrons. However, even the behavior below the interband transition is strongly affected by the d-bands through the hybridization

of sp and d bands (copper ($3d-4sp$), silver ($4d-5sp$), and gold ($5d-6sp$)), which gives rise to mutual polarization of the s and d electrons and deviations from the ideal free electron behavior. The d-band is therefore an integral part of the collective electron excitation, even in the regime where the free electron model effectively describes the optical response.

The contribution of the d-bands can empirically be accounted for by writing $\varepsilon(\omega)$ in the form $\varepsilon(\omega) = \varepsilon_{Drude}(\omega) + \varepsilon_d(\omega)$. The Drude term is well described by the sp electron density behavior. The d-bands can then be described in an extended Drude or Drude-Lorentz type model, with the d-electrons assigned an effective Coulomb restoring force, and with a certain density of oscillators, to parametrize the response.

7.1.7 Optics at Metal Interfaces

Following the general discussion of the optical properties of basic metals expressed through their dielectric function, we proceed to the description of surface plasmon polaritons (SPPs) as a collective excitation. A unified description of optical surface wave phenomena, in particular the notion of SPPs, was developed from the work of Sommerfeld, Zenneck, and Wood at the beginning of twentieth century, with subsequent experimental studies by Ritchie, Stern, Kretschmann, and Raether, together with related work by Mie. The optical excitation of the free electrons at the interface of a metal with vacuum or a dielectric medium gives rise to a collective oscillation of the carriers. This surface charge density oscillation is associated with a time varying optical field, hence the notion of a surface plasmon polariton.[3]

The electron charge density at a metal surface decays to zero on a scale comparable to the Fermi wavelength λ_F (the de Broglie wavelength of electrons at the Fermi energy, ~0.5 nm for Au and Ag) (Fig.7.3). Despite being a dynamic surface response

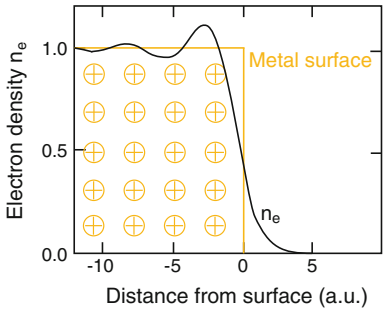

Fig. 7.3 Bulk normalized electron density n_e along the surface normal direction across the metal-vacuum interface, with distance. Friedel oscillations due to electron wavefunction scattering at the interface characterize the density behavior inside the metal, with decay into the vacuum. The decaying electron density can extend several tenths of a nanometer beyond the geometric interface

[3] Instead of an electronic excitation underlying a surface plasmon polariton, collective excitation of lattice vibrations can give rise to a surface phonon polariton. The scope of this chapter is limited to surface plasmon polaritons, but the concepts discussed can readily be extended to phonon polaritons.

the SPP is determined by the local dielectric properties of the bulk (for structures with dimensions above the onset of finite size effects).

This surface wave phenomenon manifests itself in two dis tinct ways: either in the form of *propagating* SPP modes, or as *localized* SPP oscillations.The former surface-bound wave allows for energy propagation and transport over distances at the dielectric-metal interface. In the latter, the additional restoring force introduced to the electron motion by geometric constraints of, for example, noble metal nanoparticles, results in spatially localized resonant charge density oscillations.These can lead to large optical polarizabilities and local optical field enhancement. Signatures of both propagating and localized SPP modes can be observed in random, percolated, or clustered media.

7.1.8 Propagating Surface Plasmon Polaritons

From the wave equation with appropriate boundary conditions at a metal/dielectric interface, the dispersion relationship for propagating SPPs is given by

$$k_{\parallel}^2(\omega) = \frac{\omega^2}{c^2}\frac{\varepsilon_m(\omega)\varepsilon_s}{\varepsilon_m(\omega)+\varepsilon_s} \tag{7.10}$$

with dielectric permittivity of the metal $\varepsilon_m(\omega)$ and its surrounding ε_s (assumed to be frequency independent in the spectral range of interest).[4] Specifically for the metal/vacuum interface the resonant condition $\varepsilon(\omega) = -1$ results in $\omega_{\mathrm{sp}} = \omega_p/\sqrt{1+\varepsilon_\infty}$ for the surface plasmon resonance for a Drude metal.[5]

Figure 7.4 shows the ω versus k dispersion relationship for the ideal Drude surface plasmon polariton with and without loss. The SPP is characterized by surface parallel wave vectors that are large compared to light at optical frequencies. Only in the region near $k_{\parallel} \approx \omega/c$ does the surface plasmon couple to free-space electromagnetic radiation. For higher frequencies, the excitation requires additional momentum via, for example, direct evanescent excitation, grating coupling, or increased index of refraction of the adjacent medium.

Momentum conservation relates the propagating in-plane and evanescent out-of-plane wavevectors to the incident free space wavevector via $k_{\perp,i}^2 + k_{\parallel,i}^2 = \varepsilon_i k_0^2$ for both the metal ($i = 1$) and adjacent dielectric ($i = 2$). Im ($k_{\parallel}$) describes the finite propagation length along the interface. Unlike localized SPP resonances discussed below, due to the large electric field component of propagating SPPs that penetrates into the dielectric medium, lifetimes generally exceed the Drude damping time. $k_{\perp,i}$ governs the spatial extent of the evanescent field in the surface normal

[4] A wide range of interesting phenomena result for the case of frequency dependent or resonant surrounding media, but are beyond the scope of this chapter.

[5] In treatments of this subject the contribution of core electrons is frequently neglected, using $\varepsilon_\infty = 1$. This results in $\varepsilon_{\mathrm{sp}} = \omega_p/\sqrt{2}$ and a plasmonic bandgap in the range of $\omega_{\mathrm{sp}} < \omega < \omega_p$.

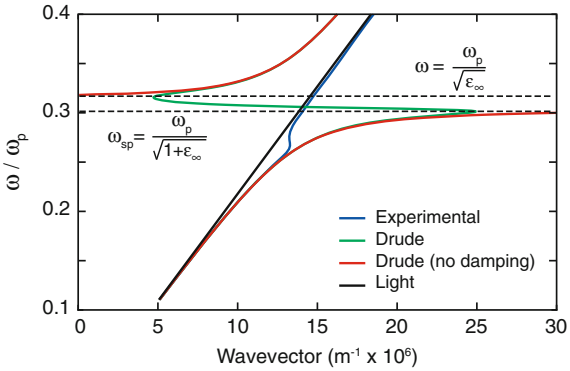

Fig. 7.4 Dispersion relation of surface plasmon polariton for the case of an ideal (undamped) Drude metal (*red*), and Au as an example of a real metal assuming both pure free electron Drude damping (*green*), and additional interband damping from experimental values (*blue*) [9, 10]. The frequency is normalized with respect to the plasma frequency ω_p

direction. The characteristic length $l_{\perp,i}$ is defined as the distance from the surface where $|E(z)/E(z=0)| = 1/e$. For medium i with complex dielectric function $\varepsilon_i = \varepsilon_i' + i\varepsilon_i''$, $l_{\perp,i}$ is given by

$$l_{\perp,i} = \frac{1}{|k_{\perp,i}|} = \frac{\lambda_0}{2\pi}\left(\frac{\varepsilon_1' + \varepsilon_2}{\varepsilon_i^2}\right)^{1/2}. \tag{7.11}$$

The penetration depth into the metal is related to the skin depth (the $1/e$ penetration depth of the optical field into the metal) $\delta = c/\kappa\omega$, with κ the imaginary part of refractive index $N = n + i\kappa$. For gold, $\delta \sim 22 - 25$ nm throughout the mid IR to visible spectral range (0.1–2 eV) [10]. The weak frequency dependence in that regime is due to the near linear dependence of $1/\kappa$ with ω.

The SPP is sensitive to a wide range of which affect surface modifications, including charge, contact to dielectrics, adsorbates, etc. which affect the surface dispersion relation through the modification of the dielectric function. This is closely related to the size and shape dependence of local SPPs in confined geometries (Mie and Rayleigh resonances) as discussed in the following section.

7.1.9 Localized SPP in Small Metal Particles

Localized SPPs are non-propagating modes that can be excited in structurally inhomogeneous environments of dimensions comparable to or smaller than the optical wavelength, i.e. where the translational invariance of the medium is lost on the length scale of the SPP wavelength. With R defining a characteristic structural dimension the parameter $1/R$ plays a role analogous to the parallel SPP wavevector for the case

of a flat surface. For a spherical particles discussed here as an example, R is the radius. In analogy to the wavelength of a surface wave of $\lambda = 2\pi/k$, with wavevector k, for a spherical particle the local mode can be described by an effective wavelength λ_{eff} given by the circumference as $\lambda_{\text{eff}} = 2\pi R$. This analogy implies $k \sim 1/R$.[6]

The optical response of a sphere of arbitrary radius can be solved exactly using (the albeit computationally intensive) *Mie theory* [11, 12], discussed here and also further below in the context of damping. Analyzing the limiting case of a spherical particle provides insight into the basic mechanisms underlying the particle response and its dependence on different input parameters. Many of the conclusions can be generalized to other simple geometries, such as rods, discs, etc.

For particles which are small compared to the wavelength, the response can be more simply described by an induced optical dipole in the lowest order approximation, when neglecting retardation. In that *quasistatic* limit ($R \ll \lambda$), provided the particle is still large enough to avoid finite size effects on the intrinsic dielectric properties,[7] the field distribution of the particle follows from the Laplace equation in spherical polar coordinates. The field outside the sphere is equivalent to the field of a point dipole at the center of the sphere with dipole moment $\boldsymbol{p} = \varepsilon_0 \varepsilon_s \alpha \boldsymbol{E}$. The (complex) polarizability α is given by the Clausius-Mossotti relation:

$$\alpha(\omega) = 4\pi R^3 \frac{\varepsilon_m(\omega) - \varepsilon_s(\omega)}{\varepsilon_m(\omega) + 2\varepsilon_s(\omega)}, \tag{7.12}$$

with $\varepsilon_m(\omega)$ the dielectric function of the metal and $\varepsilon_s(\omega)$ the dielectric function of the surrounding medium. The corresponding absorption cross section is given by $\sigma(\omega)_{\text{abs}} = k\text{Im}\alpha(\omega)$. Since $\sigma(\omega)_{\text{abs}}$ scales as R^3, it dominates for small radii over the scattering cross section $\sigma(\omega)_{\text{scat}} = (k^4/6\pi)|\alpha|^2$, which scales as R^6.

As seen from Eq. 7.12, for a particle in vacuum or air the Fröhlich resonance condition is given by $\text{Re}(\varepsilon_m(\omega)) = -2\varepsilon_s$, provided Im (ε_m) has a negligible frequency dependence. In a Drude metal the localized SPP resonance frequency is then given by $\omega_{\text{res}} = \omega_p/\sqrt{2 + \varepsilon_\infty}$. The resonance frequency redshifts with increasing index of refraction of the environment. Above $R \sim 50$ nm the onset of retardation and multipole excitation gives rise to spectral broadening and decrease in peak amplitude, which necessitates the application of the full Mie treatment.

The different dielectric properties of Au and Ag lead to pronounced differences in the spectral behavior close to the SPP resonance. For small spherical Au particles the plasmon resonance at $\lambda_{\text{SPP}} \sim 530$ nm is already superimposed on a pronounced increase in scattering and absorption due to the interband transition. In contrast, for Ag with the interband transition at ~4 eV, the SPP peak at 350 nm is dominated by

[6] $k \sim 1/R$ also describes to first order the wavevector distribution cut-off of the near-field modes of a structural element with characteristic dimension R, and their corresponding spatial localization.

[7] For particle size with radius $R \gg \lambda_{\text{TF}}$ with λ_{TF} the Thomas-Fermi screening wavelength, the response can be treated as that of the homogeneous bulk electron density. However, electron spill-over, as depicted in Fig. 7.3, may become significant in sub-nanometer particles [12, 13]. In addition, surface scattering becomes relevant when particle sizes approach the effective mean free paths of the excited electrons. (See also footnote 1.)

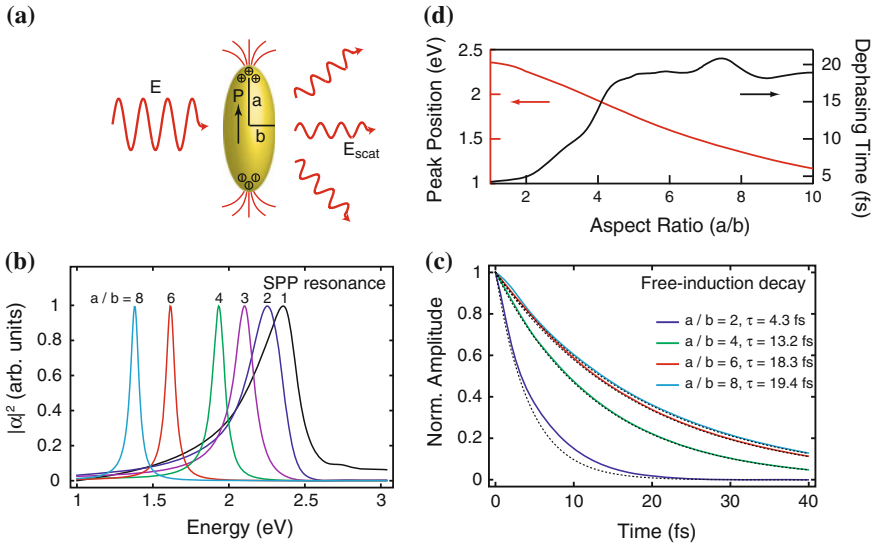

Fig. 7.5 Schematic of local SPP for a prolate spheroid with a/b for the aspect ratio of long to short principal axis (**a**). Normalized polarizability squared ($|\alpha|^2$) calculated using using experimental Au dielectric function [9] (**b**). The SPP resonance *red-shifts* with increasing aspect ratio starting with a spherical particle ($a = b$) with fixed major axis length. The interband continuum response increasing to high energies, has been subtracted for clarity. Fourier transform showing the underlying time domain evolution of the plasmon dynamics (**c**). SPP dephasing time as a function of aspect ratio and thus SPP energy exhibits a decrease from $T_2 \sim 20$ fs for the free electron behavior for energies above ~ 2 eV due to the interband contribution (**d**). *Dotted lines* in panel (**c**) show exponential fits for extracting dephasing times shown in (**d**)

the free electron response. Under otherwise identical conditions, the cross sections for Ag spheres are about one order of magnitude larger than those for Au.

A useful extension of the sphere model, also in the discussion of the damping of the plasmon response, is the SPP of an ellipsoid ($a \neq b \neq c$) or spheroid ($a = b \neq c$) shaped particle, treated in the quasistatic approximation. The longitudinal polarizability for a *prolate spheroid* (see Fig. 7.5a) with aspect ratio a/b is given by

$$\alpha(\omega) = \frac{4\pi ab^2}{3} \frac{\varepsilon_m(\omega) - \varepsilon_s(\omega)}{\varepsilon_s + L(\varepsilon_m(\omega) - \varepsilon_s)}, \tag{7.13}$$

where L is the so-called depolarization factor, an integral reflecting the particle aspect ratio. As shown in Fig. 7.5b the plasmon resonance shifts to lower energies with increasing aspect ratio. The red-shift can be viewed as a result of the increase in spatial charge separation and thus a decrease in effective restoring force. As we will see below this allows us to predict the frequency dependence of the plasmon dephasing and its correlation with the damping of the underlying dielectric function.

7.2 Damping of Surface Plasmon Polaritons

7.2.1 Theory of Radiative and Nonradiative Decay

The coherent electronic excitation of a medium is followed by fast electronic dephasing and the subsequent absorption and decay of the polarization into electron hole pairs.[8] Here we will discuss the radiative and non-radiative relaxation dynamics of SPPs as the fastest initial processes describing the light-matter interaction. We will restrict the discussion to the homogeneous SPP response, i.e., in the absence of ensemble effects and different inhomogeneities. We will discuss the basic physics of plasmon dephasing in this section, followed in subsequent sections by different frequency- and time-domain experimental linear and nonlinear spectroscopic results for its experimental determination.

Of primary interest is the electronic *dephasing*, that is, the eventual loss in phase coherence of the collective and initially phase coherent oscillation of the free electron gas (plasma oscillation). In contrast to semiconductors, which allow for a low and variable carrier density through controlled doping, the carrier density in metals is comparatively high and fixed (Table 7.1). Those high carrier densities immediately imply a high scattering and thus high dephasing rate. The SPP decoherence time is therefore fundamentally linked to the effective relaxation time in the Drude dielectric function as the response function that determines the temporal evolution of the induced optical polarization in response to an applied optical field. Consequently, to first order, the Drude relaxation time τ_D sets an upper limit for the dephasing time T_2 for a localized SPP.

The macroscopic optical response of metals in general, including the SPP resonance for plasmonic metal nanostructures, reflects the underlying elementary electron dynamics of the bound and conduction electrons involved. Specifically, the linewidth and shape of the SPP resonance in the frequency domain, or its Fourier transform in the form of the *free-induction decay* in the time domain, describes the loss in phase coherence, which in turn is directly linked to the dielectric function. In the following we discuss the ultrafast electron dynamics of spherical and spheroidal metal nanoparticles as model systems using analytical treatments. The results can readily be generalized for more complex geometries using numerical techniques. For small enough particles ($R \ll \lambda$) the excitation is dominated by the dipolar SPP response, with polarizability given by Eq. 7.12 for a sphere or Eq. 7.13 for a spheroid.

[8] Following the typically up to 10s fs coherent evolution of electronic excitations, different processes govern the incoherent carrier cooling and equilibration. The decay of the coherent excitation into electron-hole pairs gives rise to hot non-equilibrium and non-thermal carrier distributions. Electron–electron scattering leads to thermalization of the hot electrons within at most a few hundred fs and can often be described by the Fermi liquid theory. Electron–phonon interaction on 100 fs to ps time scales leads to the subsequent equilibration with the lattice degrees of freedom. Although not the subject of this review, these processes lead to transient variations of the dielectric function and its frequency dependence, ultimately due to the deposited energy in the form of heat. The processes need to be considered in time resolved experiments, especially with high pump intensities and large excitation densities, giving rise to a nonlinear response.

From the calculated spectra for a Au sphere or spheroids with different aspect ratios as shown in Fig. 7.5b, the corresponding time traces for the polarization decay are obtained by Fourier transform as shown in Fig. 7.5c. The associated SPP lifetimes can then either be directly deduced from the $1/e$ value of the maximum amplitude, or obtained from the FWHM (Γ) spectral line width by $T_2\Gamma = 2\hbar$. Note that the deviation from an ideal Lorentzian spectral or exponential time behavior in this model calculation is due to the use of the experimentally measured dielectric function $\varepsilon(\omega)$ as an input parameter and associated deviations from the ideal Drude behavior. The resulting *plasmon lifetimes* are then shown in Fig. 7.5d as a function of resonant energy (or aspect ratio). The dephasing times are found to be in the range of 18–22 fs for energies between 1.0 and 1.7 eV, i.e., the free electron regime. The dramatic decrease in lifetime at 2 eV is associated with the onset of the interband transition.

The interpretation of the dephasing time is complicated by the various possible mechanisms contributing to the loss of phase coherence in the plasmon oscillation. In general, the measured dephasing time T_2 is related to a population relaxation time T_1 of participating quantum states, and pure dephasing T_2^* by $1/T_2 = 1/2T_1 + 1/T_2^*$. The pure dephasing contribution T_2^* corresponds to elastic collisions of electrons, which destroy only the phase coherence. Because of the high carrier density and high electron scattering rate, this is expected to be negligible for SPPs. However, this relation has limited applicability in this case. The SPP classical polarization is described in terms of T_2, i.e. the polarization decay through inelastic electron scattering processes, but there is no associated population prior to the decay into electron-hole pairs after decoherence. Instead, the underlying momentum scattering, which gives rise to the loss in phase coherence, can be associated with the Drude scattering τ_D, with $T_2 \sim 2\tau_D$.

Drude scattering leads to electron-hole excitation and corresponding absorption effects, alternatively described via ohmic loss. It competes with radiative decay. The larger effective oscillator size and polarization with increasing particle size leads to an increase in the radiation damping contribution. For particles larger than ~20 nm this produces a monotonic trend of decreasing dephasing time with increasing particle diameter.

The quasistatic ellipsoid model discussed above and shown in Fig. 7.5 neglects radiative decay and thus provides only an upper limit for the dephasing time. In order to account for radiation damping, we use the rigorous solution for the scattering of a particle given by Mie theory [14, 11]. The scattering and extinction efficiencies for the m-th multipole order is related to the scattering and extinction cross sections $\sigma_{\mathrm{sca},m}$ and $\sigma_{\mathrm{ext},m}$, and the geometrical cross section $G = \pi R^2$, by

$$\begin{aligned} Q_{\mathrm{sca},m} &= \tfrac{\sigma_{\mathrm{sca},m}}{G} = \frac{2}{x^2}(2m+1)(|a_m|^2 + |b_m|^2), \text{ and} \\ Q_{\mathrm{ext},m} &= \tfrac{\sigma_{\mathrm{ext},m}}{G} = \frac{2}{x^2}(2m+1)Re(a_m + b_m), \end{aligned} \tag{7.14}$$

with $x = kR = \omega n_d(\omega)R/c$. The scattering coefficients a_n and b_n are given by

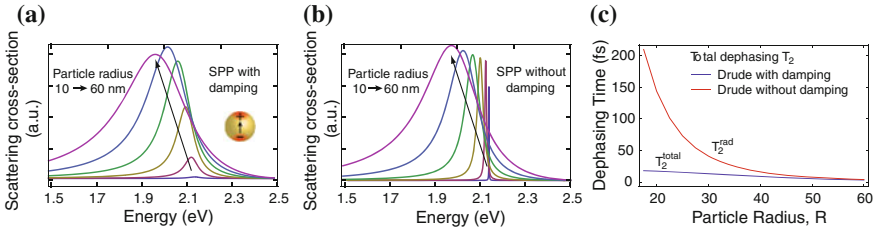

Fig. 7.6 Comparison of scattering cross-sections for the dipolar mode in spherical particles using Mie theory and the Drude model with (**a**) and without (**b**) intrinsic Drude damping of the metal electrons. The linewidths without that damping reveal the pure radiation damping contribution. Resulting dephasing times (**c**) based on linewidth analysis, demonstrating the relative contribution of radiative and nonradiative contributions with increasing particle radius [15]. Calculations utilize the Drude parameters for Au listed in Table 7.1, with $\varepsilon_\infty = 9.84$. The surrounding medium has index of refraction $n = 1.5$

$$a_m = \frac{N\psi_m(Nx)\psi_m'(x) - \psi_m(x)\psi_m'(Nx)}{N\psi_m(Nx)\xi_m'(x) - \xi_m(x)\psi_m'(Nx)}$$
$$b_m = \frac{\psi_m(Nx)\psi_m'(x) - N\psi(x)\psi_m'(Nx)}{\psi_m(Nx)\xi_m'(x) - N\xi(x)\psi_m'(Nx)} \tag{7.15}$$

with the relative refractive index $N = n_p(\omega)/n_d(\omega)$ of the particle (n_p) and the dielectric medium (n_d), and the Ricatti-Bessel functions ψ_m and ξ_m.

Shown in Fig. 7.6 is the result of the calculated spectral dependence of the scattering cross sections for spherical Au particles with increasing radius from Mie theory, using the Drude model parameters with (a) and without (b) damping (based on Eq. 7.9), and the resulting variation in dephasing times (c) [15]. The finite linewidths in the hypothetical absence of material damping (b) reveal the *radiation* contribution to the plasmon dephasing. The broader linewidths when including material damping (a) are due to contributions from both *radiative and nonradiative* dephasing, i.e.,

$$\frac{1}{T_2} = \frac{1}{T_2^{\text{rad}}} + \frac{1}{T_2^{\text{non-rad}}}, \tag{7.16}$$

with $T_2^{\text{non-rad}} \sim 18 - 22$ fs as discussed above. The increasing dephasing rate for larger particles is a result of the increasing contribution of radiation damping. As a result, the dephasing times for the damped and undamped Drude models converge for the case of large particles where the radiation damping due to the increasing dipole moment dominates over Drude scattering.

As SPPs oscillate in the visible spectral range with periods in the $\sim 2-4$ fs range, radiative decay times for electronic excitations in the 10 s of fs to sub 10 fs range thus imply a very good coupling of the optical dipole to the electromagnetic density of states in the far-field. The results can be compared to the emission of radiation from a classical dipole or the spontaneous emission from a quantum two level system.

In order to describe the radiative emission of a oscillating charge it must be recognized that the radiation field in turn influences the motion of the charge itself, termed *radiation reaction*. Assuming the radiation reaction force $\boldsymbol{F}_r$ as the only damping term, the equation of motion can be written as:

$$m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{dt}^2} + \omega_0^2 m\boldsymbol{r} = \boldsymbol{F}_r = -m\Gamma_0\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{dt}} = \frac{q^2}{6\pi\varepsilon_0 c^3}\frac{\mathrm{d}^3\boldsymbol{r}}{\mathrm{dt}^3}, \tag{7.17}$$

with the Abraham-Lorentz equation to describe the reaction force coefficient:

$$\Gamma_0 = \frac{1}{4\pi\varepsilon_0}\frac{2q^2\omega_0^2}{3mc^3}. \tag{7.18}$$

This gives rise to radiative lifetimes $\tau = 1/\Gamma_0 \simeq 20$ ns for optical frequencies. An additional term, conventionally introduced to describe the damping of a Lorentzian oscillator of the form $\Gamma \mathrm{d}\boldsymbol{r}/\mathrm{d}t$, contains both radiative and non-radiative contributions.

Similarly to Eq. 7.17, one can start with the induced optical polarization of the form:

$$\boldsymbol{P}(\omega) = \chi(\omega)\left(\boldsymbol{E}_{inc} + i\frac{2k_0^3}{3}\boldsymbol{P}(\omega)\right) \tag{7.19}$$

with particle susceptibility $\chi(\omega)$. The second term corresponds to the radiation reaction field with:

$$\boldsymbol{F}_r = e\boldsymbol{E}_{rad} = \frac{2}{3}\frac{e^2}{c^3}\ddot{\boldsymbol{v}} = i\frac{2}{3}\frac{\omega^3}{c^3}e\boldsymbol{x} = i\frac{2}{3}k^3\boldsymbol{P}, \tag{7.20}$$

using $\boldsymbol{x} = \mathrm{e}^{-i\omega t}$ and $\ddot{\boldsymbol{v}} = i\omega^3\boldsymbol{x}$ for the harmonic oscillator. Hence, both approaches are equivalent, with the difference that the damping for the resonant denominator for $\chi(\omega)$ already contains the a priori indistinguishable radiative and non-radiative terms.

Interestingly, the quantum description for the spontaneous emission of a two level system provides a qualitative intuition for the high radiative emission rate as derived from Mie theory in the femtosecond regime. The transition rate follows from Fermi's golden rule as

$$\Gamma_{sp} = \frac{\pi\omega_0}{3\varepsilon_0\hbar}|\langle a|\hat{\boldsymbol{\mu}}|b\rangle|^2\rho_\mu(\boldsymbol{r}_0, \omega_0), \tag{7.21}$$

with transition dipole moment operator $\hat{\boldsymbol{\mu}}$ and ρ_μ the partial local density of states (LDOS) at the location $\boldsymbol{r}_0$ of the system, given by $\rho_v(\omega) = \omega_0^2/\pi^2c^3$ in vacuum. With $\boldsymbol{\mu_{ba}}^2 = |\langle a|\hat{\boldsymbol{\mu}}|b\rangle|^2 = q^2r_{21}^2$ the spontaneous emission rate becomes:

$$\Gamma_{sp} = \frac{\omega_0^3}{3\pi\varepsilon_0\hbar c^3}\boldsymbol{\mu_{ba}}^2. \tag{7.22}$$

For an atomic emitter with typically $\boldsymbol{\mu_{ba}}$ = (1 electron charge)·(0.1 nm) (or 1.602×10^{-29}C·m), and $\omega_0 = 2\,\text{eV}/\hbar = 3.04 \times 10^{15}$ rad/s, the corresponding spontaneous emission lifetime is $\tau = 1/\Gamma_{sp} = 33$ ns. Since the results in the weak perturbation regime are similar for the classical and quantum treatment (oscillator strength ~ 1), we can rewrite Eq. 7.22 in the following semi-classical form:

$$\Gamma_{sp} = \frac{8e^2\pi^2}{3\varepsilon_0\hbar}\left(\frac{r}{\lambda_0}\right)^2 \frac{1}{\lambda_0}. \tag{7.23}$$

This equation highlights the size mismatch r/λ_0 giving rise to the long ns radiative lifetimes for atomic emitters. Considering the SPP nanoparticles as an optical dipole with $r = 10...100$ nm, compared to the 0.1 nm of atomic dimensions, will increase the effective size of the dipole moment and thus reduce the radiative impedance mismatch. For a one-electron oscillator of that size the radiative rate would increase by $10^2 - 10^6$ and with that the dephasing time would decrease from the ns into the fs regime as seen for a localized SPP of a metal nano-particle. Note that this model merely qualitatively describes the general trend of an increase in radiative rate with increasing oscillating charge separation, with details depending sensitively on geometry.

7.2.2 Experimental Studies of Plasmon Lifetimes

A range of studies have investigated dephasing times from both time resolved and spectral line width analysis (see, e.g., [16, 17] and references therein). Here we discuss frequency-domain measurements of T_2, with further time-resolved experiments provided in Sect. 7.4. Since the time scales for plasmon dephasing are in the few femtosecond regime and the relative contributions of radiative and nonradiative decay pathways are size-dependent, accurate measurements of the intrinsic dephasing time usually require either a homogeneous sample or individual nanostructure, and various model assumptions are employed. Dark field scattering of individual particles or persistent spectral hole burning give access to the homogeneous sub-ensemble of an inhomogeneous sample. In hole burning, for example, the sub-ensemble with resonance close to that of the exciting laser frequency is bleached, and the linewidth of the spectral hole at different fluences is extrapolated to zero fluence to establish the dephasing time.

As shown in Figs. 7.5 and 7.6, the linewidth and related quality factor $Q = \omega_0/\Gamma$ can be derived with a particle SPP calculation using either the Drude model or experimental dielectric values. Most experimental results indicate 5–10 fs for dephasing time T_2, i.e. reduced from the theoretical maximum nonradiative value of $\sim$18 fs. While the limiting nonradiative case has been demonstrated [18], the shorter dephasing times often reported may result from structural inhomogeneities, surface scattering, and radiation damping. Measurements of the dephasing times of

Ag and Au particles as a function of their geometrical aspect ratio found that the dephasing times for higher aspect ratios are longer than for similarly sized spherical particles. Larger particles have shorter radiative dephasing times, consistent with observations from Mie theory and the quasistatic approximation. The SPP decay time for long rods approaches 20 fs, indicating that dephasing for these geometries is dominated by nonradiative Drude relaxation τ_D.

The different geometric behavior is important for applications of plasmonic structures. For mediating the coupling of nanoscopic emitters to far-field emission, increased radiation rates and spherical particles are preferred. However, the dephasing time is directly related to the field enhancement $T_2 \propto F$, so for many applications it is desirable to instead maximize the plasmon lifetime.

There are various *momentum scattering* contributions to the T_2 SPP dephasing discussed above. Electron–electron, electron–phonon, electron-defect, impurity, and surface scattering can all contribute, so that the total decay rate is the sum of these different contributions,

$$\Gamma = \sum_i \tau_i^{-1} = \tau_{e-e}^{-1} + \tau_{e-ph}^{-1} + \tau_{e-defect}^{-1} \tag{7.24}$$

All of these processes have been found to be largely temperature independent with the exception of electron–phonon scattering, which shows a linear increase with temperature, explained with a Debye model for the material-dependent electron–phonon interaction [19].

In addition to the extrinsic dependence of particle plasmon properties on size, with dielectric constant $\varepsilon = \varepsilon_{\text{bulk}}$, intrinsic *size effects* occur for particles where the size approaches the mean free path of the conduction electrons. This regime, relevant for few nanometer sized particles, is characterized by increased electron scattering from the particle surface and $\varepsilon \neq \varepsilon_{\text{bulk}}$. A radius-dependent correction to the Drude damping can be introduced empirically [12, 20]:

$$\Gamma(R) = \Gamma_\infty + \frac{A v_F}{R} \tag{7.25}$$

with the bulk Drude damping Γ_∞ and Fermi velocity v_F. A has a value near unity depending on particle geometry and the $1/R$-dependence follows from the ratio of the surface area to particle volume.

The short timescales and multiple relaxation processes involved in SPP dephasing lead to difficulty in interpreting results and separating the various effects in both frequency and time domain measurements. For time domain measurements, a challenge arises that for plasmon resonances in the visible to near-IR, the SPP dynamics on few femtosecond timescales are comparable to the shortest possible laser pulses in that energy range (e.g. ~2 fs optical cycle period at $\lambda_{\text{SPP}} = 600$ nm). In the following section as an application of the nonlinear SPP response we will also discuss nonlinear optical time-resolved techniques for the investigation of the ultrafast plasmon dynamics.

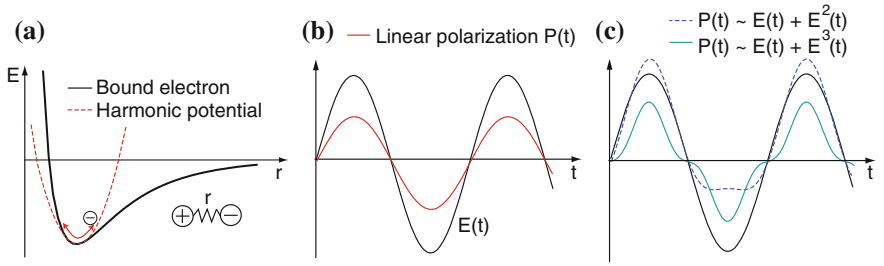

Fig. 7.7 **a** Interaction potential experienced by a bound electron in a medium. The deviation from a purely harmonic potential leads to a nonlinear optical polarization response under high driving fields. **b** The far off-resonant linear polarization $P(t)$ (*red*) in response to a weak driving field $E(t)$ (*black*). **c** The corresponding induced polarization incorporating a second-order response (i.e. $P(t) \propto E^2(t)$), for example in a non-centrosymmetric material (*blue*), and a third-order response (*green*), for large driving fields

7.3 Nonlinear Plasmon Optics

In this section we will first discuss the nonlinear optical response of metallic nanostructures, nonlinear resonant effects, and selection rules. We will then show how they can provide a means of separating the complex interaction of dephasing processes, for example from investigation of their relative phase, and also enable precise characterization of electric fields and response functions.

Thus far we have been assuming that the optical polarization P of the metal is linear with respect to the applied optical field, which applies for the case of a relatively weak driving field. However, if the incident driving field is comparable to electric fields within the medium a *nonlinear response* can result due to the deviation from a perfect harmonic oscillator potential experienced by the charge carriers coupling to the optical field. This anharmonic oscillator behavior is shown schematically in Fig. 7.7 for a bound electron in a medium. In metals, the polarization perpendicular to the surface is particularly important for second-order nonlinearities, since at the surface the electrons will experience an additional surface asymmetric potential.

A small nonlinearity can be treated *perturbatively*, so that the polarization is expressed as a power series expansion in the driving field:

$$\boldsymbol{P} = \varepsilon_0 \boldsymbol{\chi}^{(1)} \boldsymbol{E} + \varepsilon_0 \boldsymbol{\chi}^{(2)} \boldsymbol{E}^2 + \varepsilon_0 \boldsymbol{\chi}^{(3)} \boldsymbol{E}^3 ...., \tag{7.26}$$

with $\boldsymbol{\chi}^{(n)}$ the susceptibility tensor describing the material and its resonances for the n-th ($n \geq 2$) order optical process. In the following we employ explicit tensor notation due to the importance of anisotropy and symmetry considerations in studying the nonlinear response. The electric field $\boldsymbol{E}$ in this description is the local electric field experienced by atoms in the medium. The local field can be modified from the incident driving field due to the polarization of the medium itself. This local field correction and its importance for plasmonic antennas is discussed further below.

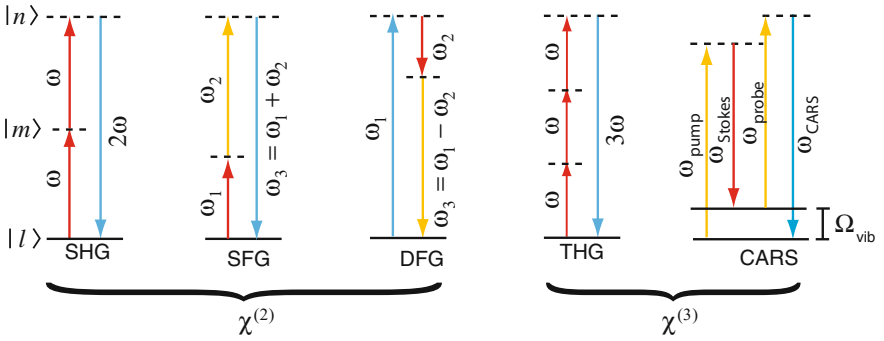

Fig. 7.8 Summary of common nonlinear optical processes with corresponding energy level diagrams. Second-harmonic generation (SHG), sum-frequency generation (SFG), and difference-frequency generation (DFG) are second-order processes. Third-harmonic generation (THG) is a third order process, in which three fundamental photons combine to produce a 3ω photon. Coherent anti-Stokes Raman scattering (CARS) is a resonant four-wave mixing process. The *dashed lines* represent the off-resonant excitation of a real state $|m\rangle$, $|n\rangle$, at a different energy

Just as in the linear case, the induced optical polarization can equivalently be described in terms of a current, but this approach is often less practical because there are typically several nonlinear source terms that may be difficult to separate in this treatment. Both basic harmonic generation wavemixing and more complex frequency conversion processes follow from Eq. 7.26, with relative efficiencies depending on the spectral and symmetry characteristics of the linear and nonlinear susceptibilities of the material.

Several representative nonlinear optical processes are summarized in Fig. 7.8. Metals typically have bulk *inversion symmetry* and therefore a vanishing $\boldsymbol{\chi}^{(2)}$, so that all even-order nonlinear responses in the bulk will vanish in the so-called *dipole approximation*, which neglects weaker higher-order, non-local contributions to the nonlinear response such as magnetic dipole and electric quarupole terms. The second-order nonlinear processes in metals are therefore dominated by the optical-surface interaction.[9] Second-harmonic generation (SHG) is the simplest second-order nonlinear process, where two photons with frequency ω combine to produce a single photon at 2ω. The material response is described by the nonlinear susceptibility $\boldsymbol{\chi}^{(2)}(-2\omega;\omega,\omega)$, which is a third rank tensor with symmetry reflecting the crystal symmetry and dependent on all frequencies $(\omega, 2\omega)$ in the nonlinear process. More generally, the second-order induced polarization can radiate at any frequency which is a linear combination of the frequencies of the incident waves (see Fig. 7.9a), allowing sum-frequency generation (SFG) with the energy conservation condition $\hbar\omega_3 = \hbar\omega_1 + \hbar\omega_2$, difference-frequency generation corresponding to

[9] The term surface is defined here with respect to the actual atomic layer surface boundary, extending over a region of only a few atomic layers in the surface normal direction, where electronic structure is distinct from translationally invariant bulk and possibly modified by surface electronic states.

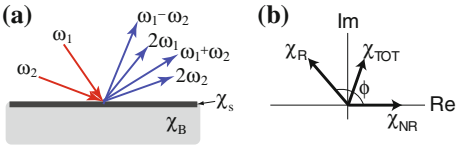

Fig. 7.9 **a** Schematic of interactions with two input fields, ω_1 and ω_2, producing output fields with different frequency, intensity, and emission direction for a planar surface. Both surface χ_s and bulk χ_B induced polarizations occur, with the nonlinear laws of reflection and refraction governing the momentum conservation of the in-plane wavevector. **b** Resonant (χ_R) and non-resonant (χ_{NR}) contributions to the second-order polarization in the complex plane and the sum χ_{TOT}, which may produce interference and asymmetric lineshapes depending on relative phase ϕ

$\hbar\omega_3 = \mid \hbar\omega_1 - \hbar\omega_2 \mid$, and the degenerate case ($\omega = \omega_1 = \omega_2$) of optical rectification giving rise to a DC field with the condition $0 = \hbar\omega - \hbar\omega$.

Third-harmonic generation (THG) produces a 3ω photon from three incident photons with frequencies ω, with a rank four susceptibility tensor $\boldsymbol{\chi}^{(3)}(-3\omega; \omega, \omega, \omega)$. The general third-order process of four wave mixing (FWM) is based on interactions of three photons, with frequencies ω_1, ω_2, and ω_3 combining to produce an output photon with frequency ω_4, with $\boldsymbol{\chi}^{(3)}(-\omega_4; \pm\omega_1, \pm\omega_2, \pm\omega_3)$. Since these are odd-order processes, they are permitted for all materials, including those with centrosymmetric point groups. The nonlinear Kerr effect is also a third-order process, but one with degenerate input and output frequencies, described by the susceptibility $\boldsymbol{\chi}^{(3)}(-\omega; \omega, \omega, -\omega)$. Here the negative sign indicates that the process involves the annihilation of a photon, instead of the simple additive combination seen in harmonic generation. This process is based on a change in the index of refraction and absorption of a material proportional to the incident intensity. Another type of four wave mixing is Coherent anti-Stokes Raman Scattering (CARS), a resonant third order interaction with $\omega_{CARS} = \omega_{pump} + \omega_{probe} - \omega_{Stokes}$ and $\boldsymbol{\chi}^{(3)}(-\omega_{CARS}; \omega_{pump}, \omega_{probe}, -\omega_{Stokes})$. Usually the pump and probe frequencies are identical, and ω_{Stokes} is typically chosen so the difference between the frequencies is resonant with a vibrational level of the material $\Omega_{vib} = \omega_{pump} - \omega_{Stokes}$. This is the coherent analog to incoherent Raman scattering, and as a vibrational spectroscopy technique provides chemical specificity.

The efficient generation of coherent nonlinear optical signals requires both energy conservation and *phase-matching* conditions, that is, momentum conservation between the nonlinear and fundamental k-vectors. In the bulk, this is achieved through the linear dispersion and associated wavelength-dependence of the index of refraction $n(\omega)$. At the interface, it arises from the selection of the input and output k-vector directions. For rough structures or particles on the order of or smaller than λ, the loss of translational invariance leads to changes in the momentum conservation conditions, giving rise to nonlinear light scattering and in certain situations allowing for, e.g. separation of non-local bulk and local surface susceptibilities, as discussed further below.

These nonlinear responses provide access to conduction electrons throughout the energy continuum, which allows probing of interband and intraband transitions. The

enhancement provided by both intrinsic material and extrinsic structural resonances can also lead to a significant increase in the efficiency of nonlinear processes. A discussion of plasmon-resonant metallic systems, where the nonlinear enhancement scales with the order of the process, can be found in Sect. 7.3.4. The coherent nature of the wavemixing processes leads to a strong dependence on the phase of the driving field and material response, which provides additional information for characterization and control. This is particularly useful for the study of the ultrafast dynamics in complex metallic nanostructures, where nonlinear techniques provide more degrees of freedom to probe multiple resonances and their coupling than linear optics. The multiple driving fields in nonlinear optics also enable probing of changes in the complex dielectric function and therefore propagation characteristics of SPPs under strong-pump illumination, an important consideration for active plasmonics.

7.3.1 Second-Order Nonlinear Optics

Here we will provide a more detailed discussion of the origins and theory associated with the lowest, second-order nonlinear response. Although the symmetry considerations associated with even-order responses are different from odd-order nonlinearities, much of what follows can be readily extended to third-order and higher nonlinear processes.

Using Einstein summation notation, the second-order optical response can be written as

$$P_i^{(2)}(\omega_1+\omega_2) = \varepsilon_0 \chi_{ijk}^{(2)}(-\omega_1-\omega_2;\omega_1,\omega_2)E_j(\omega_1)E_k(\omega_2) \tag{7.27}$$

with i, j, k denoting the Cartesian coordinates x, y, z. Within the classical theory of nonlinear optics, an expression for the nonlinear susceptibility can be derived from *perturbation theory* based on a driven, damped harmonic oscillator, analogous to the linear case, with the addition of a quadratic term as a first order perturbation. This approximation results in Lorentzian resonances at the fundamental and wavemixing frequencies,

$$\chi_{ijk}^{(2)}(-\omega_1-\omega_2;\omega_1,\omega_2) \tag{7.28}$$

$$= \frac{Ne^3a}{\varepsilon_0 m^2}\frac{1}{(\omega_0^2-(\omega_1+\omega_2)^2-2i(\omega_1+\omega_2)\Gamma)}\frac{1}{(\omega_0^2-\omega_1^2-2i\omega_1\Gamma)(\omega_0^2-\omega_2^2-2i\omega_2\Gamma)}$$

$$= \frac{\varepsilon_0^2 ma}{N^2e^3}\chi^{(1)}(\omega_1+\omega_2)\chi^{(1)}(\omega_1)\chi^{(1)}(\omega_2) \tag{7.29}$$

with resonance frequency ω_0 (for a single oscillator), nonlinear parameter a and number density of atoms N. This simple model provides an intuitive description of the optical nonlinearity, for the case of weak absorption in the material.

In the dipole approximation of the quantum mechanical description, by considering symmetry operations, the second-order susceptibility tensor can be written as a sum of terms of the form [21, 22]

$$\chi^{(2)}_{ijk}(-(\omega_1+\omega_2);\omega_1,\omega_2) = \frac{Ne^3}{2\varepsilon_0\hbar^2}\sum_{lmn}\rho_l\Big[\frac{\langle l|r_i|n\rangle\langle n|r_j|m\rangle\langle m|r_k|l\rangle}{(\Omega_{nl}+\omega_1+\omega_2)(\Omega_{ml}-\omega_2)}+ \frac{\langle l|r_i|m\rangle\langle m|r_k|n\rangle\langle n|r_j|l\rangle}{(\Omega_{nl}+\omega_1+\omega_2)(\Omega_{ml}-\omega_1)}\cdots\Big] \tag{7.30}$$

This expression describes transitions from state $|l\rangle$ (not necessarily the ground state), through two intermediate states $|m\rangle$ and $|n\rangle$, followed by the emission of a photon with the remaining net energy difference, e.g. $\hbar(\omega_1+\omega_2)$ when returning to the initial state. ρ_l is the population of the initial state, $\langle l|r_i|n\rangle$ is the transition dipole moment operator in the density matrix formalism, and $\hbar\Omega_{nl}$ is the energy difference for this transition. For driving fields with frequencies far off-resonance, all components of $\boldsymbol{\chi}^{(n)}$ are real and additive, corresponding to almost-instantaneous transitions involving "virtual" energy levels, as shown in Fig. 7.8. Close to resonance, $\Omega_{nl} = \omega_{nl} + i\Gamma_{nl}$, with Γ_{nl} describing the line width of the transition, arising from damping. Therefore, $\boldsymbol{\chi}^{(n)}$ is generally complex, with *resonant* (R) and *nonresonant* (NR) contributions to the nonlinear response,

$$\boldsymbol{\chi}^{(n)} = \boldsymbol{\chi}^{(n)}_{\mathrm{R}} + \boldsymbol{\chi}^{(n)}_{\mathrm{NR}} \tag{7.31}$$

As shown in Fig. 7.9b, the SHG signal then arises from the sum of these complex contributions. Since $\boldsymbol{\chi}^{(n)}_{\mathrm{R}}$ will have a strong frequency dependence, the interference of the two terms will produce dispersive lineshapes and even destructive interference depending on the relative phase.[10] The resonances that lead to this behavior can involve single or multiphoton processes, with different degrees of coupling [23].

Figure 7.10 shows possible resonant SHG interactions within the Au band structure, with a plasmon resonant process from the Fermi level and an electronic resonance from the *d*-band.The mixing of the two fundamental ω photons is essentially an instantaneous process if the intermediate $|1\rangle$ state is a virtual energy level, as shown for the 2ω electronic resonance. If the intermediate state is resonant with an eigenfrequency of the material, e.g. in the form of an extrinsic SPP resonance, it has a finite lifetime and the SHG process can accordingly involve fundamental pulses separated by a time interval, denoted τ.

[10] These asymmetric lineshapes resemble those observed in the case of the quantum interaction of two competing pathways connecting discrete and continuous energy levels, called Fano resonances. However, since the interference of the different nonlinear contributions does not arise from quantum interference, but rather from the classical interference of different linear and nonlinear, and resonant and non-resonant polarizations, the use of the Fano lineshape terminology for describing asymmetric linear or nonlinear lineshapes may only be seen as an analogy.

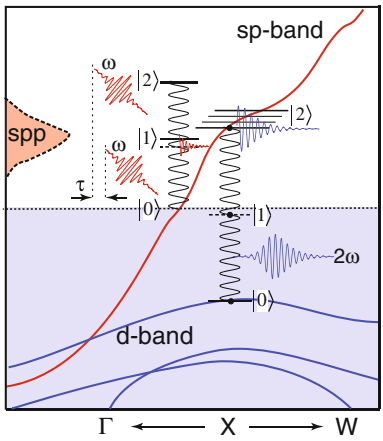

Fig. 7.10 Schematic representation of second-harmonic generation (SHG) in a three level system, superimposed on a band structure diagram for Au. Enhancement of the SHG response can occur when either ω or 2ω corresponds to an eigenfrequency of the material. As an example, resonant excitation can occur via an intermediate extrinsic SPP resonance at the 800 nm fundamental frequency, or a two-photon $sp-d$ resonance involving the d-bands

7.3.2 SHG Response at Metal Surfaces

Since the discovery of nonlinear optics, the nonlinear response of metals has received substantial attention. However, establishing and accurately modeling the microscopic signal sources has been difficult. The nonlinear response does not follow simply from the linear case and includes several contributions to the nonlinear polarization, which are typically hard to separate. These contributions, and the sensitivity of SHG to surface modifications, hampered efforts to quantify the magnitude of $\boldsymbol{\chi}^{(2)}$ in metals.

For centrosymmetric crystals, the lowest order, bulk dipole response is forbidden, since $\chi^{(2)} \equiv 0$ is the only solution to satisfy the inversion operation. The second-order response therefore originates from surfaces and interfaces where symmetry is broken in the sample normal direction, and higher order bulk contributions. The higher-order terms arise primarily from magnetic dipole and electric quadrupole interactions. They are usually small compared to the dipolar response, yet as a bulk response might overall be comparable to a pure surface dipole response. For a cubic crystal, the bulk polarization from these sources can be expressed as an isotropic and anisotropic term,

$$\boldsymbol{P}_B(2\omega) \propto \gamma(\omega)\nabla(\boldsymbol{E}\cdot\boldsymbol{E}) + \xi(\omega)\boldsymbol{E}\nabla\boldsymbol{E}. \tag{7.32}$$

In the free-electron model the anisotropic second term above is zero, but can appear if lattice effects are taken into account. The first description of SHG in metals therefore treated SHG as generated by the isotropic bulk term within the skin depth of the metal [24, 25]. However, this treatment neglects the broken inversion symmetry at the metal-dielectric interface, which leads to the additional dipole-allowed surface SHG.

The surface polarization arises due to the rapid change of the electric field at the metal-air interface, which produces a surface second-order polarization perpendicular and parallel with respect to the surface. The normal component of the electric field

at the surface varies over approximately the Thomas-Fermi screening length, which leads to the spatial confinement of the induced nonlinear current to a subnanometer region. Therefore, a classical electromagnetic description to model surface SHG fails and a quantum mechanical treatment is necessary in order to accurately incorporate the surface charge density and screening effects.

The nonlinear surface polarization is described by susceptibility tensor components, with $\chi^{(2)}_{zzz,s}$ describing the surface normal current, which is expected to be the largest contribution to surface SHG as it is the most sensitive to the structural and electric field change across the interface. The other components for the in-plane surface current are $\chi^{(2)}_{xxz,s}$, or equivalently $\chi^{(2)}_{yyz,s}$, $\chi^{(2)}_{xzx,s}$, etc., and $\chi^{(2)}_{zxx,s} = \chi^{(2)}_{zyy,s}$, due to the symmetries of the tensor.

Rudnick and Stern [26] parametrized three contributions to SHG in terms of the phenomenological constants $a(\omega) \propto \chi^{(2)}_{\perp,s}$, $b(\omega) \propto \chi^{(2)}_{\parallel,s}$, and $d(\omega) \propto \chi^{(2)}_{B}$. In the Drude model, $b(\omega) = -1, d(\omega) = 1$, and $a(\omega)$ was initially assumed to be close to 1. Several models for calculating the spatial distribution of the electron density close to the surface and deriving $a(\omega)$ were developed, chiefly using hydrodynamic arguments to derive the surface potential within a jellium framework, which treats the metal surface as a homogeneous gas of interacting free electrons in a background of uniform positive charge. These models provided an intuitive description of the system, but underestimated the magnitude of the SHG by an order of magnitude [27]. Subsequent models used density functional theory to describe the electron–electron interactions at the surface, which incorporates the screening of the external electric field [28]. While these models typically agree qualitatively with experimental observations, particularly in the long wavelength limit, other effects can also become significant and change the relative contributions of the different polarization terms. Additional susceptibility components may also appear when the lattice is considered. For example, close to resonances the bound electrons may contribute more strongly to the nonlinear polarization, producing a bulk response larger than the surface, even in centrosymmetric materials. Consistent with this, a strong enhancement in SHG has been observed in noble metals close to the interband transition, in addition to the usual off-resonant nonlinear signal [29, 30].

For noble metals, the k-dependence of the electronic structure is typically neglected in modeling the SHG response. The high density of states and overlapping d-bands allow for a continuum of transitions with different symmetries, as shown in Fig. 7.10, producing broad SHG peaks. When the excitation frequency is such that the band gap E_g is less than 2ω, the SHG response is generally dominated by transitions where both the initial and the intermediate states are in the d-band. This sensitivity of SHG to the d-band can provide spectroscopic material specificity.

The literature disagrees on quantitative measurements of the magnitude of the SHG signal and its components, due to the high sensitivity of SHG to surface structure and contamination. In particular for Ag and Al, accurate measurements require ultra high vacuum to ensure clean surfaces. The second-harmonic responses from Ag(111) and Au(111) surfaces, far off-resonant at $\hbar\omega = 0.81$ eV, were found to be dominated by the surface normal susceptibility, as expected [31]. Significant contributions were

also found for other susceptibility components, with their relative magnitude varying with fundamental frequency. In this early work, the nonlinear response of Au was measured to be $\chi^{(2)}_{s\perp\perp\perp} \sim 2 \times 10^{-8}$ m/V, approximately four times larger than that of Ag, close to 1 eV. Later measurements on thin films found that Ag has the strongest SHG intensity of the metals, with Au slightly smaller and Cu approximately 50 % of the Ag response [32]. With careful angle and polarization-dependent measurement enabling separation of the bulk and surface responses in the experiment, the $\chi^{(2)}_{s\perp\perp\perp}$ component was determined to be approximately 200 times larger than $\chi^{(2)}_B$ for Au, and ∼100 $\chi^{(2)}_B$ in Ag, at a fundamental photon energy of 1.55 eV. Interestingly, Al has an inherently high bulk nonlinear response, with $\chi^{(2)}_B$ values an order of magnitude higher than Au and Ag. However, it also suffers from a short skin depth and high losses in the visible and NIR, in addition to a tendency to oxidize, so is generally not considered as suitable for plasmonic applications.These experiments also demonstrated a wide variation in surface susceptibility values depending on surface roughness and growth conditions of the thin films, an effect that is the subject of Sect. 7.3.4. Hence, the development of both accurate quantitative experiments and an accurate quantitative theory has remained difficult.

While the discussion to this point has been limited to the nonlinear response of metal involving single particle excitations, the next section is concerned with nonlinear interactions involving SPPs, where the surface-sensitivity of the second-order response becomes particularly important for the enhancement of the nonlinear signal.

7.3.3 Nonlinear Wavemixing with Surface Plasmons

The momentum mismatch between the incident and emitted light and in plane SPP wavevectors as shown in Fig. 7.4 means that linear excitation of SPPs on planar surfaces typically requires an effective momentum change of the incident field in the form of a coupling element or increase in index of refraction, for example a grating or Kretschmann prism, respectively. However, the phase-matching conditions associated with the different wavevectors participating in a nonlinear wavemixing process provide additional flexibility, allowing free-space launching of SPPs via the generated nonlinear polarization, or SPPs as the source for the nonlinear output field, or both. While even-order processes are intrinsically surface-confined, interactions involving SPPs are also in practice limited to the near-surface region given by the skin-depth, even for odd-order processes. Any of the participating wavemixing fields can be an SPP, with a wavevector that is defined by the surface dispersion relation of the specific SPP frequency. Figure 7.11 shows a set of different possible configurations. The surface-parallel components of the free-space k-vectors $k_i = \omega_i n(\omega_i) \sin\theta_i / c$, where θ_i is the angle with respect to the surface normal and $n(\omega_i)$ is the index of refraction of the dielectric medium, can then be summed appropriately with the

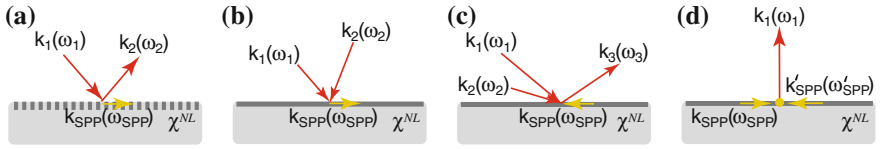

Fig. 7.11 Schematics for several examples of wavemixing processes involving SPPs. Launching of SPPs through the use of a grating or a modification of the index of refraction can lead to enhancement in the nonlinear response (**a**), with surface-parallel momentum conservation condition $k_1(\omega_1) + \frac{2\pi n}{a_0} = k_2(\omega_2) + k_{\mathrm{SPP}}(\omega_{\mathrm{SPP}})$, for integer n and grating period a_0. SPPs can also be generated through appropriate phase matching conditions between several input waves (**b**), e.g., $k_1(\omega_1) - k_2(\omega_2) = k_{\mathrm{SPP}}(\omega_{\mathrm{SPP}})$ for DFG. One or more of the free-space waves in a wavemixing process can also be substituted by an SPP, e.g. $k_1(\omega_1) + k_2(\omega_2) - k_{\mathrm{SPP}}(\omega_{\mathrm{SPP}}) = k_3(\omega_3)$ for FWM (**c**), or $k_{\mathrm{SPP}}(\omega_{\mathrm{SPP}}) - k_{\mathrm{SPP}}(\omega_{\mathrm{SPP}}) = k_{1,\parallel} = 0$ (**d**)

SPP wavevectors to achieve energy and momentum conservation for the desired wavemixing process.

Nonlinear SPP wavemixing can also provide enhanced efficiency of the nonlinear response. The field amplitudes of SPP modes that drive the wavemixing process are enhanced near the surface due to the spatial field confinement, thus enhancing the nonlinear polarization generated. An example of this process is the enhancement of SHG observed when an SPP is excited through prism coupling onto a silver film in the Kretschmann geometry [33]. Similar effects have been seen in third-harmonic generation (THG) with total internal reflection [34]. The plasmon-enhanced nonlinear response can interfere with other sources of nonlinear polarization in the system. Due to their different phase relationship with the driving field, this interference will also depend on the incident k-vector.

Early in the development of nonlinear optics, four-wave mixing (FWM) was proposed as a mechanism for launching surface waves such as exciton polaritons, phonon polaritons [35] or SPPs [36], by tuning the angle of illumination to achieve wavevector matching at the sample-air interface. The efficiency of this approach is determined by the local field enhancement and nonlinearity of the metal. While the nonlinearity of metals is high in general, the interaction volume is limited by the skin depth. This leads to a low efficiency in generating SPPs by wavemixing, compared to direct excitation of an SPP of the corresponding frequency. Another approach to achieve SPP coupling via a nonlinear process is create a transient temperature grating by interfering two incident waves on the surface. This is a incoherent pump induced, rather than a coherent wavemixing process. The resulting thermal gradient gives rise to a spatial variation in the index of refraction, and thus allows for launching SPPs. This process has a much higher efficiency than FWM with femtosecond pulses, but a much long timescale, given by thermal diffusion [37]. In order to maintain the ultrafast timescale of wavemixing, higher efficiencies could be possible with a second-order process such as DFG [38] rather than FWM. These and other combinations of free space and propagating SPP waves (examples shown in Fig. 7.11) have

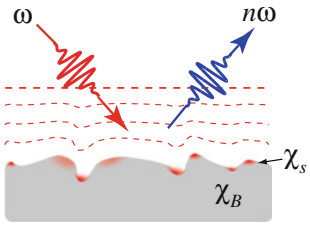

Fig. 7.12 Areas of local field enhancement on rough metallic surfaces lead to large enhancement in both the local and the overall nonlinear response, here for the example of harmonic generation $n\omega$, with $n = 2, 3$ etc

been considered for second and third order wavemixing SPPs, and recently received renewed attention [39].

SPPs can also act as one or more of the driving fields in a nonlinear optical process [40–42] (Fig. 7.11b, c). For the right conditions, SPPs contribute in a phase-matched fashion to the wavemixing process, e.g. SHG generation from two SPP fields [43].

7.3.4 Surface-Enhanced Nonlinear Processes

The sensitivity to symmetry-breaking of even-order nonlinear processes makes them an effective tool for the study of, e.g., surface electronic and vibrational resonances and their coupling. However, the nonlinear response is weak in general, and further limited by the small volume of surface material involved in the nonlinear interaction. Enhancement can arise from the localization and concentration of the optical fields near a surface or at a nanostructure. Localized plasmon resonances in noble metal nanoparticles, clusters, and rough metal surfaces can provide a further increase in nonlinear optical effects, and substantially change the relative bulk to surface contributions in a nonlinear response.[11] These "hot spots" provide enhancement in linear optical processes as well, but with regard to an aggregate bulk response are reduced, since energy conservation conditions require that enhancement of the field is balanced by lower local fields and thus reduced optical response in other regions. In nonlinear processes, in contrast, in one sample location the total signal enhancement can be much higher due to the nonlinear dependence of the response on the local optical field. The breaking of translational symmetry and spatial redistribution of the optical field is therefore beneficial to the higher order response (See Fig. 7.12).

The enhancement of an optical response is described phenomenologically in terms of a local field enhancement factor $L(\omega)$, which modifies the driving electric field, analogous to the Fresnel factors for planar interfaces in reflection or the bulk local field correction factor discussed earlier, as

[11] In random, fractal, or percolated media, a mixed mode between localized and propagating SPPs is possible. The interference of this collective mode of the local excitation and multiple scattering in the disordered media can give rise to Anderson localization for typically uncorrelated disorder with associated nonlinear optical effects [44].

$$E_{\mathrm{loc}}(\omega) = L(\omega)E(\omega). \tag{7.33}$$

A local field factor needs to be considered for all optical fields contributing to the nonlinear process, so that the total enhancement is the combination of all enhancement factors incorporating the order and coherence of the nonlinear process.[12]

Raman scattering is an incoherent, linear optical process, but the enhancement in the field is approximately proportional to $L^2(\omega)$ since the fundamental and Stokes shifted Raman signal have only a small frequency separation compared to the typical spectral variation of $L(\omega)$ for the supporting metal, and will both be enhanced.[13] This effect has been exploited for *surface-enhanced Raman scattering* (SERS), where the increase in the effective cross section by a rough metal film can provide single-molecule sensitivity [46], and is also the basis of *tip-enhanced Raman scattering* (TERS). In both cases, the additional sensitivity arises from the redistribution and near-field localization of the field in the surface normal direction. A higher-order, coherent process such as SHG also benefits from the *lateral* redistribution of the field, with areas of high field enhancement increasing the signal nonlinearly. For SHG, the enhancement in the polarization is given by

$$\boldsymbol{P}(2\omega) = L(2\omega)\boldsymbol{\chi}^{(2)}(-2\omega;\omega,\omega)L^2(\omega)\boldsymbol{E}^2(\omega) \tag{7.34}$$

where $L(\omega)$ and $L(2\omega)$ are the local field factors at the fundamental and SHG frequencies respectively. The total intensity enhancement is then $\propto L^2(2\omega)L^4(\omega)$. Both fields in this case might not simultaneously be enhanced due to their spectral separation, in which case either $L(\omega)$ or $L(2\omega)$ is typically approximately equal to 1. The same arguments apply to higher-harmonic generation processes, with the general field enhancement behavior

$$\boldsymbol{P}(n\omega) = L(n\omega)\boldsymbol{\chi}^{(n)}(-n\omega;\omega,\omega,...)L^n(\omega)\boldsymbol{E}^n(\omega). \tag{7.35}$$

In a spatially distributed nanoparticle system, the regions of highest local field enhancement for different wavemixing processes can be in different locations, depending on the resonant frequency and mode behavior. Degenerate four wave mixing in general displays higher enhancement on rough surfaces than third harmonic generation, due to more than one driving laser field being enhanced simultaneously. It has also been observed that harmonic generation tends to show lower enhancement than incoherent processes such as the nonlinear Kerr effect, since the coherence of the process can produce destructive interference in random metallic systems [44].

---

[12] Equivalently, the enhancement can be incorporated into a modification of the susceptibility tensor, but this description may be less intuitive for the case of, for example, surface-enhanced Raman scattering, where the susceptibility tensor is not the intrinsic metallic system but rather a coupled metal-molecule system.

[13] Because of symmetry considerations arising from the Raman tensor, this coupling of the incident and radiative fields is not rigorously accurate. In reality, the relative orientation of the local field and the molecular dipole or crystallographic orientation can lead to more complex enhancement behavior. For more details, see, e.g., Ref. [45].

The general principles of surface- and plasmon-enhanced nonlinear optical effects are well understood. However, details in terms of the relative surface and bulk modifications to the susceptibility, the interband and intraband transitions, finite size effects on band structure, plasmon mediated effects in the nanostructure, and interactions with the substrate are not yet well understood. Furthermore, the spectral dependence and magnitude of the field enhancement varies critically depending on the surface morphology, which is difficult to model. Grating structures can be useful for the treatment of surface-enhanced nonlinear processes, since they provide a model system for rough surfaces. Experimentally, the enhancement of nonlinear optical effects has been demonstrated on various samples, with roughness controlled to a certain extent through film thickness and growth conditions [47], and SHG enhancement of 10^4 on a roughened Ag surface was observed early on [48], in addition to surface-enhanced higher order processes [49]. However, just as for a planar geometry, the surface and bulk contributions to the SHG are difficult to separate, and both may be modified by the roughness [50].

7.3.5 Nonlinear Light Scattering

One of the complications of surface-enhanced nonlinear interactions is that the roughness can lead to extrinsic dephasing and depolarization. Similarly, nonlinear processes in particle systems, e.g. in gas and liquid phase, where there is a substantial spatial inhomogeneity in local fields and nonlinear susceptibilities, will be accompanied by scattering. For the small particle limit, where the particles can be treated as dipole sources (i.e. 5–10 nm for visible light), nonlinear light scattering is known as hyper-Rayleigh scattering, in analogy to Rayleigh scattering [11]. Some confusion in terminology exists in the literature, but according to the strict definition, the nonlinear response in particles larger than 10 nm arises from coherent effects, even when the contributions from the particles add incoherently, and so hyper-Rayleigh scattering can be a misleading term [51].

The change in momentum conservation rules in scattering processes compared to bulk media produces new and additional symmetry selection rules, which are described in the context of nonlinear Mie and Rayleigh scattering with an effective surface susceptibility $\chi_s^{(n)}$ [52–54]. In particular, the lack of translational invariance and $k \propto 1/r$ for a single nanoscopic system lifts the phase matching condition, so that the projection of the nonlinear k-vectors to the far field is not restricted to a particular direction. Of the different nonlinear interactions in individual nanoparticles, third-order processes such as THG behave similarly to linear scattering, while second-order processes such as SHG have additional sensitivity to the particle surface and geometric details. Furthermore, the susceptibilities for nanoparticles can be very strongly affected by grain size and crystallinity. Analogous to the linear case, if the particles are large enough to allow for retardation effects over the particle diameter, higher order multipolar contributions to the nonlinear polarization occur.

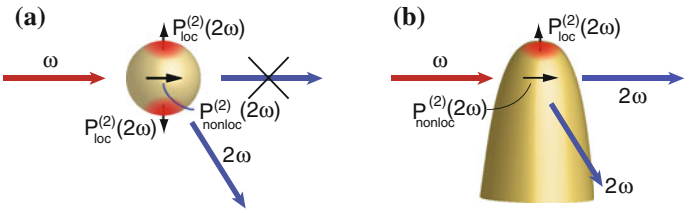

Fig. 7.13 Symmetry considerations for a sphere of centrosymmetric material (**a**) and a conical tip (**b**). No second-order response appears in the exact forward or backscattering directions for the sphere since the surface contributions are out of phase and interfere destructively, but a non-local response can produce SHG in other directions. For the conical tip, the broken symmetry along the tip axis allows both local dipolar forward- and back-scattering and non-local scattering

The case of SHG from a *spherical nanoparticle* of a centrosymmetric material is particularly interesting for reasons of symmetry of the second-order nonlinear response. Inversion symmetry is broken at the surface, but the usual linear dipole mode aligned in the direction of the pump polarization, as is responsible for linear Rayleigh scattering, will not produce SHG as the surface contributions are 180° out of phase and thus cancel (Fig. 7.13a. Instead, a non-local nonlinear polarization in the direction of the pump wavevector arises due to retardation in the phase across the particle diameter [55], in addition to a possible higher-order bulk response. The orientation of the dipole and bulk quadrupole sources is such that no SHG will radiate in the exact forward and backward directions, but radiates in non-collinear directions with spatial distribution determined by particle size.

A *conical tip*, such as those used in near-field optical experiments, while semi-infinite, can also be considered within the context of nanoscale particles. It possesses broken mirror symmetry along the tip axis (∞mm point group symmetry). This leads to fully *local dipole-allowed* SHG polarization $P^{(2)}_{\mathrm{loc}}(2\omega)$ along the tip axis [53]. This symmetry breaking produces different polarization selection rules for SHG in nanoscopic metal tips than for surfaces or spherical particles. In particular it is possible to distinguish the *non-local* bulk $P^{(2)}_{\mathrm{nonloc}}(2\omega)$ and *local* surface $P^{(2)}_{\mathrm{loc}}(2\omega)$ SHG response, since these two contributions are perpendicular (Fig. 7.14) and produce correspondingly cross-polarized SHG. As discussed above, this separation of local and non-local SHG contributions is typically difficult for planar surfaces due to nonlinear laws of reflection, which limit emission to the direction defined by the incident k-vector direction. The conical geometry therefore provides a model system for characterizing nonlinear enhancement and scattering effects, since it is a single element structure with well-defined symmetry and permits the separation of different SHG responses.

The symmetry-breaking behavior of a Au conical tip with apex radius ~20 nm is demonstrated in Fig. 7.14, for sagittal illumination of the tip exciting a local, purely dipolar surface nonlinear polarization $P^{(2)}_{\mathrm{loc}}(2\omega)$ oriented along the tip axis, leading to radiation of SHG in the forward direction. In addition, the non-local source perpendicular to the tip axis can radiate in the 90° direction. This arises from retarda-

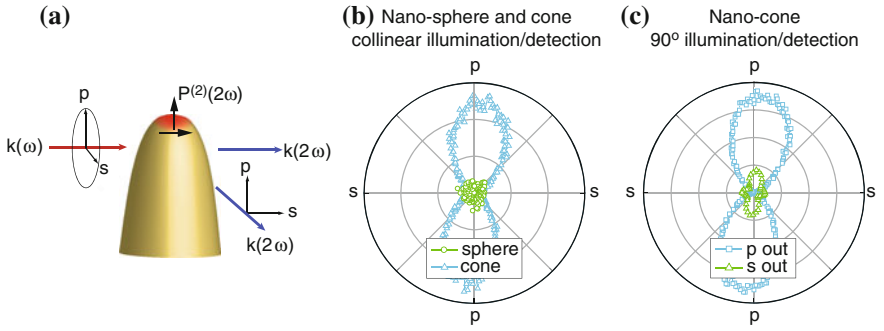

Fig. 7.14 Geometry of SHG scattering from a conical nano-tip (**a**), with forward scattering allowed for sagittal *p*-polarized illumination due to the broken symmetry along the tip axis (**b**). For 90° detection both a bulk non-local and dipolar response are possible (**c**). Input polarization dependence for collinear SHG (**b**) from a nano-sphere reference (*green*) and nano-cone (*blue*). No SHG is observed in this geometry for the sphere, but the cone demonstrates the expected dipolar response. SHG in 90° sagittal illumination/detection geometry, for *p* and *s* polarized output (**c**), demonstrating the separation of dipolar surface and bulk response

tion from spatially-distributed surface nonlinear polarizations and higher-order bulk contributions. Experimental results are shown, first with no SHG observed for a nano-sphere in the forward-scattering direction used as a reference (Fig. 7.14b, green). In contrast, for the tip, SHG in this geometry is dominated by the local dipole-allowed $p_{\text{in}} - p_{\text{out}}$ contribution (blue), with the expected two-fold anisotropy, i.e. intensity $I_{\text{SHG}} \propto \cos^4(\theta)$. Similar to a planar surface, the response is due to the strong $\chi^{(2)}_{s,\perp\perp\perp}$ tensor element. For the tip, the weak $s_{\text{in}} - p_{\text{out}}$ response (data not shown) suggests that the $\chi^{(2)}_{s,\perp\parallel\parallel}$ susceptibility component is negligible. With sagittal illumination and 90° detection for a tip (c), both the local dipolar $p_{\text{in}} - p_{\text{out}}$ and non-local (distributed) bulk $p_{\text{in}} - s_{\text{out}}$ and $s_{\text{in}} - s_{\text{out}}$ response appear.

These results provide a demonstration that the additional degrees of freedom that arise from the combination of *intrinsic* material response and *extrinsic* nanoscale geometric properties enables separation of bulk and surface SHG. The SHG properties again depend sensitively on the morphology and local environment of the nanostructure, but the symmetry selection rules derived above are generally applicable to asymmetric nanostructure systems. For example, the presence of a substrate will break symmetry and relax the polarization selection rules for metal particles, and the tip in a near-field optical experiment will have a similar effect on a local scale. With the capability to probe both surface and bulk properties on the nanoscale, the study of plasmonic behavior with high specificity can be achieved.

7.3.6 Nonlinear Optical Antennas

As discussed above, particles and rough surfaces can provide large field enhancements, but with the increasing interest in plasmonic applications such as imaging, sensing, cloaking, or harvesting, a need for controllable and reproducible linear and nonlinear responses has developed. Recent advances in chemical synthetic methods allow the production of crystalline metal nanostructures with a wide range of shapes and sizes with nanometer-scale structural control. Single-crystal nanoparticles and nanowires often exhibit strong plasmonic resonances due to their low defect density and well-defined shape. Additionally, lithographic techniques, focused ion beam milling, and template stripping now provide a means to generate arrays of nanoparticles and other more complex structures such as coupled nanowires and bowtie antennas, which have a large field enhancement in the nanogap region. For an antenna, the SHG polarization is again given by

$$\boldsymbol{P}^{(2)}(2\omega) = \boldsymbol{\chi}^{(2)}(-2\omega;\omega,\omega)L^2(\omega)L(2\omega)\boldsymbol{E}^2, \tag{7.36}$$

where the nonlinear response is given by the material susceptibility. The local field factor L now describes enhancements due to antenna resonances in addition to localized plasmon resonances, and so depends sensitively on geometrical and environmental properties and the coupling of plasmonic modes.

Similar to radio-frequency antennas, antenna resonances for plasmonic antennas such as rods occur when the length of the antenna is equivalent to an integer multiple of half the wavelength. However, the wavelength is modified from the free-space wavelength by the SPP dispersion on the surface of the metal [56]. The precise resonant behavior of the antenna depends sensitively on geometrical details such as diameter, cross-section, and shape, as well as roughness. For such simple antennas, the spectral dependence of the antenna resonances is often approximated by Lorentzian lineshapes,

$$L(\omega) \propto \sum_l \frac{A_l}{\omega_l - \omega + i\Gamma_l}, \tag{7.37}$$

where ω_l are the resonance frequencies, Γ_l is the damping of the antenna, and A_l the relative strength of the resonance.

The local field correction arising from optical antenna resonances can provide an enhanced nonlinear response, up to several orders of magnitude, arising from the enhancement of the linear electric field. However, the spectral dependence of the local field enhancement can lead to a shift in the emission spectrum, as represented in Fig. 7.15, and therefore also an apparent spectral shift in the nonlinear response. The nonlinear response may also not be accurately predicted by the linear far-field response, due to the different near-field spectral density of states distribution compared to the far-field [57].

For non-degenerate wavemixing processes, coupled antennas can be designed such that several input frequency components are simultaneously enhanced [58, 59].

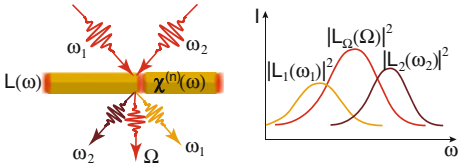

Fig. 7.15 Interaction of light with a rod antenna, showing the spectral shift in emission due to the local field enhancement associated with antenna and plasmon resonances at incident and wavemixing frequencies $L(\omega_1)$, $L(\omega_2)$, and $L(\Omega)$

The localization and concentration provided by optical antennas can also be utilized to couple to highly nonlinear media, such as GaAs, ZnO, or $BaTiO_3$, to generate a strong nonlinear response.

7.4 Femtosecond Time-Domain Measurement of Plasmon Dynamics

In this section we resume the discussion on plasmon dynamics from above (7.2), demonstrating the use of the nonlinear SPP response itself for the determination of the dynamic response underlying a localized SPP excitation, and with that the electron dynamics of the supporting metal. An SPP, as with any optical response, is defined in terms of both amplitude and phase, whether in the spectral or temporal domain. The characterization of SPP dynamics, however, is frequently incomplete, with only amplitude but no phase information obtained, e.g., in incoherent dark field scattering. The underlying dynamics inferred from these spectral measurements therefore rely on model assumptions such as a Lorentzian lineshape from a harmonic oscillator model with flat spectral phase, a transform limited driving laser pulse, or constant relative phase of the response with respect to the non-resonant background. In addition, in frequency space the fast initial dynamics of the plasmon evolution are encoded in the spectral wings, where the signal level is low and thus sensitive to background and noise. Consequently, the spectral wings are very susceptible to possible constructive or destructive interference with the background. Conventional techniques are therefore unsuitable for the study of complex, multi-resonance, or coupled plasmonic systems.

In contrast to incoherent techniques such as dark field scattering, nonlinear optical techniques such as harmonic generation provide access to full amplitude and phase for optical waveform characterization, enabling the direct measurement of plasmon dephasing time and other electron interaction and relaxation behavior. Access to the ultrafast nonlinear response is therefore important for developing an understanding of field enhancement and resonance effects, since a resonance with a plasmon excitation in a system can enhance the linear and nonlinear response, but will also lead to a prolonged dephasing time [60]. Ultrafast nonlinear measurements can ad-

ditionally provide insight into more complicated effects such as the interaction of the various resonance decay channels and interface relaxation time, and the effect of spatial confinement on scattering. Measurements of the temporal dynamics of plasmon resonances in the visible range are challenging, since the few-femtosecond resolution necessary is comparable to the shortest possible pulse duration in the desirable visible to near-IR local SPP resonance energy range. However, achieving the required temporal resolution is not necessarily limited by the shortest pulse available, but rather by the signal to noise ratio and precise characterization of the full optical transient in amplitude and phase.

Plasmonic interactions can be probed with interferometric homodyne or heterodyne techniques or electro-optic sampling to extract the response function defined in Eq. 7.4. These techniques can resolve the ultrashort dynamics of the plasmon response $R(t)$ with exact reconstruction of the response function by deconvolution from autocorrelation and cross-correlation measurements. Spectrally resolved nonlinear techniques can provide the simultaneous phase and amplitude information needed for the unambiguous reconstruction of both the driving field and the resonant polarization transient response, for example through a frequency resolved optical gating (FROG)-based technique [61].

FROG allows the determination of a pulse amplitude and phase through measurement of the self-gated pulse in the time-frequency domain. The most common implementation of FROG is based on the SHG response arising from two pulses interacting in a medium, given by $(X(t) + X(t-\tau))^2$, where $X(t)$ is the field transient of the pulse and τ is the time delay between the two pulses. In a non-collinear implementation, only the cross-term is detected, so that the FROG spectrogram corresponds to a spectrally-resolved intensity autocorrelation, i.e.,

$$S(2\omega, \tau) \propto \left| \int_{-\infty}^{\infty} X(t) X(t-\tau) e^{-i\omega t} dt \right|^2, \tag{7.38}$$

For a nonlinear medium which is far off-resonant, where the response is essentially instantaneous, the field transient $X(t)$ is simply proportional to the electric field of the driving laser pulse $E(t)$, gated by the time-delayed pulse $E(t-\tau)$. For a material close to resonance, the finite response time leads to an induced polarization transient, so that $X(t) = P(t)$, with free-induction decay behavior. From the spectrogram, the full electric field or polarization transient information can be reconstructed using an iterative algorithm. The resonant response function $R(t)$ can be extracted through deconvolution.

The spectrogram can be measured in a collinear geometry, producing additional terms and a spectrally resolved interferometric autocorrelation or IFROG:

$$S(2\omega, \tau) \propto \left| \int_{-\infty}^{\infty} (X(t) + X(t-\tau))^2 e^{-i\omega t} dt \right|^2 \tag{7.39}$$

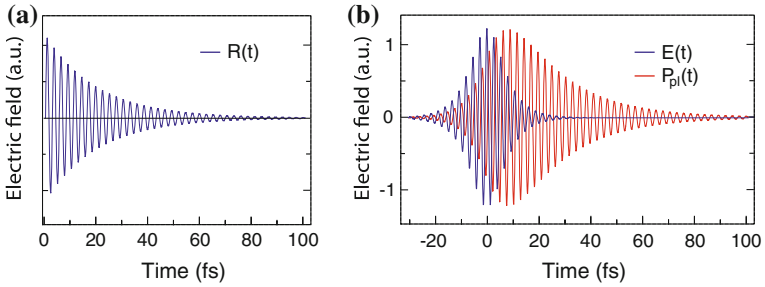

Fig. 7.16 Modeled plasmon response function $R(t)$ (**a**), with dephasing time $T_2 = 20$ fs, and resulting resonant polarization response $P(t)$ subject to driving electric field $E(t)$ (**b**). Incident pulse duration $\tau_0 = 10$ fs, and $\omega_0 = \omega_{\rm pl}$ corresponding to 800 nm wavelength (details in text)

A model of the effects of a finite response function $R(t)$ on the resulting induced polarization $P(t)$ transient in the time domain is shown in Fig. 7.16. The plasmonic response is modeled as a damped harmonic oscillator in the time domain,

$$R(t) = Ae^{i\omega_{\rm pl}t}e^{-\gamma t} \tag{7.40}$$

where A gives the effective oscillator strength, $\omega_{\rm pl}$ is the plasmon resonant frequency, here taken to be resonant with the laser pulse, and the linewidth is given by $\gamma = 1/T_2 = \frac{1}{20}$ fs^{-1}. A sech^2 laser pulse with flat spectral phase is used to simulate the driving field $E(t)$, with full width at half-maximum $\tau_0 = 10$ fs and carrier frequency $\omega_0 = \omega_{\rm pl}$ corresponding to 800 nm center wavelength, i.e.,

$$E_{\rm sim} = E_0 {\rm sech}\frac{1.763t}{\tau_0}e^{i\omega_0 t}. \tag{7.41}$$

The resulting polarization arising from the driving field demonstrates the increased response time from relaxation of the damped harmonic oscillator model of the plasmon resonance, with its free-induction decay lasting past the end of the laser pulse.

A possible experimental geometry for measuring SPP dynamics in the time domain is shown in Fig. 7.17. A high quality, well-aligned parabolic mirror is used as the focusing element in order to minimize dispersion and maintain short pulses and a spatially well-defined Gaussian beam profile. Phase and amplitude of the driving laser pulse are determined using an instantaneously responding reference medium. The BBO acts as the non-resonant medium for pulse characterization, mounted interchangeably with the plasmonic system without further alignment. Results are shown in Fig. 7.18a for BBO, and the corresponding IFROG for a resonant plasmon tip response in b, with the characteristic spectrally narrowed and temporally broadened plasmon excitation [62]. The tip, as a resonant medium, shows spectral narrowing due to the temporal broadening from the finite response function, and a frequency shift in the spectrogram due to the difference between the plasmon resonance and

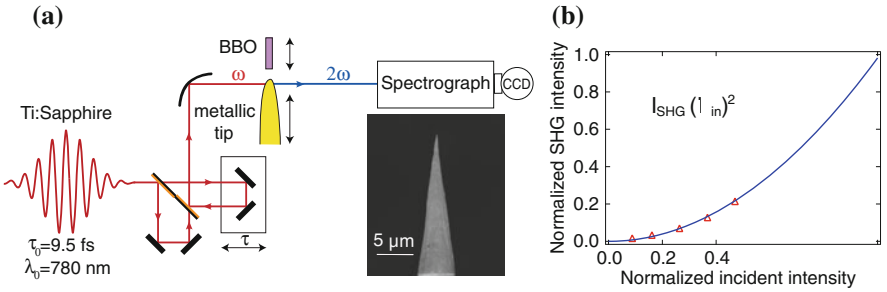

Fig. 7.17 Schematic of the experimental set-up for IFROG characterization of the plasmon dynamics in a single nanoscopic plasmonic structure, here with a plasmonic conical tip as an example of a localized SPP system (**a**). A Mach-Zehnder interferometer with special beamsplitter and parabolic mirror are employed to minimize dispersion and provide diffraction-limited excitation. *Inset* SEM image of a Au plasmonic tip. Power dependence of tip apex SHG, showing the expected quadratic behavior (**b**)

laser carrier frequencies. Figure 7.18c, d show the electric field $E(t)$ and polarization $P(t)$ amplitude and phase, reconstructed from a FROG retrieval algorithm, for the time and frequency domains. Panels (d) and (f) are the plasmon response function from deconvolution of $P(t)$ and $E(t)$, with comparison to the decay for a damped Lorentzian fit as given in Eq. 7.40. From $R(t)$ a dephasing time of 20 ± 5 fs can be directly determined without model assumptions, while for a tip with plasmon frequency not resonant within the bandwidth of the driving field, the tip response is essentially instantaneous (data not shown). The deviations from a flat phase behavior indicate possible inhomogeneities arising from structural imperfections in the nanoscale tip.

The value of T_2 is in agreement with the low-energy limit and energy-independent damping, i.e. $T_2 \sim 2\tau_D \sim 20$ fs for Au, as shown in Table 7.1. This corresponds to the non-radiative limit for decay of the plasmon response in the Drude model. Note that this value was directly extracted from the envelope of the reconstructed plasmon response function (green curve, Fig. 7.18) without any model assumptions.

Other nonlinear processes can also provide the necessary nonlinear polarization and pulse characterization, and may be required since SHG relies on a non-centrosymmetric structure. However, THG for example cannot distinguish bulk, surface, local and non-local effects and is therefore not ideal for extracting the pure plasmonic response. In addition to FROG-based measurements, time-resolved two-photon photoemission can also provide information on plasmonic dephasing. Here a two-pulse cross correlation measurement is used to measure the photoemission current from electrons excited above the vacuum level as a function of pump-probe delay, providing phase information, sub-femtosecond time resolution, and sub-100 nm spatial resolution in combination with photoemission electron microscopy (PEEM). As an early example, Ag nanoparticles on a grating were studied to determine both morphology and dynamics of the nanostructures, with dephasing times as short as 5 fs measured [63].

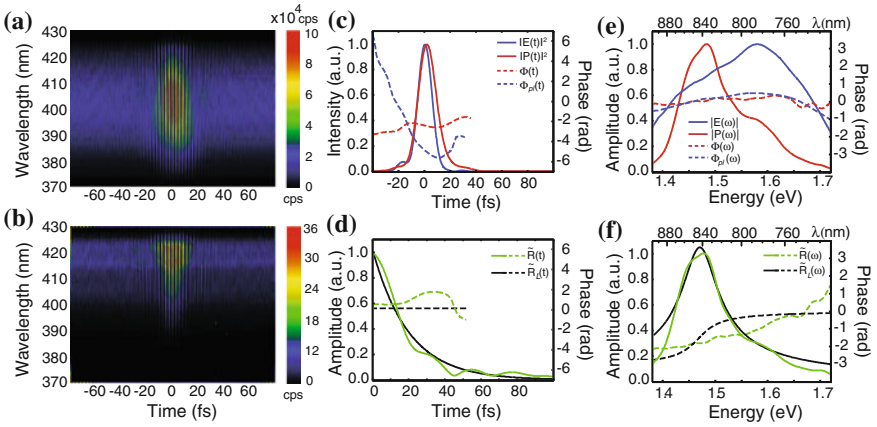

Fig. 7.18 Interferometric SHG FROG measurement of BBO (**a**) and a plasmon resonant Au tip (**b**). Phase (*dashed*) and intensity (*solid line*) of $\boldsymbol{E}(t)$ (*blue*) and $\boldsymbol{P}(t)$ (*red*) (**c**), derived from **a** and **b**. Corresponding Fourier transforms $\boldsymbol{E}(\omega)$ (*blue*) and $\boldsymbol{P}(\omega)$ in frequency domain (**e**). Response function $R(t)$ (**d**) and $R(\omega)$ (**f**) (*green*) from deconvolution of (**c**) and (**e**). Damped harmonic oscillator model response function fit $R_L(t)$ and $R_L(\omega)$ shown in *black*. Reprinted with permission from Ref. [62]. *Copyright 2010 American Chemical Society*

7.5 Ultrafast Spatio-Temporal Control with Plasmonic Antennas

The capability of optical antennas to generate high spatial localization and enhancement of optical fields is also important for characterization of nanoscale materials. While most implementations of optical antennas rely on planar geometries, spatial control of optical fields for nano-imaging can be realized using free-standing conical tip geometries, such as those used in scanning probe applications and discussed in Sect. 7.3.5 as individual nanoscopic nonlinear antennas.

The sensitivity and efficiency of optical antennas can be improved by reducing the *mode-mismatch* between the exciting far-field waveform and the near-field excitation, for example using wedges, grooves, or cascaded structures to achieve a continuous transformation from the micro- to the nano-scale. Taking advantage of the radius-dependent index of refraction experienced by SPP modes on a conical waveguide such as a noble metal tip is one approach for achieving high localization for background-free spectroscopy and imaging [64, 65]. This *adiabatic nanofocusing* approach has the advantage that scattering losses due to structural discontinuities and the decreasing SPP wavelength are minimized until the apex, where a nanoscale optical emitter is efficiently generated.

Furthermore, the nanofocusing mechanism does not rely on a resonant response and therefore is only weakly wavelength and phase dependent, unlike most optical antenna concepts which rely on the spectrally-limited plasmonic response. A broad bandwidth and thus short pulse delivery to the apex is possible. Other approaches to achieve localization rely on the interference of plasmon modes in an arbitrary

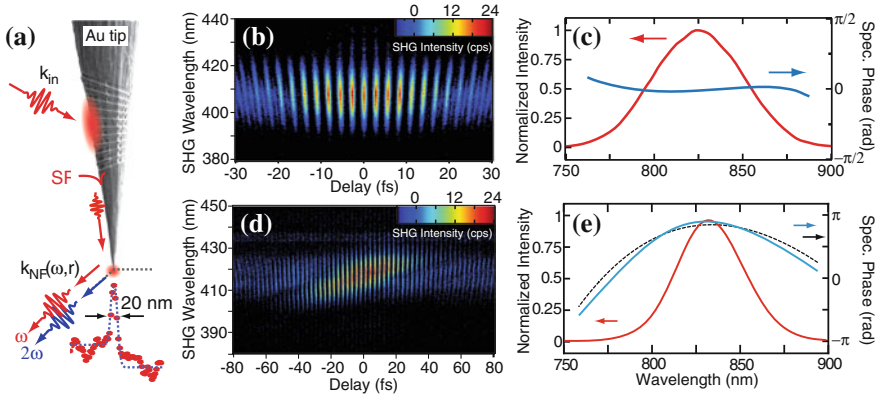

Fig. 7.19 SEM image of Au tip, illustrating grating coupling for SPP launching and nanofocusing to $\sim 20\,\text{nm}$, followed by localized apex emission (**a**). Interferometric spectrogram of apex-emitted SHG radiation (**b**) and corresponding spectral amplitude and phase (**c**) from reconstruction using a FROG algorithm, showing a transform limited pulse with bandwidth corresponding to a 16 fs pulse duration. Demonstration of deterministic pulse control at the tip-apex, with interferometric spectrogram of transform limited pulse and pulse with $200\,\text{fs}^2$ applied chirp (**d**). Corresponding reconstructed spectral amplitude and phase (**e**), showing the close agreement between the applied and extracted phase. After Ref. [67]. *Copyright 2011 American Chemical Society*

metallic structure [66]. This requires some combination of spatial, spectral, and phase pulse shaping of the driving laser field, and often uses *adaptive* algorithms in order to generate the desired nanofocus at a particular spatial location. However, the necessary local interference relies on a specific phase relationship between modes and therefore limits the spectral and temporal degrees of freedom available at the nanofocus. In contrast, the adiabatic nanofocusing process retains essentially all degrees of freedom to *deterministically* control the optical transient at the tip apex.

A demonstration of femtosecond-nanometer spatio-temporal control based on plasmonic nanofocusing on a tip is shown in Fig. 7.19. Femtosecond SPPs are launched onto an electrochemically etched Au tip using a grating structure formed by focused ion beam milling to overcome the photon-SPP momentum mismatch (a). The grating is spatially chirped for maximum coupling bandwidth. The SPPs then propagate along the tip towards the apex experiencing an increasing effective index of refraction, which leads to an increasing wavevector, decreasing group velocity, and increasing spatial confinement, thereby concentrating the electric field into the tip apex. Figure 7.19a shows the 20 nm spatial field localization at the tip apex. The efficiency of the process is high enough that, combined with the symmetry-breaking along the cone axis, SHG can be generated at the tip apex. This enables full characterization of the electric field transient at the apex, for example through IFROG, as discussed previously. Furthermore, frequency-domain pulse-shaping can be used to compensate dispersion with a multiphoton intrapulse interference phase scan (MIIPS) algorithm [68], and also to generate pulse pairs with controllable delay for the IFROG measurements themselves. The resulting spectrogram for a few-

femtosecond pulse at the apex is shown in Fig. 7.19b, with reconstructed amplitude and phase (c) corresponding to a 16 fs transform limited pulse. Arbitrary waveform generation and full deterministic control of the re-emitted apex radiation is possible through feedback on the nonlinear response of the tip. This capability is demonstrated here by applying a 200 fs^2 chirp to one of the pulses at the apex, as shown in Fig. 7.19 d) and e) with the reconstructed amplitude and phase characteristics from cross-correlation FROG (XFROG).

The localized plasmon resonance for the tips in these experiments was red-shifted relative to the laser bandwidth, so that the tips were non-resonant and had an almost instantaneous response. A tip with a plasmon resonance close to the laser wavelength would provide higher field enhancement, but with achievable minimum pulse duration now limited by the plasmon dephasing time to a few 10s of fs, rather than the SPP coupling bandwidth. The adiabatic nanofocusing process is necessarily accompanied by a decrease in the SPP group velocity on approaching the apex of the tip [69]. This SPP slow-down could provide a further increase in the nonlinear response of the tip-apex.

The grating-coupled tips demonstrate how the combination of the ultrafast optical properties of metals and intrinsic and geometry-related SPP behavior allow for spatio-temporal nano-imaging in a scanning probe configuration. Together with the nonlinear optical response of the tips and associated symmetry selectivity, this opens the door for deterministic few-femtosecond optical control on the nanoscale.

More generally, the design and optimization of optical antennas for nonlinear applications requires the ability to accurately characterize field enhancement and mode distribution properties within an antenna. Both electron-based techniques and photon-based techniques have been used for antenna characterization. Conventional far-field optical characterization can provide information about the interaction between an optical antenna and propagating light, such as the relationship between device geometry and resonant frequency. While this can be applied over a broad frequency range, it suffers from comparatively low spatial resolution, and the linear response of an antenna does not necessarily predict the nonlinear response, due to, for example, coupling between the driven plasmon and surrounding dielectric resonances. SHG, two-photon photoluminescence, or FWM can provide slightly higher spatial resolution and more accurate determination of nonlinear spectral properties, but do not provide knowledge of the underlying resonant modes and their associated spatial field distribution within an antenna. Electronic techniques, such as electron energy loss spectroscopy (EELS), transmission electron microscopy (TEM), and cathodoluminescence can facilitate the extraction of spatially detailed information, with nanometer resolution of modes and plasmonic field enhancement.

In recent years, near-field optical techniques such as scattering-scanning near-field optical microscopy (*s*-SNOM) have been utilized for high spatial resolution mapping of linear and nonlinear antenna properties, offering information about the local optical electric field magnitude and phase and interactions of modes in coupled nano-optical plasmonic and optical antenna structures. With these techniques it is possible to spatially and spectrally probe the microscopic electric field distribution, and correlate details of the field with geometrical features. Such measurements

have demonstrated spectral shifts between near-field and far-field, and confirmed that within nanoparticle geometries the nonlinear response remains sensitive to deviations in shape and other defects. These techniques are applicable to various antenna geometries and wavelengths, provide design criteria for more complex architectures and modeling, and can facilitate impedance matching of optical antennas to quantum systems [56].

The adiabatic focusing tips provide an avenue to extend ultrafast spectroscopic imaging from its conventional far-field spatial resolution limit to the nanoscale. Historically, the development of femtosecond pulsed lasers has enabled the investigation of ultrafast dynamics on the characteristic time scales of the elementary electronic and vibrational excitations in matter, with direct and selective spectroscopic access to the corresponding energy levels. The combination with spectral pulse shaping [70] provides the additional capability to control the coherent evolution of these quantum excitations (quantum coherent control), which allows steering of chemical reactions or control of optical and electronic material properties [71].

It would be highly desirable to extend the spatial resolution to the nanoscale to simultaneously access ultrafast dynamics on their associated natural length scales of the elementary electronic, visible, and spin excitations. The potential use of plasmonic properties to achieve this nanometer-femtosecond spatio-temporal control of optical excitations for imaging and spectroscopy has attracted much interest. With such a nano-optical technique, individual molecules, quantum dots, and plasmonic structures, for example, could be investigated even in dense inhomogeneous media, in addition to spatial and non-local dynamics to probe transport, propagation, and spatial coupling properties. The adiabatic focusing tips are one optical antenna concept for overcoming the diffraction limit and providing the desired high spatial field localization, with high bandwidth, and high off-resonant field enhancement thus supporting even the shortest possible few-fs optical pulses.

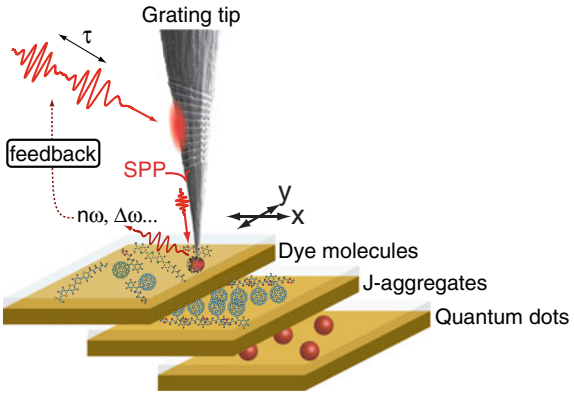

Fig. 7.20 Implementation of nanofocused, background-free coherent control and interaction dynamics experiments on single quantum emitter. Feedback for optimization is based on the luminescent or nonlinear response of the coupled tip-sample system

A possible implementation of adiabatic nanofocusing for imaging and quantum coherent control is shown in Fig. 7.20. Nanofocusing in combination with a pulse shaper can be used for linear, nonlinear, and ultrafast imaging of individual quantum systems or mesoscale variations in bulk and dense media. The ability to control pulse shape and phase enables pump-probe techniques including collinear FWM and 2D spectroscopy. With feedback on the nonlinear response $n\omega$, $\Delta\omega$, etc. of the coupled tip-sample system, the extension to quantum coherent control of the excitation pathways and evolution of the quantum system can also be realized.

7.6 Outlook

In spite of the importance of the nonlinear and ultrafast properties of metals in general, and metal surfaces and nanostructures in particular, for the understanding of phenomena such as enhancement and dephasing associated with plasmonic excitations, considerably less work has focused on these aspects compared to linear and continuous wave spectroscopies. Here we have attempted to summarize the fundamental properties of metals and their plasmonic excitations and their effect on nonlinear behavior in order to provide a guide for future extensions of plasmonic studies. Optimization and control of these properties will be important for increasing sensitivity and efficiency in a wide variety of sensing and optical switching applications, chemical spectroscopy, nano-scale imaging, or coherent control on the single quantum limit. Thus far little nanoscale imaging using a nonlinear response of a material system has been demonstrated [72], though there has been interest in further increasing the nonlinear response through engineered nanostructures with a bulk nonlinear material at the tip, which in combination with field enhancement could provide high wave-mixing conversion efficiencies and provide access to the additional spectroscopic and symmetry degrees of freedom enabled by nonlinear techniques.

In the ultrafast regime, recent work has demonstrated the control of mode coherence in plasmonic systems [73], in spite of the extremely short plasmon dephasing times. Control of coupled plasmonic-photonic modes, which have longer dephasing times and are therefore easier to direct, has been demonstrated, as well as nonlinear wavemixing in plasmonic-photonic waveguides [57]. Furthermore, taking advantage of SPP properties such as nanofocusing provides one of the most promising routes to achieving full spatio-temporal control of individual nanostructures and nanosystems. This would enable the investigation of ultrafast dynamics on the characteristic time scales of the elementary electronic and vibrational excitations in matter, and with direct and selective spectroscopic access to the corresponding energy levels.

Plasmonic field enhancement has also been proposed as a means to achieve the necessary peak intensities for high harmonic generation (HHG). However, the nanoscopic interaction volume for plasmonic structures in comparison to conventional intracavity HHG suggests that the HHG yield would be small [74]. Multiphoton or high field fluorescence processes can however be effective below the ionization threshold [75], in addition to electron emission [76]. Strong field gradients could

be utilized for nanomanipulation and trapping [77]. Additionally, strong coupling of quantum states to light can lead to qualitatively new nonlinear optics, beyond the perturbative regime discussed here. Structures based on molecular, quantum-dot, or quantum-wire exciton resonances coupled to plasmonic metal nanostructures can be optimized to form hybrid modes with large optical nonlinearities as a result of the quantum interference of the exciton dipole oscillation and the plasmonic modes, providing new quantum states and allowing new avenues for ultrafast control [78]. Further improvements in design and control or plasmonic structures could also provide access to non-local nonlinear effects.

Acknowledgments Funding was provided by the National Science Foundation (NSF CAREER Grant CHE 0748226). Part of the work was performed at the Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility from DOE's Office of Biological and Environmental Research at Pacific Northwest National Laboratory (PNNL). PNNL is operated by Battelle for the US DOE under the contract DEAC06-76RL01830.

References

1. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer, New York, 1987)
2. M. Kerker, *The Scattering of Light, and Other Electromagnetic Radiation* (Academic Press, New York, 1969)
3. S.A. Maier, *Plasmonics: Fundamentals and Applications* (Springer, New York, 2007)
4. N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Brooks Cole, Belmont, 1976)
5. F. Ladstädter, U. Hohenester, P. Puschnig, C. Ambrosch-Draxl, Phys. Rev. B **70**, 235125 (2004)
6. D.J. Roaf, Philos. T. R. Soc. A **255**(1052), 135 (1962)
7. A. Sommerfeld, Z. Phys. A **47**(1), 1 (1928)
8. M. Dressel, G. Grüner, *Electrodynamics of Solids: Optical Properties of Electrons in Matter* (Cambridge University Press, New York, 2002).
9. P. Johnson, R. Christy, Phys. Rev. B **6**, 4370 (1972)
10. R.L. Olmon, B. Slovick, T.W. Johnson, D. Shelton, S.H. Oh, G.D. Boreman, M.B. Raschke, Phys. Rev. B **86**, 235147 (2012)
11. C.F. Bohren, D.R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 2008)
12. U. Kreibig, M. Vollmer, *Optical Properties of Metal Clusters, Springer Series in Materials Science*, vol. 25 (Springer, Berlin, 1995)
13. C. Ciracì, R.T. Hill, J.J. Mock, Y. Urzhumov, A.I. Fernández-Domínguez, S.A. Maier, J.B. Pendry, A. Chilkoti, D.R. Smith, Science **337**(6098), 1072 (2012)
14. G. Mie, Ann. Phys. **330**, 377 (1908)
15. H.U. Yang, S. Berweger, J.M. Atkin, M.B. Raschke (in preparation)
16. S. Link, C. Burda, M.B. Mohamed, B. Nikoobakht, M.A. El-Sayed, Phys. Rev. B **61**(9), 6086 (2000)
17. G.V. Hartland, Chem. Rev. **111**(6), 3858 (2011)
18. C. Sönnichsen, T. Franzl, T. Wilk, G. von Plessen, J. Feldmann, O. Wilson, P. Mulvaney, Phys. Rev. Lett. **88**(7), 077402 (2002)
19. M. Liu, M. Pelton, P. Guyot-Sionnest, Phys. Rev. B **79**, 035418 (2009)
20. T.V. Shahbazyan, I.E. Perakis, J.Y. Bigot, Phys. Rev. Lett. **81**, 3120 (1998)
21. R. Boyd, *Nonlinear Optics* (Academic Press, New York, 2003)
22. Y.R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984)

23. M.B. Raschke, M. Hayashi, S.H. Lin, Y.R. Shen, Chem. Phys. Lett. **359**(5–6), 367 (2002)
24. S. Jha, Phys. Rev. **140**(6A), 2020 (1965)
25. N. Bloembergen, R.K. Chang, S.S. Jha, C.H. Lee, Phys. Rev. **174**, 813 (1968)
26. J. Rudnick, E.A. Stern, Phys. Rev. B **4**, 4274 (1971)
27. J.E. Sipe, V.C.Y. So, M. Fukui, G.I. Stegeman, Phys. Rev. B **21**, 4389 (1980)
28. A. Liebsch, W.L. Schaich, Phys. Rev. B **40**, 5401 (1989)
29. C.M. Li, L.E. Urbach, H.L. Dai, Phys. Rev. B **49**, 2104 (1994)
30. G. Petrocelli, S. Martellucci, R. Francini, Appl. Phys. A Mater. Sci. Proces. **56**, 263 (1993)
31. E.K.L. Wong, G.L. Richmond, J. Chem. Phys. **99**(7), 5500 (1993)
32. D. Krause, C.W. Teplin, C.T. Rogers, J. Appl. Phys. **96**(7), 3626 (2004)
33. H.J. Simon, D.E. Mitchell, J.G. Watson, Phys. Rev. Lett. **33**, 1531 (1974)
34. T.Y.F. Tsang, Opt. Lett. **21**(4), 245 (1996)
35. F. De Martini, Y.R. Shen, Phys. Rev. Lett. **36**, 216 (1976)
36. K. Ujihara, Opt. Commun. **42**(1), 1 (1982)
37. N. Rotenberg, M. Betz, H.M. van Driel, Phys. Rev. Lett. **105**, 017402 (2010)
38. A.T. Georges, N.E. Karatzas, Phys. Rev. B **85**, 155442 (2012)
39. J. Renger, R. Quidant, N. van Hulst, S. Palomba, L. Novotny, Phys. Rev. Lett. **103**, 266802 (2009)
40. M. Fukui, J. Sipe, V. So, G. Stegeman, Solid State Commun. **27**(12), 1265 (1978)
41. G.I. Stegeman, J.J. Burke, D.G. Hall, Appl. Phys. Lett. **41**(10), 906 (1982)
42. Q. Zhang, K. Lin, Y. Luo, Opt. Express **19**(6), 4991 (2011)
43. N.B. Grosse, J. Heckmann, U. Woggon, Phys. Rev. Lett. **108**, 136802 (2012)
44. V.M. Shalaev, *Nonlinear Optics of Random Media: Fractal Composites and Metal-dielectric Films*, vol. 158 (Springer Tracts in Modern Physis, Heidelberg, 2000)
45. E.L. Ru, J. Grand, N. Felidj, J. Aubard, G. Levi, A. Hohenau, J. Krenn, E. Blackie, P. Etchegoin, J. Phys. Chem. C **112**, 8117 (2008)
46. M. Moskovits, Rev. Mod. Phys. **57**, 783 (1985)
47. A. Wokaun, J.G. Bergman, J.P. Heritage, A.M. Glass, P.F. Liao, D.H. Olson, Phys. Rev. B **24**, 849 (1981)
48. C.K. Chen, A.R.B. de Castro, Y.R. Shen, Phys. Rev. Lett. **46**, 145 (1981)
49. D.S. Chemla, J.P. Heritage, P.F. Liao, E.D. Isaacs, Phys. Rev. B **27**, 4553 (1983)
50. F.X. Wang, F.J. Rodriguez, W.M. Albers, M. Kauranen, New J. Phys. **12**, 063009 (2010)
51. S. Roke, G. Gonella, Ann. Rev. Phys. Chem. **63**(1), 353 (2012)
52. J.I. Dadap, J. Shan, T.F. Heinz, J. Opt. Soc. Am. B **21**, 1328 (2004)
53. C. Neacsu, G. Reider, M. Raschke, Phys. Rev. B **71**(20), 201402 (2005)
54. S. Roke, W.G. Roeterdink, J.E.G.J. Wijnhoven, A.V. Petukhov, A.W. Kleyn, M. Bonn, Phys. Rev. Lett. **91**, 258302 (2004)
55. J. Nappa, I. Russier-Antoine, E. Benichou, C. Jonin, P.F. Brevet, J. Chem. Phys. **125**(18), 184712 (2006)
56. R. Olmon, M. Raschke, Nanotechnology **23**, 444001 (2012)
57. T. Utikal, T. Zentgraf, T. Paul, C. Rockstuhl, F. Lederer, M. Lippitz, H. Giessen, Phys. Rev. Lett. **106**, 133901 (2011)
58. M. Danckwerts, L. Novotny, Phys. Rev. Lett. **98**, 026104 (2007)
59. M. Kauranen, A.V. Zayats, Nat. Photon. **6**(11), 737 (2012)
60. C. Timm, K.H. Bennemann, J. Phys. Condens. Matter. **16**(4), 661 (2004)
61. R. Trebino, *Frequency Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses* (Kluwer, New York, 2002)
62. A. Anderson, K.S. Deryckx, X.G. Xu, G. Steinmeyer, M.B. Raschke, Nano Lett. **10**, 2519 (2010)
63. H. Petek, S. Ogawa, Prog. Surf. Sci. **56**(4), 239 (1997)
64. A.J. Babadjanyan, N.L. Margaryan, K.V. Nerkararyan, J. Appl. Phys. **87**(8), 3785 (2000)
65. M.I. Stockman, Phys. Rev. Lett. **93**, 137404 (2004)
66. M. Aeschlimann, M. Bauer, D. Bayer, T. Brixner, S. Cunovic, F. Dimler, A. Fischer, W. Pfeiffer, M. Rohmer, C. Schneider, F. Steeb, C. Strüber, D.V. Voronine, Proc. Nat. Acad. Sci. **107**, 5329 (2010)

67. S. Berweger, J.M. Atkin, X.G. Xu, R.L. Olmon, M.B. Raschke, Nano Lett. **11**, 4309 (2011)
68. B. Xu, J.M. Gunn, J.M.D. Cruz, V.V. Lozovoy, M. Dantus, J. Opt. Soc. Am. B **23**(4), 750 (2006)
69. V. Kravtsov, J.M. Atkin, M.B. Raschke, Opt. Lett. **38**(8), 1322 (2013)
70. A.M. Weiner, Rev. Sci. Intr. **71**, 1929 (2000)
71. A. Assion, T. Baumert, M. Bergt, T. Brixner, B. Kiefer, V. Seyfried, M. Strehle, G. Gerber, Science **282**(5390), 919 (1998)
72. C.C. Neacsu, B.B. van Aken, M. Fiebig, M.B. Raschke, Phys. Rev. B **79**(10), 100107 (2009)
73. A. Kubo, K. Onda, H. Petek, Z. Sun, Y.S. Jung, H.K. Kim, Nano Lett. **5**, 1123 (2005)
74. M.B. Raschke, Annalen der Physik **525**(3), 40 (2013)
75. M. Sivis, M. Duwe, B. Abel, C. Ropers, Nature **485**, 7397 (2012)
76. C. Ropers, D.R. Solli, C.P. Schulz, C. Lienau, T. Elsaesser, Phys. Rev. Lett. **98**, 043907 (2007)
77. M.L. Juan, M. Righini, R. Quidant, Nat. Photon. **5**, 349 (2011)
78. P. Vasa, R. Pomraenke, G. Cirmi, E. De Re, W. Wang, S. Schwieger, D. Leipold, E. Runge, G. Cerullo, C. Lienau, ACS Nano **4**(12), 7559 (2010)