Chapter 1
Optical Properties of Plasmonic Materials

Maxwell's equations state that a dielectric–metallic interface can support surface plasmon polaritons (SPPs), which are coherent electron oscillation waves that propagate along the interface with an electromagnetic wave. The unique properties of the interface waves result from the frequency-dependent dispersion characteristics of metallic and dielectric materials. This chapter provides an introduction to alternative plasmonic materials, as well as the rationale for each material choice. The comprehensive optical properties of various materials, including noble metals and semiconductors, are presented. The optical properties are evaluated based on the permittivity and permeability defined by either the Drude or Lorentz model. Furthermore, the noble metals are described from the generally approved data in a general handbook of solid materials, such as the *Handbook of Optical Constants of Solids*, edited by Palik. This chapter outlines the effective medium approaches for describing the effective dielectric functions of composite nanostructures. It also provides a reference for finding better plasmonic materials at specific frequencies.

1.1 Electromagnetic Waves Propagating through Materials

1.1.1 Fundamental equations of electromagnetic waves

The modern electromagnetic wave theory has been developed from the group work in theory by James Clerk Maxwell, Oliver Heaviside, and Josiah Willard Gibbs from 1861 to 1865, and verified experimentally by Heinrich Hertz in 1887. During the establishment processes of the electromagnetic wave theory, Maxwell's contribution was settled in his two famous papers: "On Physical Lines of Force" in 1861 and "A Dynamical Theory of the Electromagnetic Field" in 1865. In his theory, four equations composed the fundamentals of the electromagnetic wave theory and are now universally known as Maxwell's equations. They demonstrate the unifying connection between electromagnetic waves and light, from the extremely long wavelengths of radio, television,

radar, and microwaves, to the shorter wavelengths of visible light and very short ultraviolet light.

The electromagnetic wave theory describes the fundamental concepts of the interaction processes between matter and electromagnetic energy in numerous scientific disciplines. Since the 20th century, it has been understood that the theory of quantum electrodynamics can give a more accurate and fundamental explanation of some phenomena related to photons, photon scattering, and quantum optics. Regardless, the sophisticated electromagnetic wave theory has been furthering our comprehension of the flourishing scientific disciplines of plasmonic optics.

In electromagnetic wave theory, the conventional Maxwell's equations are formulated via vector notation in two forms of integral equations and differential forms. The two forms are mathematically equivalent and useful in same physical meaning. The integral equations are often used to directly calculate fields in the conditions of symmetric distributions of electric charges and electric currents. At the same time, the different equations are convenient formulations in more complicated situations, such as using the finite element analysis method.

There are also microscopic and macroscopic variants of Maxwell's equations. The differential forms of macroscopic Maxwell equations are as follows:

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} B, \quad \nabla \times \mathbf{H} = \frac{\partial}{\partial t} \mathbf{D} + \mathbf{J}_{ext}, \tag{1.1a}$$

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \cdot \mathbf{D} = \rho, \tag{1.1b}$$

where $\mathbf{E}$ (volts/m) and $\mathbf{H}$ (amperes/m) describe the electric and magnetic field, respectively. $\mathbf{D}$ (coulombs/m^2) and $\mathbf{B}$ (webers/m^2) correspond to the electric displacement and the magnetic flux density vectors, respectively. These four variables are generally time-dependent and position-dependent parameters, which are created by electric charges or electric currents and thus expressed by the local charge density per unit volume ρ (coulombs/m^3) and the external current density $\mathbf{J}_{ext}$ (amperes/m^2).

From the macroscopic point of view, Eqs. (1.1a) describe Faraday's law of induction and Ampère's circuital law. The former states that a time harmonic magnetic field induces an electric field, whereas the latter states that both an electric current and a varying electric field can generate a magnetic field. Equations (1.1b) give Gauss' laws for magnetic and electric fields. Gauss' laws for electric fields formulate the conversion between a static electric field and the electric charges; Gauss' law for magnetic field states that the sum total magnetic flux through any finitude volume surface is zero, which means that there are no magnetic charges. In a 3D vector system, the parameters $\mathbf{B}$, $\mathbf{H}$, $\mathbf{D}$, ρ, and $\mathbf{J}_{ext}$ are considered to be functions of both position and time ($\mathbf{r}$, t).

In the macroscopic Maxwell's equations, the total electric charge density $\boldsymbol{\rho}$ is factorized into a bound component $\boldsymbol{\rho}_b$ and a free counterpart $\boldsymbol{\rho}_f$, just as the total current density $\mathbf{J}$ is separated into a free component $\mathbf{J}_f$ and a bound component $\mathbf{J}_b$:

$$\boldsymbol{\rho} = \boldsymbol{\rho}_b + \boldsymbol{\rho}_f, \quad \mathbf{J} = \mathbf{J}_b + \mathbf{J}_f. \tag{1.2}$$

The bound charge density $\boldsymbol{\rho}_b$ and current density $\mathbf{J}_b$ are defined by introducing the terms of polarization $\mathbf{P}$ and magnetization $\mathbf{M}$ in the form of

$$\boldsymbol{\rho}_b = -\nabla \cdot \mathbf{P}, \quad \mathbf{J}_b = \nabla \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t}. \tag{1.3}$$

The conservation law for the electric charge density and the external current density is described as

$$\nabla \times \mathbf{J}_{ext} = -\frac{\partial}{\partial t}\boldsymbol{\rho}. \tag{1.4}$$

Equation (1.4) indicates that the divergence of the current density leaving a volume equals the decrease rate of the charge density in the same volume. This equation can be obtained from Eqs. (1.1) by taking the divergence of Ampère's circuital law and introducing Gauss' laws for electric fields.

In the case of a nonmagnetic medium, where only the free current density remains, the magnetization and the subsequent bound current density are zero. The continuity equation [Eq. (1.4)] becomes the partial component corresponding to the free charges

$$\nabla \times \mathbf{J}_f = -\frac{\partial}{\partial t}\boldsymbol{\rho}_f. \tag{1.5}$$

In a bulk matter without electric charges or currents, the macroscopic Maxwell's equations are presented in the form of the microscopic variant:

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{B}}{\partial t}, \tag{1.6}$$

$$\nabla \cdot \mathbf{H} = 0, \quad \nabla \times \mathbf{H} = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{1.7}$$

Equations (1.6) and (1.7) describe an electric and magnetic field in a vacuum, where μ_0 and ε_0 are the permittivity and permeability in the free space of vacuum ($\mu_0 \approx 4\pi \times 10^{-7}$ henry/meter and $\varepsilon_0 \approx 8.85 \times 10^{-12}$ farad/meter). They refer only to the cases that do not account for the medium without charges and currents; they do not mean that the space of the medium is empty of charge or current. For an oscillating wave $\exp(j\omega t)$, Eqs. (1.1a) can be rewritten in the simple form of multiplication by the time derivatives:

$$\nabla \times \mathbf{E} = -j\omega\mathbf{B}, \quad \nabla \times \mathbf{H} = j\omega\mathbf{D} + \mathbf{J}_{ext}, \tag{1.8}$$

where $\omega = 2\pi f$ represents the angular frequency of an oscillating wave. A time-oscillatory plane wave propagating in 3D notation is

$$u(\mathbf{r},t) = A\exp(j\omega t + jk\mathbf{r}), \tag{1.9}$$

where the complex amplitude A refers to any one component of the aforementioned field vectors. The wavenumber k is defined by the number of wavelengths per 2π units

$$k = 2\pi/\lambda = 2\pi f/\upsilon_p = \omega/\upsilon_p, \tag{1.10}$$

where f refers to the frequency of the propagating wave, λ is the wavelength in a media, and υ_p is the phase velocity of the wave field. A series of sources generate a linear combination of time-dependent wave fields:

$$U(\mathbf{r},t) = \sum_{i=1}^{N} A_i \exp(j\omega_i t + jk_i\mathbf{r}). \tag{1.11}$$

In general cases, the complex amplitude A_i depends on the initial and boundary conditions. The combined fields are called the spatial spectrum. The time-oscillating propagating waves can also be described in any specific axis. The notation r in Eqs. (1.9) and (1.11) is replaced by another displacement vector, such as in the x, y, and z axes.

1.1.2 Constitutive equations of inhomogeneous media

The constitutive equations of a medium specify the physical kinetic response to external stimuli combined with other physical laws. In electromagnetic wave theory, they are applied to describe the electrical and magnetic response of various media. In Maxwell's macroscopic equations, the constitutive equations describe the relations between the displacement field **D** and the electric field **E**, and the magnetizing field **H** and the magnetic field **B**. In other words, they describe the dynamic response of bound charges and current to external applied fields.

Within the vast majority of isotropic materials, the electric and magnetic fields are often investigated separately. The constitutive relations of general materials without polarization and magnetization are commonly written as:

$$\mathbf{D} = \varepsilon\mathbf{E}, \quad \mathbf{H} = \mathbf{B}/\mu, \tag{1.12}$$

where ε and μ are the absolute permittivity and permeability of a general medium, respectively. Even for simple linear media, the constitutive relations Eq. (1.12) have various complications. For example, in homogeneous materials, ε and μ are constant values throughout the media, while in inhomogeneous

materials, they are position dependence within the materials. For anisotropic materials, ε and μ are in the form of vectors, while for isotropic materials, they are denoted as scalars. Because general materials are dispersive, both ε and μ are dependent on the frequency of electromagnetic wave.

The vast majority of natural materials are electrically neutral at the macroscopic level. This macroscopic electrical neutrality is the consequence of the internal equilibrium of collective charge interactions. When an electromagnetic wave impinges on electrical neutrally media, the time-dependent electric and magnetic field will induce separate oscillatory charge displacements. Such local separation of positive and negative charges from their original positions in opposite directions manifests the media to present in the form of induced electric dipole momentum. These phenomena are called polarizations. Similar effects induced by magnetic field are magnetization or magnetic polarization. For an inhomogeneous medium with polarization and magnetization, the continuous constitutive relations are described to be the functions of space coordinates and time variables:[2]

$$\mathbf{D}(r,t) = \varepsilon_0 \mathbf{E}(r,t) + \mathbf{P}(r,t), \tag{1.13a}$$

$$\mathbf{B}(r,t) = \mu_0 \mathbf{H}(r,t) + \mathbf{M}(r,t). \tag{1.13b}$$

Equation (1.13a) indicates that a dielectric medium is characterized by a free-space part $\varepsilon_0\mathbf{E}(r, t)$ and a polarization vector $\mathbf{P}(r, t)$. The latter represents the electric dipole moment. The electric flux density $\mathbf{D}$ represents the organization of electric charges induced by an external electric field $\mathbf{E}$. In the presence of the external field, the polarization vector is caused by induced dipole moments, alignment of the permanent dipole moments, and the migration of electric charges. Equation (1.13b) states that a magnetic medium can also be described by a free-space part $\mu_0\mathbf{H}(r, t)$ and a magnetization vector $\mathbf{M}(r, t)$. A medium is called diamagnetic if $\mu < \mu_0$, whereas it is paramagnetic if $\mu > \mu_0$. For a diamagnetic medium, the induced magnetic moments tend toward the opposite direction of the external magnetic field, whereas a paramagnetic medium features an alignment of magnetic moments within the medium.

The instantaneous polarizations of a medium to an electromagnetic field are the convolution integration over the previous history response:[2]

$$\mathbf{P}(r,t) = \varepsilon_0 \int_{-\infty}^{t} \kappa_e(r,t-\tau)\mathbf{E}(r,\tau)d\tau, \tag{1.14a}$$

$$\mathbf{M}(r,t) = \mu_0 \int_{-\infty}^{t} \kappa_m(r,t-\tau)\mathbf{H}(r,\tau)d\tau, \tag{1.14b}$$

where the scalar functions of κ_e and κ_m are the electric and magnetic susceptibilities, both of which are time-dependent scalar. In the frequency domain, the aforementioned induced polarizations are represented by a multiplication operation:

$$\mathbf{P}(r,\omega) = \varepsilon_0 \kappa_e(r,\omega)\mathbf{E}(r,\omega), \tag{1.15a}$$

$$\mathbf{M}(r,\omega) = \mu_0 k_m(r,\omega)\mathbf{H}(r,\omega), \tag{1.15b}$$

where $\kappa_e(r, \omega)$ and $\kappa_m(r, \omega)$ are the Fourier-transformed kernels corresponding to the electric and magnetic susceptibilities κ_e and κ_m, respectively. The constitutive relations of an inhomogeneous medium in the frequency domain are:[2]

$$\mathbf{D} = \varepsilon_0[1 + \kappa_e(r,\omega)]\mathbf{E} = \varepsilon_0\varepsilon_r(r,\omega)\mathbf{E} = \varepsilon(r,\omega)\mathbf{E}, \tag{1.16a}$$

$$\mathbf{B} = \mu_0[1 + \kappa_m(r,\omega)]\mathbf{H} = \mu_0\mu_r(r,\omega)\mathbf{H} = \mu(r,\omega)\mathbf{H}, \tag{1.16b}$$

where $\varepsilon_r(r, \omega) = 1 + \kappa_e(r, \omega)$ and $\mu_r(r, \omega) = 1 + \kappa_m(r, \omega)$ are the relative permittivity and relative permeability in frequency region, respectively. $\varepsilon(r, \omega) = \varepsilon_0\varepsilon_r(r, \omega)$ and $\mu(r, \omega) = \mu_0\mu_r(r, \omega)$ are the corresponding absolute permittivity and absolute permeability. Here, both the electric field $\mathbf{E}$ and the magnetic field $\mathbf{H}$ are complex amplitudes. From the macroscopic point of view, these vector quantities are the averaging of the microscopic electric and magnetic field vectors over the unit cell of a medium.

In plasmonic optics, the complex artificial nanostructures, such as cubic lattice, may be constructed by kinds of nanoparticles, even with different natural substrates. In such inhomogeneous conditions, the parameters $\varepsilon_r(r,\omega)$ and $\mu_r(r,\omega)$ in Eq. (1.16) are spatially variant.

The dependency characteristics of the frequencies of time-harmonic fields are called the materials' dispersion. In general, all physical materials have some dispersion because no material can respond instantaneously to external acting fields. Additionally, in a nonlinear medium, both the permittivity and the permeability depend on the strength of the electric field and the external magnetic field, respectively.

1.1.3 Isotropic and anisotropic media

The responses of an isotropic medium to an external electromagnetic field are invariant in notational directions. Thus, the magnetization, polarization, and other parameters of an isotropic medium can be expressed simply by the scalar coefficients of permittivity and permeability. However, numerous composite materials, especially artificial nanostructures emerging in the field of plasmonic optics, are anisotropic media because their internal microstructures are an asymmetric configuration of lattice patterns.

Although they remain spatially homogeneous, these anisotropic materials have distinct axis directions. The responses of anisotropic materials to an external field depend on the orientation of the external field with respect to its internal alignment. In such cases, the constitutive equations should be generalized by introducing medium parameters in the form of the permittivity and permeability tensors.

Another complication of anisotropic materials related to their complex geometry is the magneto-electric coupling effect, i.e., an external electric field induces not only electric dipole momentum but also magnetic dipole momentum inside an anisotropic material. An applied magnetic field generally generates both magnetic and electric polarizations. For a simply anisotropic material, the magneto-electric coupling effect occurs at perpendicular polarization directions, consisting of the orthogonal components of electric and magnetic components.

In such bi-anisotropic media with magneto-electric coupling effects, the general constitutive equations are formulated by introducing coupling tensors as

$$\mathbf{D} = \varepsilon\mathbf{E} + \mathbf{P}_0 + \sqrt{\varepsilon_0\mu_0}\kappa_{em}\mathbf{H}, \tag{1.17a}$$

$$\mathbf{B} = \mu\mathbf{H} + \mathbf{M}_0 + \sqrt{\varepsilon_0\mu_0}\kappa_{me}\mathbf{E}, \tag{1.17b}$$

where $\varepsilon = \varepsilon_0(1 + \kappa_e)$ and $\mu = \mu_0(1 + \kappa_m)$ are the absolute permittivity and permeability, respectively. $\mathbf{P}_0$ and $\mathbf{M}_0$ are called the static electric polarization and static magnetization in the absence of external alternating fields. The dimensionless tensors of κ_{em} and κ_{me} are introduced to describe the coupling effect between the electric polarization induced by a magnetic field and the magnetization induced by an electric field. The magnetic susceptibility describes the linear magneto-electric effect, which is often observed in artificial bi-anisotropic materials, such as split-ring resonators and wires.

1.1.4 Constitutive equations of dielectric media

The discussed constitutive equations covering polarization and magnetization expansions address the major phenomena encountered for the vast majority of artificial nanostructures and natural materials. In the cases of a symmetric lattice, the number of tensor components can be reduced. For example, the magneto-electric coupling effect within isotropic media is orthorhombic independent, i.e., four scalar parameters characterize such bi-isotropic materials instead of four tensors.

Dielectric media are some of the most dominant materials used for traditional optical components due to their effective manipulation of light waves. The physical understanding for propagating waves within dielectric materials can also be analyzed using Maxwell's equations and constitutive

relations. When the instantaneous responses of homogeneous dielectric media are not considered, the constitutive relations are expressed as:[5]

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 (1 + \chi_e) \mathbf{E}, \tag{1.18a}$$

$$\mathbf{B} = \mu_0 \mathbf{H} + \mathbf{M} = \mu_0 (1 + \chi_m) \mathbf{H}, \tag{1.18b}$$

where $\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E}$ and $\mathbf{M} = \mu_0 \chi_m \mathbf{H}$ are the polarization density and the induced magnetization within dielectric media, respectively. The introduced terms of χ_e and χ_m are called the electric and magnetic susceptibility, respectively. The relative permittivity and permeability are denoted as $\varepsilon_r = 1 + \chi_e$ and $\mu_r = 1 + \chi_m$. Both the electric displacement $\mathbf{D}$ and magnetic field $\mathbf{B}$ can be considered to be spatially dependent in inhomogeneous media (and likely time-harmonic dependent, too).

Likewise, the instantaneous electromagnetic responses of a homogeneous dielectric medium at a certain time depend on the fields at that time and the evolutionary progress over a period of past time. Thus, the constitutive relations involve the time evolutionary variation as:

$$\mathbf{D}(t) = \varepsilon_0 \mathbf{E}(t) + \varepsilon_0 \int_{-\infty}^{t} \chi_e (t - \tau) \mathbf{E}(\tau) d\tau, \tag{1.19a}$$

$$\mathbf{B}(t) = \mu_0 \mathbf{H} + \mu_0 \int_{-\infty}^{t} \chi_m (t - \tau) \mathbf{H}(\tau) d\tau. \tag{1.19b}$$

The constitutive relations are still valid as long as the susceptibilities are independent of the strength of the electric and magnetic fields. The evolutionary constitutive relations in the frequency domain are written as

$$\mathbf{D}(\omega) = \varepsilon_0 [1 + \chi_e(\omega)] \mathbf{E}(\omega), \tag{1.20a}$$

$$\mathbf{B}(\omega) = \mu_0 [1 + \chi_m(\omega)] \mathbf{H}(\omega). \tag{1.20b}$$

In plasmonic optics, both dielectric and metallic materials are employed to construct the unit cells in plasmonic devices, such as metamaterials and metasurfaces. Metals are much more dispersive than dielectric materials in the visible and near-infrared region. Nevertheless, the constitutive relations of dielectric materials enable us to understand the origin of the frequency dispersion in plasmonic components. In particular, when the range of interest is extended to the mid-infrared or longer wavelengths, the properties of the frequency dispersion for most dielectric materials begin to change. For example, silicon dioxide, which is transparent in wavelengths less than 4.0 μm, becomes opaque when the wavelengths extend to the mid-infrared and longer than 5.0 μm. Silicon dioxide shows evident reflectivity in the wavelength range of 8.0–10.0 μm. The development of plasmonic optical devices should account for the constitutive relations of both dielectric and

metallic constituents. The material constituents and their patterns determine the performance of the plasmonic nanostructures in most cases.

1.2 Electromagnetic Properties of Materials

1.2.1 Permittivity and permeability

In the theory of electromagnetic wave propagation presented here, the terms "permittivity" and "permeability" require clarification. For an optical medium, they define the specifications for how electromagnetic waves propagate through a given medium.

Regarding electromagnetic fields with very low frequencies, the auxiliary magnetic field is proportional to the magnetic field through a scalar permeability. However, at a very high frequency of oscillatory fields, a medium behaves like a dynamic system. The quantities will respond to each other with a phase delay described in the form of:

$$\mathbf{H}(\omega) = \mathbf{H}_0 e^{j\omega t}, \quad \mathbf{B}(\omega) = \mathbf{B}_0 e^{j(\omega t - \delta)}, \tag{1.21}$$

where δ denotes the phase delay between the magnetic fields $\mathbf{B}(\omega)$ and the corresponding auxiliary magnetic fields $\mathbf{H}(\omega)$. In electromagnetic wave theory, permeability describes the magnetization that a medium undergoes in response to an acting magnetic field. The permeability is defined as the ratio of the magnetic flux density to the magnetic field. In the conditions of phase delay existing, the permeability becomes a complex value:

$$\mu(\omega) = \frac{\mathbf{B}_0 e^{j(\omega t - \delta)}}{\mathbf{H}_0 e^{j\omega t}} = \frac{\mathbf{B}_0}{\mathbf{H}_0} e^{-j\delta}. \tag{1.22}$$

The complex permeability is a scalar for an isotropic medium and a tensor for an anisotropic medium. The complex permeability is translated from a polar coordinate form to a rectangular coordinate one as

$$\mu(\omega) = \mathbf{B}_0/\mathbf{H}_0(\cos\delta - j\sin\delta) = \mu_1(\omega) - j\mu_2(\omega), \tag{1.23}$$

where $\mu_1(\omega)$ and $\mu_2(\omega)$ refer to the real and imaginary parts of the complex permeability, respectively. Likewise, a phase delay emerges between a displacement field and an electric field when the frequency of an electromagnetic wave increases a certain amount. A complex permittivity is used to describe the polarization within a medium; it is expressed in the form of real and imaginary parts as

$$\varepsilon(\omega) = \mathbf{D}_0/\mathbf{E}_0(\cos\delta - j\sin\delta) = \varepsilon_1(\omega) - j\varepsilon_2(\omega). \tag{1.24}$$

In Eqs. (1.23) and (1.24), the real parts of $\varepsilon_1(\omega)$ and $\mu_1(\omega)$ are connected with the stored energy within a medium, while their imaginary parts are related to the energy dissipation. The complex permittivity and permeability are usually

a complicated description of dispersion phenomena (which will be discussed later in the optical properties of media). In standard international (SI) units, the permittivity and electric susceptibility are related to each other by

$$\varepsilon = \varepsilon_0(1 + \chi_e) = \varepsilon_0\varepsilon_r, \tag{1.25a}$$

$$\mu = \mu_0(1 + \chi_m) = \mu_0\mu_r. \tag{1.25b}$$

The actual permittivity is calculated by multiplying the relative permittivity and the electric permittivity of free space, just as the actual permeability is calculated.

For conventional materials working at visible and near-infrared wavelengths, the relative permeability is often considered to be unity because the magnetic susceptibility becomes negligible. Such conditions will simplify the description of the optical refractive index. The particular example is that the refraction and reflection properties of dielectric media, including both the magnitude and phase differences, are characterized by Fresnel equations, where neither permeability nor magnetic susceptibility is considered.

For the media with a unit value of magnetic permeability at optical frequency, the complex permittivity $\varepsilon(\omega)$ is determined experimentally through the reflectivity measurement. The complex refractive index is linked with the complex permittivity in the relations of

$$\varepsilon_1(\omega) = n_1^2 - n_2^2, \quad \varepsilon_2(\omega) = 2n_1k, \tag{1.26a}$$

where n_1 and n_2 are the real and imaginary value of the refractive index. The complex refractive index can then be determined by

$$n_1^2 = \frac{1}{2}\left(\varepsilon_1 + \sqrt{\varepsilon_1^2 + \varepsilon_2^2}\right), \quad n_2 = \frac{\varepsilon_2}{2n_1}. \tag{1.26b}$$

The imaginary part n_2 implies the energy absorption within a medium. Taking into account the permeability, the refractive index is expressed as

$$(n_1 + jn_2)^2 = (\varepsilon_1\mu_1 - \varepsilon_2\mu_2) + j(\varepsilon_1\mu_2 + \varepsilon_2\mu_1). \tag{1.27}$$

In the emerging fields of metamaterials and metasurfaces, nanostructure materials with a negative or near-zero refractive index are now being achieved. To achieve a negative real part of the refractive index, the imaginary part in Eq. (1.26) must be negative. It indicates that the negative refractive index can be achieved by a sufficiently large imaginary part of permittivity and/or permeability. It is not necessary that the real parts of both ε_1 and μ_1 are negative. In the visible light region, most noble metals have a large imaginary part of permittivity $|\varepsilon_1| >> |\varepsilon_2|$. This property enables them to be frequently used as materials to construct plasmonic devices.

1.2.2 Loss tangent

The curl equation of a magnetic field in Maxwell's equations is expressed as

$$\nabla \times \mathbf{H} = (\omega\varepsilon_2 + \sigma)\mathbf{E} + j\omega\varepsilon_1\mathbf{E}. \tag{1.28}$$

The real part ε_1 relates to the lossless component. The imaginary part ε_2 represents the energy loss component. The quantity $\sigma(\omega)$ is defined by a constitutive relation between the internal current density $\mathbf{J}$ and the electric field $\mathbf{E}$ as[4]

$$\mathbf{J}(\omega) = \sigma(\omega) \times \mathbf{E}(\omega). \tag{1.29}$$

The quantity $\sigma(\omega)$ refers to the conductivity, which is an intrinsic characteristic of a medium determined by the relation

$$\sigma(\omega) = \frac{N \times q_e^2}{m_e(1/\tau - j\omega)} = \frac{\sigma_0}{1 - j\omega\tau}, \tag{1.30}$$

where m_e is the effective mass of an electron, N refers to the electron number in unit volume, and γ is the damping ratio of the oscillating electrons. The term σ_0 is defined as

$$\sigma_0 = \frac{N \times q_e^2\gamma}{m_e}. \tag{1.31}$$

In the presence of an external electric field, the loss tangent is defined by the ratio of the loss component to the lossless component in Eq. (1.28):

$$\tan\eta = \frac{\omega\varepsilon_2 + \sigma}{\omega\varepsilon_1}. \tag{1.32}$$

For dielectric media with a small amount of energy loss, Eq. (1.32) approaches a simple form by the approximation $\tan\eta \approx \eta$. Likewise, the electric and magnetic loss tangent of a transparent dielectric medium are quantified similarly in the subsequent definitions of

$$\tan\eta_m = \frac{\mu_2}{\mu_1}, \quad \tan\eta_e = \frac{\varepsilon_2}{\varepsilon_1}. \tag{1.33}$$

The loss tangent quantifies the inherent dissipation of electromagnetic energy due to the imaginary parts of its dielectric properties.

1.2.3 Penetration depth and skin depth

When an electromagnetic wave impinges on the surface of a medium, part of the wave reflects from the surface, and a small fraction of the wave will penetrate into the medium. The penetrated wave interacts with the atoms and electrons of the medium, which will contribute significantly to the SPPs. The

Beer–Lambert law indicates that the intensity of the electromagnetic wave penetrating into a medium falls off exponentially from the air–medium interface.[5,6] The penetrating intensity is described by the exponential function of

$$I(z) = I_0 e^{-z\alpha(\omega)}, \tag{1.34}$$

where z denotes the penetrating distance from the air–medium interface. The attenuation factor $\alpha(\omega)$ depends on the frequency of the impinging electromagnetic wave. $I(z)$ and I_0 are the penetrating intensity at depth z and the initial intensity at the interface, respectively.

A penetration depth is defined as the penetrating distance at which the wave intensity decays to the value $1/e$ ($\sim$0.37) of its initial intensity. Conventionally, the penetration depth is defined in terms of the attenuation factor

$$\delta_p(\omega) = \alpha^{-1}(\omega). \tag{1.35}$$

In the condition of a static electric field, the current density within a conductor wire distributes uniformly throughout the cross-section of the wire. On the other hand, when the frequencies of an electromagnetic wave increase, the current flow tends to approach the surface of a conductor medium. The penetration depth decreases sharply with the frequency increasing. A similar concept of skin depth is used to describe the current flow in a conductor medium.

The energy of a wave propagating through a particular medium is proportional to the square of its intensity. From the energy point of view, when the amplitude of the energy attenuates to $1/e$ of its initial energy, the depth below the air–medium interface is defined to be the skin depth of a medium, which relates with the penetration depth as

$$\delta_s(\omega) = \frac{1}{2\alpha(\omega)} = \frac{1}{2}\delta_p(\omega). \tag{1.36}$$

The skin depth is one characteristic of an electromagnetic medium whose value also depends on the frequencies of electromagnetic waves.

For a solid wire, as shown in Fig. 1.1(a), the current concentrates on the outer surface. In the case of a micro-strip layout, as shown in Fig. 1.1(b), the current concentrates nearest to the substrate dielectric material. As suggested in Eqs. (1.34)–(1.36), there is an exponential decay of the electromagnetic amplitude or energy. A more gradual penetration of the electromagnetic energy into the metal occurs, although the high-amplitude regions are depicted as having a sharp interface between the "bulk" metal and the near-surface regions.

In the case of a poor conductor, such as a dielectric medium, or in the presence of an electromagnetic wave with a high frequency, a general calculation with a more exact solution is expressed as:

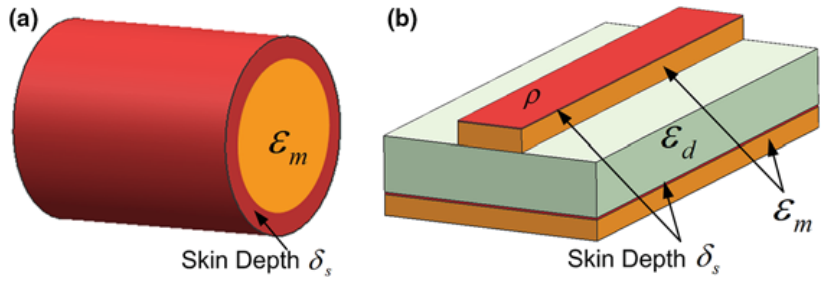

Figure 1.1 Schematic diagram of current distribution within conductor media. (a) Current flows through a cylindrical conductor wire and (b) in metal strips of a composite structure. The dark color indicates regions with a higher electromagnetic amplitude. The regions of metal and dielectric materials are labeled by their permittivity.

$$\delta_S(\omega) = \sqrt{2\rho_c/\omega\mu}\sqrt{\sqrt{1+(\rho_c\varepsilon\omega)^2}+\rho_c\varepsilon\omega}, \tag{1.37}$$

where ρ_c is the resistivity of conductors (ohm $\times$ meter), ε and μ are the absolute permittivity and permeability, and ω is the angular frequency of an impinging wave.

When the frequency is much lower than the value $1/\rho_c\varepsilon$, the calculated value of the square root in Eq. (1.37) approaches unity. In such a case, the skin depth of a medium is estimated by

$$\delta_S(\omega) = \sqrt{2\rho_c/\omega\mu}. \tag{1.38}$$

Equation (1.38) applies to the materials with a larger skin depth at a much lower frequency. For instance, it would hold true for metal copper at frequencies below 10^{18} Hz. Nevertheless, for materials with a rather low conductivity corresponding to a shallow skin depth or materials that work at frequencies much higher than the value of $1/\rho_c\varepsilon$, the skin depth approaches an asymptotic value rather than monotonously decreasing:

$$\delta_S(\omega) \approx 2\rho_c\sqrt{\varepsilon/\mu}. \tag{1.39}$$

Equation (1.39) applies to materials with low conductivity working at high frequencies. For instance, the skin depth of undoped silicon is ~40 m irradiated by waves with a frequency of 100 kHz. When the frequency of the wave increases to the microwave region (1 GHz to 30 GHz), with a corresponding wavelength of 30 cm to 10 mm, the skin depth of the same undoped silicon closes to the asymptotic value of 11 m.

For good conductors working at a sufficiently high frequency, the skin depth becomes tiny, which implies that the electric charges approach the air–medium interface. This characteristic deserves attention with respect to nanostructures and plasmonic devices. In such conditions, the dimensional

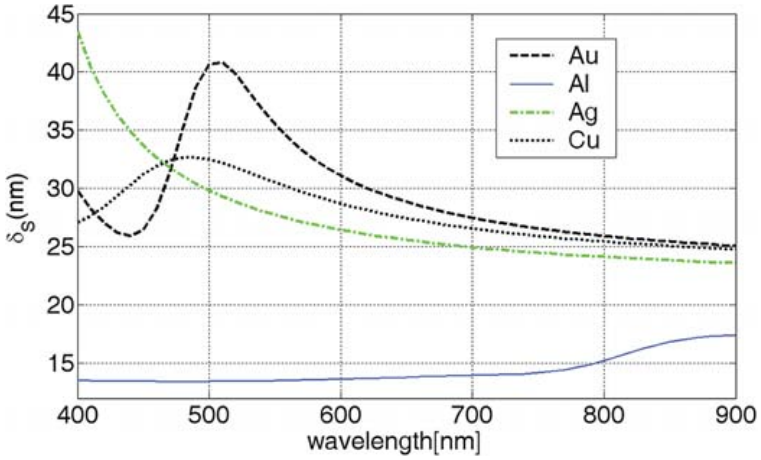

Figure 1.2 Spectral dependence of skin depth for frequently used metals when a plane wave impinges on the air–metal interface at normal incidence. The parameter values for the estimation of skin depth are adopted from Rodrigo et al.[7]

sizes approach the scale of skin depth at the visible or near-infrared wavelength.

Figure 1.2 displays the skin depth spectra of a few frequently used noble metals. For those metals with good conductivity working at the visible and near-infrared wavelength, a majority of the electric current flows within a thin region, which is less than tens of nanometers beneath the air–medium interface. For example, the skin depth of gold at a wavelength of 500 nm reaches a peak value of ~40 nm and then decays to a skin depth of ~25 nm at a near-infrared wavelength of 900 nm.[7] Table 1.1 shows the skin depths of a few frequently used noble metals at certain frequencies.

In the condition of normal incidence of an electromagnetic wave, the attenuation factor is also proportional to the imaginary part of the refractive index. The following relationship holds true for the attenuation factor defined by Eq. (1.34):

Table 1.1 Bulk resistivity and skin depths of frequently used noble metals.

Material	Bulk Resistivity ($\Omega \times 10^{-8}$ m)	Skin Depth (μm)					
		1 MHz	10 MHz	100 MHz	1 GHz	10 GHz	100 GHz
Aluminum	2.65	81.9	25.9	8.19	2.59	0.819	0.259
Copper	1.69	65.4	20.7	6.54	2.07	0.654	0.207
Gold	2.2	74.7	23.6	7.47	2.36	0.747	0.236
Graphite	783.7	1409	446	141	44.6	14.1	4.46
Lead	20.6	228	72.2	22.8	7.22	2.28	0.722
Nickel	6.9	9.3	3.0	.093	0.30	0.093	0.030
Silver	1.63	64.3	20.3	6.43	2.03	0.643	0.203
Titanium	54	370	117	37.0	11.7	3.70	1.17

$$\alpha(\omega) = 2n_2(\omega)\omega/c, \tag{1.40}$$

where $n_2(\omega)$ denotes the imaginary part of the complex refractive index. So both the penetration depth and the skin depth can be estimated directly from the imaginary part of the refractive index.

1.3 Optical Properties of Metals

The obvious difference between the optical responses of metals and dielectrics enables them to be fundamental elements for designing plasmonic devices. In plasmonic optics, metals are incorporated into nanostructures for the emerging characteristics of the SPPs and other light–matter interactions. This section emphasizes the optical performance of metals following the theory model for free electrons and bound electrons, pursing the approach for realizing the optimal plasmonic materials at the nanoscale.

1.3.1 Free electrons and interband transitions

To investigate the optical properties of solid materials, we start from the conduction electrons within metals and semiconductors. There are conduction electrons and interband transition electrons within media.[8] Conduction electrons oscillate freely within bulk metals, whereas interband transition electrons can only be induced by the incident photons with energy greater than the bandgap energy. Both of them determine the complex dielectric functions, such as permittivity and permeability, that describe the optical properties of metals.

Conduction electrons are associated principally with free carriers within metal media. The mechanism of this intraband electronic conduction is explained by the Drude theory. For the electrons bounded into occupied states, the electrons can jump up from the energy band below the Fermi level to the conduction bands only when excited by photons with enough energy. This process of electron transition is called an interband transition, as shown in Fig. 1.3, where no energy states are allowed between the filled valence band and an empty conduction band. Higher-energy photons can promote the

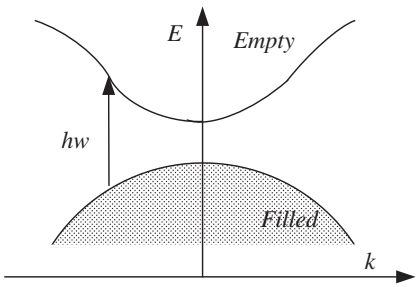

Figure 1.3 Schematic diagram of an interband transition of bound electrons.

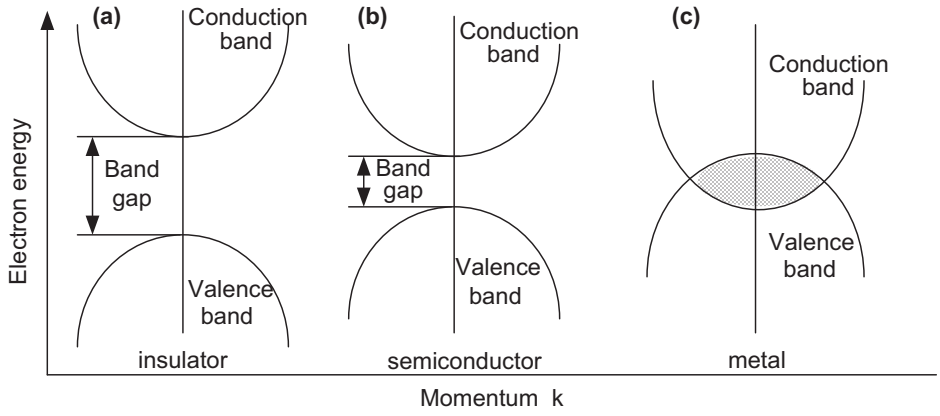

Figure 1.4 Typical band structures of insulator, semiconductor, and metal materials.

bound electrons from lower energy levels into the empty conduction band. Such interband transitions often occur in the light–matter interaction at the visible or infrared wavelength. The process of interband transitions produces a pair carrying an electron and a hole. In practice, only a portion of interband transitions contribute to the conduction mechanism, which corresponds to a few specific energy bands.

Solid materials have their own energy band structures, as shown in Fig. 1.4. The various characteristics of energy band structures determine the multitudinous electrical properties of solid materials. For example, the bandgap of the undoped silicon is ~1.1 eV, whereas diamond has a 5-eV bandgap.[8]

The electron transitions in a medium between the energy levels determine its dielectric function and then the optical properties. For an insulator, a large gap exists between the valence and conduction band, as shown in Fig. 1.4(a). No electron jumps into the conduction band without enough energy. The electrons in insulators are rare compared with semiconductors. As shown in Fig. 1.4(b), semiconductor materials have basically the similar band structure as insulators have; however, the bandgap E_g of semiconductors is much smaller than that of insulators. The number of electrons within semiconductors would be increased greatly after the photons are excited to achieve electron conduction. A deep doping will greatly reduce their semiconductor band gaps, which enables more available electrons to contribute to the interband transition.

In the case of metal materials, however, the energy bands overlap each other or, in certain conditions, are partially filled, as shown in Fig. 1.4(c). The overlap of valence bands and free conduction bands enables electrons to move freely across the bands. Interband transitions within metals can easily occur under the excitation by photons with low frequencies. The free carriers contribute largely to the conduction. Although the densities of free carriers are low, the high mobility of free carriers contributes to the conductivity of metals.

For semiconductors, the energy gaps are often temperature dependent. For example, the energy gap of germanium depends on the temperature as

expressed by an empirical formula $E_g = 0.742 - 4.8 \times 10^{-4}\, T^2/(T + 235)(\text{eV})$,[12] where T is the temperature in kelvin. Also, the energy gap bands of semiconductors are determined by the doping density of either donors or acceptors. For example, deep doping will increase the band gap of germanium by the following approximation formula:[12]

- For n-type Ge, $\Delta E_g = 4.31 \times 10^{-9}\sqrt{N_d} + 8.67 \times 10^{-6}\sqrt[3]{N_d} + 8.14 \times 10^{-5}\sqrt[4]{100 \times N_d}(\text{eV})$,
- For p-type Ge: $\Delta E_g = 5.77 \times 10^{-9}\sqrt{N_a} + 8.21 \times 10^{-6}\sqrt[3]{N_a} + 9.18 \times 10^{-5}\sqrt[4]{100N_a}(\text{eV})$,

where N_d and N_a are the doping density of the donor and the acceptor, respectively. Increases to the doping concentration enlarge the energy bandgap, slowly at first and then sharply.

1.3.2 Harmonic oscillator model

The Drude and Lorentz models were developed in the electronic kinetic theory of microscopic electrons to explain the optical properties of materials. The models were further extended into the Drude–Lorentz model, which describes the dielectric properties of solid materials.

As shown in Fig. 1.5, a charge within a medium is treated as a harmonic oscillator, which is bounded with a nucleus. Under the excitation of an incident electromagnetic wave, the oscillator will oscillate in the oppositional phase relative to the electric field. Here, only the conditions of a nonpolar molecule are discussed, in which no dipole moment exists in the electron–nucleus system.

From the dynamic point of view, the charge oscillation will lead to the charge redistribution, which will create an additional induced electric field. The induced field will restore the charge to its equilibrium position.

Newton's second law states that the production of mass multiplying the acceleration equals the sum of the forces applied on the system. The governing equation for the system in Fig. 1.5 is

$$m_e \frac{\partial^2 \mathbf{r}(t)}{\partial t^2} = \sum_i \mathbf{F}_i = \mathbf{F}_E(t) + \mathbf{F}_D(t) + \mathbf{F}_S(t), \tag{1.41}$$

where m_e denotes the effective mass of a charge, $\mathbf{r}(t)$ is the instantaneous distance deviation from its equilibrium position. $\mathbf{F}_E(t) = q_e \mathbf{E}_0 e^{-j\omega t}$ is the local

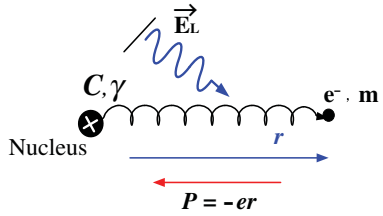

Figure 1.5 Schematic of electron motion for a harmonic oscillator model in an external electric field.

force attributed to an external alternating electric field, where q_e is the electric quantity of an charge, and $\mathbf{E}_0$ is the electric field amplitude. $\mathbf{F}_S(t) = -q \times \mathbf{r}(t)$ is the restoration force proportional to the distance deviation from the equilibrium position $q = m_e\omega_0{}^2$, where ω_0 is the natural frequency of the bound oscillator. The restoration forces are zero for free electrons within conductor materials because they are not bound to a particular nucleus. $\mathbf{F}_D = -\gamma[\partial\mathbf{r}(t)/\partial t]$ refers to the damping force due to the collision energy loss, where γ is the damping coefficient in hertz.

Note that the restoration and damping forces are negative because they are opposite the direction of the motion. Thus, the differential form in Eq. (1.41) becomes

$$m_e \frac{\partial^2 \mathbf{r}}{\partial t^2} + m_e \gamma \frac{\partial \mathbf{r}}{\partial t} + m_e \omega_0^2 \mathbf{r} = q_e \mathbf{E}_o e^{-j\omega t}. \tag{1.42}$$

In general, the condition $\gamma > 0$ leads to the damping attenuation.

The instaneous distance solution of Eq. (1.42) for a monochromatic electric field is solved to be

$$\mathbf{r}(\omega) = \frac{q_e}{m_e(\omega_0^2 - \omega^2 - j\gamma\omega)} \mathbf{E}(\omega). \tag{1.43}$$

The polarization of a dipole moment causes the electric-charge timing distance to deviate from its balanced position. The gross polarization of the whole electrons in unit volume is

$$\mathbf{P} = N_e q_e \mathbf{r}(\omega) = \frac{N_e q_e^2}{m_e(\omega_0^2 - \omega^2 - j\gamma\omega)} \mathbf{E}(\omega), \tag{1.44}$$

where N_e is the electron density per unit volume. The induced dielectric polarization density is proportional to an electric field by the constant electric susceptibility as

$$\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E}(\omega) = \varepsilon_0(\varepsilon_r - 1)\mathbf{E}(\omega). \tag{1.45}$$

The susceptibility is $\chi_e = \varepsilon_r - 1$. The permittivity function is thus obtained:[11]

$$\varepsilon(\omega) = 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - j\gamma\omega}, \tag{1.46}$$

where ω_p is the plasma frequency of the bulk metal, which is calculated by[5]

$$\omega_P^2 = \frac{N_e q_e^2}{m_e \varepsilon_0}. \tag{1.47}$$

The dielectric function [Eq. (1.46)] presents explicitly the real and imaginary parts in the form of

$$\varepsilon_1(\omega) = 1 + \frac{\omega_p^2(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}, \tag{1.48a}$$

$$\varepsilon_2(\omega) = \frac{\gamma\omega_p^2\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}. \tag{1.48b}$$

1.3.3 Drude model and Lorentz model

The interactions between metals and electromagnetic waves are firstly determined by the collective movement of free electrons. In this case, the electrons are not bound to any particular nucleus, which are considered to move about freely around the metal lattice in the absence of a restoration force. The motion equation of a free electron in an alternating electric field is described by

$$m_e\frac{d^2\mathbf{r}}{dt^2} + m_e\gamma\frac{d\mathbf{r}}{dt} = -q_e\mathbf{E}_o e^{-j\omega t}, \tag{1.49}$$

where q_e is the electric charge of a free electron. the damping effect γ is proportional to the Fermi velocity $\gamma = \nu/l = 1/\tau$, where ν denotes the Fermi velocity, and l is the mean free path of an electron between successive collision events. The relaxation time τ is the averaged interval time between subsequent collisions of an electron. In general, the relaxation time is about 10^{-14} s. The solution of instaneous distances in Eq. (1.49) for a monochromatic electric field is solved to be

$$\mathbf{r}(\omega) = \frac{q_e}{m_e(\omega^2 + j\omega\gamma)}\mathbf{E}(\omega). \tag{1.50}$$

The polarization density is the total dipole moment per unit volume. The gross polarization of all of the electrons in the unit volume is

$$\mathbf{P} = N \times q_e \times \mathbf{r}(\omega) = \frac{N \times q_e^2}{m_e(\omega_0^2 + j\gamma\omega)}\mathbf{E}(\omega), \tag{1.51}$$

where N is the electron density per unit volume. A comparison of Eqs. (1.45) and (1.51) leads to the frequency-dependent permittivity of metals as

$$\varepsilon(\omega) = 1 - \frac{\omega_P^2}{\omega^2 + j\omega\gamma}, \tag{1.52}$$

where ω_p is the plasma frequency of the bulk media, which has a similar physical mechanism as the definition in Eq. (1.47). At the plasma frequency, the electromagnetic response of a material changes from the metallic behaviors to those of dielectric materials. At the frequency below the plasma

frequency, the optical properties of a medium exhibit a metal-like behavior. However, at a frequency greater than the plasma frequency, their optical properties look more like those of dielectric media. Equation (1.52) describes the contribution of free electrons to the permittivity of metals. The permittivity in Eq. (1.52) is rewritten into the real and imaginary parts

$$\varepsilon(\omega) = \varepsilon_1 + j\varepsilon_2 = \left(1 - \frac{\omega_P^2}{\omega^2 + \gamma^2}\right) + j\frac{\gamma\omega_P^2}{\omega(\omega^2 + \gamma^2)}. \tag{1.53}$$

In the conditions of $\omega >> \gamma$, Eq. (1.53) is simplified to be

$$\varepsilon(\omega) \approx 1 - \frac{\omega_P^2}{\omega^2} + j\frac{\omega_P^2}{\omega^3}\gamma. \tag{1.54}$$

The permittivity of metal gold is calculated by Eq. (1.53), ranging from visible to infrared wavelengths. As shown in Fig. 1.6, the real part of the permittivity for metal gold remains negative throughout the wavelengths in which the frequencies are less than the plasma frequency. The negative permittivity results in an imaginary part of the refractive index. For metals, the imaginary part of the permittivity implies an energy dissipation that is relevant with the thermal motion of electrons.

Figure 1.6 indicates that, in the infrared wavelengths longer than ~650 nm, the Drude model gives an accurate permittivity of gold, and the experimental data clearly agree well with the Drude theory. However, there is an obvious deviation of the imaginary part in the visible wavelengths shorter than 650 nm. The measured imaginary part increases much more sharply than that predicated by the Drude theory. This difference is attributed to the fact that the interband transitions of the bound electrons excited by the photons

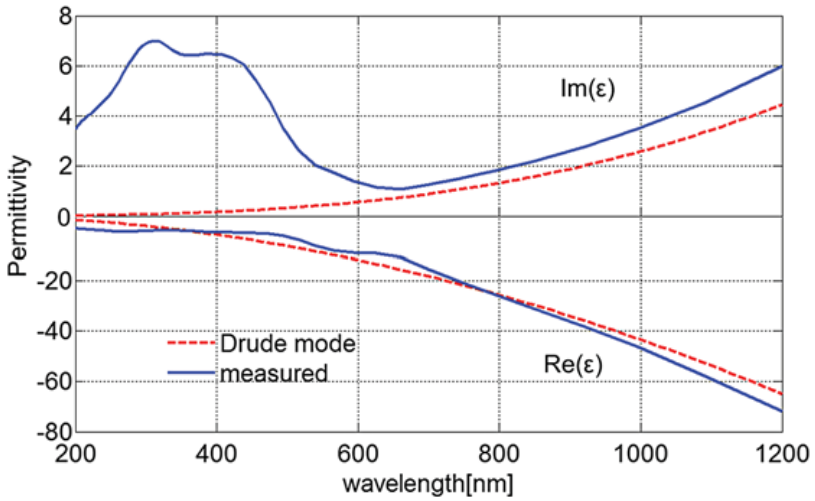

Figure 1.6 Permittivity of metal gold estimated by the Drude model (dotted line) and the measurement (solid line). The measured data come from Ref. [10]. The parameter values in the Drude model are adopted from Table 1.2.

with higher energy have not been taken into account in the Drude model, where, however, the interband transitions become significant.

The contribution from the interband transitions of bound electrons to the dielectric permittivity looks like the corresponding resonance in dielectric media. These transitions are obtained from the governing equations without consideration of the free electrons as

$$\varepsilon(\omega) = 1 + \frac{\omega_1^2}{\omega_0^2 - \omega^2 - j\gamma\omega}, \tag{1.55}$$

where ω_0 is the oscillation frequency of a bound electron under an electric field, and ω_1 and γ are related to the corresponding frequency and damping of a bound electron, respectively.

Figure 1.7 shows the dielectric permittivity of metal gold considering only the interband contribution of bound electrons. The real part shows a dispersion behavior while the imaginary part displays a clear resonant behavior. The vertical dotted line represents the resonant frequency ω_0 that is consistent with the zero value of the second part on the right side of Eq. (1.52). The resonant frequency ω_0 also agrees well with the peak value of the imaginary part, so it is sometimes called the absorption frequency. For the real parts, there is a peak value in the lower frequency and a valley value in the higher frequency that is greater than the resonant frequency. The frequency region between the two extreme values is called the abnormal dispersion region, where the permittivity values decrease with the increasing frequency.

For a Lorentz oscillator, the real part of permittivity is a positive value in the range less than the resonance frequency, while there is a negative valley at the slightly higher frequency than the resonance frequency. Another feature in Fig. 1.7 is that the dielectric permittivity has a nonzero asymptotic value when

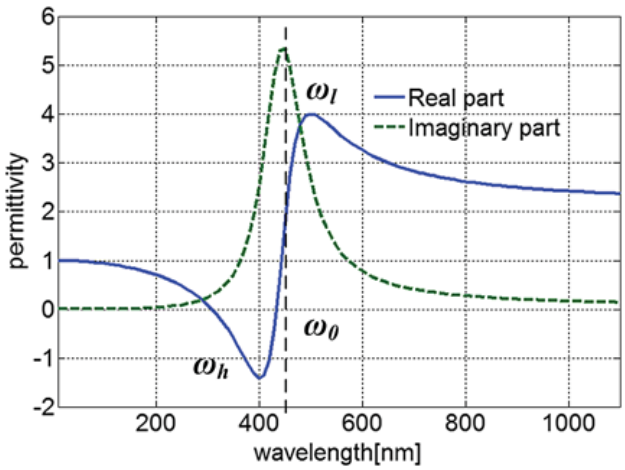

Figure 1.7 Contribution of the bound electrons in metal gold to the permittivity. The values of the parameters used are $\omega_1 = 4.55 \times 10^{15}$ s^{-1}, $\gamma = 9.10 \times 10^{16}$ s^{-1}. The central resonance wavelength is 450 nm ($\omega_0 = 4.19 \times 10^{17}$ s^{-1}).

the wavelength increases much longer than the resonance wavelength. After all, if the interband electron transitions within media have been taken into account, such an asymptotic constant should approach one summary value.[11]

If the higher order of interband electrons transition is taken into account, the dielectric permittivity is expressed as the model superposition of the Lorentz model as

$$\varepsilon(\omega) = 1 + \sum_{i}^{K} \frac{f_i \omega_i^2}{\omega_i^2 - \omega^2 - j\gamma_i \omega}, \tag{1.56}$$

where i refers to the resonant modes, ω_i corresponds to the resonance frequencies, f_i indicates the weighting coefficients, and γ_i is the damping effect.

Each electron is assumed to contribute to the dipole polarization. The atomic nucleus also makes a slight contribution, which is ignored because the mass of the nucleus is much greater than that of the electron. The aforementioned formulae of dielectric permittivity in the Drude theory are derived under the approximation of infinite nucleus mass. Another assumption is that only one electron works in one electronic-dynamics system. In such conditions, the electric properties of commonly used noble metals and their corresponding bulk plasma frequencies (as well as the damping constant) are listed in Table 1.2.

Figure 1.8 compares the characteristics of metals using the Drude model and the Lorentz model. At the plasma frequency in the Drude model, the dielectric function that passes through zero corresponds to the collective

Table 1.2 Electric properties of noble metals and bulk plasma frequencies.

Metal	Plasma Frequency (eV)	Plasma Frequency (10^{15} s^{-1})	Damping Constant (10^{15} s^{-1})
Au	9.1	13.8	0.011
Ag	9.2	14.0	0.032
Al	15.1	22.9	0.92
Cu	8.8	13.4	0.14

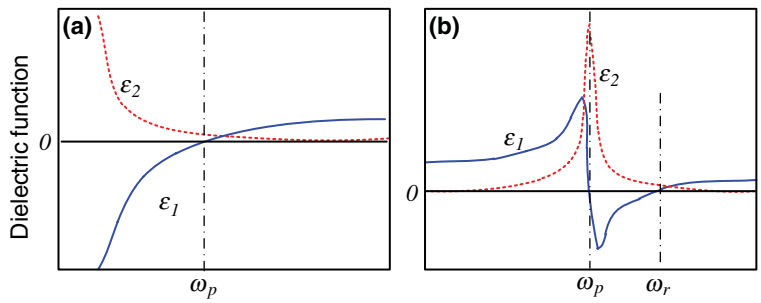

Figure 1.8 Components of the complex dielectric functions for (a) metals described using the Drude model and (b) dielectrics using the Lorentz equations.

electronic oscillation within metals. Figure 1.8(b), described by the Lorentz equations, shows a resonant frequency that corresponds to the bound electrons, which is consistent with the absorption peak.

1.3.4 Drude–Lorentz model

Both bound electrons and free ones contribute to the optical properties of a general metallic medium. Therefore, the resembling complex dielectric permittivity contains both the Drude component for the intraband effect and the Lorentz term for the interband transition in the form of the Drude–Lorentz model

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + j\omega\gamma_d} + \sum_{i=1}^{K} \frac{f_i \omega_i^2}{\omega_i^2 - \omega^2 - j\omega\gamma_i}, \tag{1.57}$$

where K is the total number of higher-energy oscillators with the corresponding resonant frequency ω_j, weighting coefficient f_j, and lifetime $1/\gamma_j$. In the second term in Eq. (1.57), ω_p refers to the plasma frequency relevant to the intraband transitions with a damping constant γ_d.

As shown in Figs. 1.6 and 1.7, the interband transition has a nonzero asymptotic value of the dielectric permittivity at a wavelength much further away from the resonance wavelength. It is noteworthy that, within noble metals, there are usually multiple interband transitions due to the various band structures of bound electrons. Therefore, after taking all pertinent transitions into account, we use a constant offset value to denote the constant offset when working at much longer wavelengths than the resonance wavelength. Thus, the modified Drude–Lorentz model considering the fundamental oscillator is

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + \gamma^2} + j\frac{\omega_p^2 \gamma}{\omega(\omega^2 + \gamma^2)}. \tag{1.58}$$

For example, the interband offset for gold is about $\varepsilon_\infty = 1.2$ from Eq. (1.58) by setting $\omega \to 0$.

Optimal values of typical metals for the Drude–Lorentz model are listed in Table 1.3. These include noble metal materials (Au, Ag, and Cu), aluminum, and transition metals (Cr, Ni, Pd, Pt, and Ti). These metals have been frequently used in optoelectronic devices. The measurements of these optical data are strongly dependent on the experimental conditions (in terms of geometry morphology and the fabrication process), so the relevant optical data must be carefully chosen.[12]

A correction may be made to the model data shown in Fig. 1.5 obtained by the Drude–Lorentz model. Taking into account the contribution from both free electrons and a single interband transition, the dielectric function of gold progresses toward the measured results. Only one interband transition has

Table 1.3 Parameter values for the Drude–Lorentz model. Adapted from Ref. 12 with permission from the American Physical Society, © 1998.

Parameters	Ag	Au	Cu	Al	Be	Cr	Ni	Pt	Ti	W
f_0	0.845	0.760	0.575	0.523	0.084	0.168	0.096	0.333	0.148	0.206
γ_d	0.048	0.053	0.030	0.047	0.035	0.047	0.048	0.080	0.082	0.064
f_1	0.065	0.024	0.061	0.227	0.031	0.151	0.100	0.191	0.899	0.054
γ_1 (eV)	3.886	0.241	0.378	0.333	1.664	3.175	4.511	0.517	2.276	0.530
ω_1 (eV)	0.816	0.415	0.291	0.162	0.100	0.121	0.174	0.780	0.777	1.004
f_2	0.124	0.010	0.104	0.050	0.140	0.150	0.135	0.659	0.393	0.166
γ_2	0.452	0.345	1.056	0.312	3.395	1.305	1.334	1.838	2.518	1.281
ω_2	4.481	0.830	2.957	1.544	1.032	0.543	0.582	1.314	1.545	1.917
f_3	0.011	0.071	0.723	0.166	0.530	1.149	0.106	0.547	0.187	0.706
γ_3	0.065	0.870	3.213	1.351	4.454	2.676	2.178	3.668	1.663	3.332
ω_3	8.185	2.969	5.300	1.808	3.183	1.970	1.597	3.141	2.509	2.580
f_4	0.840	0.601	0.638	0.030	0.130	0.825	0.729	3.576	0.001	2.590
γ_4	0.916	2.494	4.305	3.382	1.802	1.335	6.292	8.517	1.762	5.836
ω_4	9.083	4.304	11.18	3.473	4.604	8.775	6.089	9.249	19.43	7.498
f_5	5.646	4.384	—	—	—	—	—	—	—	—
Γ_5	2.419	2.214	—	—	—	—	—	—	—	—
ω_5	20.29	13.32	—	—	—	—	—	—	—	—

been considered in Fig. 1.9, so the mode curves still fail to reproduce the experimental results at wavelengths below 500 nm. For a practical metal, there are usually multiple interband transitions in the short-wavelength spectrum because the higher energy of photons promotes interband transitions of bound electrons.

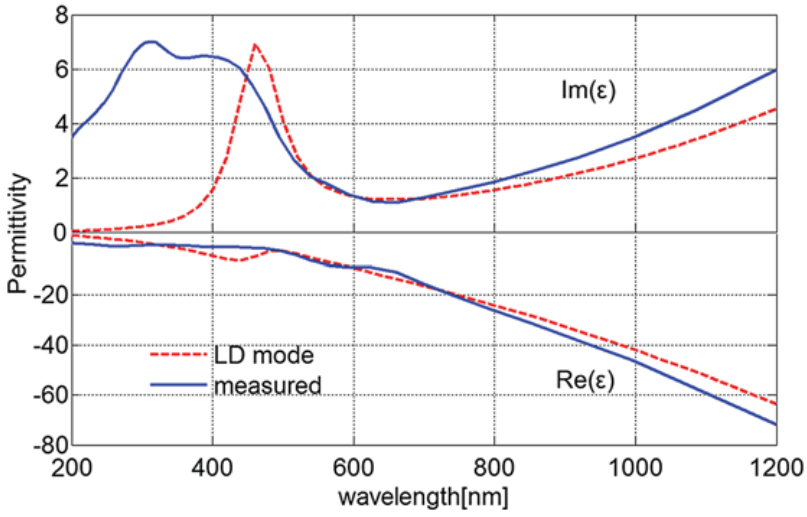

Figure 1.9 Permittivity of gold estimated by the Drude–Lorentz model (dotted line) considering the contribution from both the free electron and a single interband transition. The experimental values (solid line) from Ref. [10] are compared. The model parameters for Au are taken from Table 1.3. Note the different scales for the real part in the lower half and the imaginary part in the upper half.

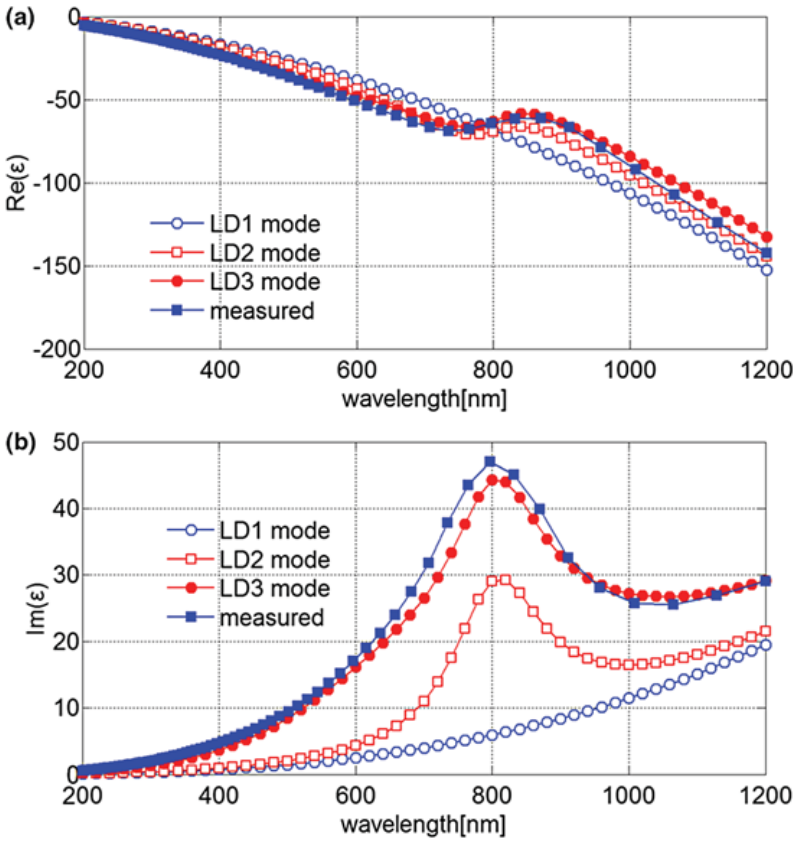

Figure 1.10 (a) Real and (b) imaginary parts of the permittivity for metal aluminum at wavelengths ranging from 200 nm to 1200 nm. The first, second, and third oscillators of interband transitions are integrated into the effect. The experimental values are taken from Ref. [10], and the parameter values for the model are from Table 1.3.

When more interband transitions are integrated into Eq. (1.57), the calculated date would progress closely to the experimental data. For example, the dielectric functions of aluminum are shown in Fig. 1.10. Both the real part and the imaginary part of the permittivity approach the experimental results when the higher-energy interband transitions are integrated into the effect. Some rapid variations occur around 800 nm, which corresponds to the lowest-lying interband transition of the bound electrons in the outer shells.

The dielectric permittivity of metal silver and the complex refractive index are plotted in Fig. 1.11. The contribution of both free and bound electrons is considered. The measured data are consistent with the results estimated by the Drude–Lorentz model. At optical frequencies, the imaginary part of the dielectric permittivity is positive, whereas the real part has negative values at wavelengths longer than the resonant wavelength. The dielectric function of $\mathrm{Im}(\varepsilon) << \mathrm{Re}(|\varepsilon|)$ holds true. However, the imaginary part of the refractive index is incrementally greater than the corresponding real part, which is slightly greater than zero.

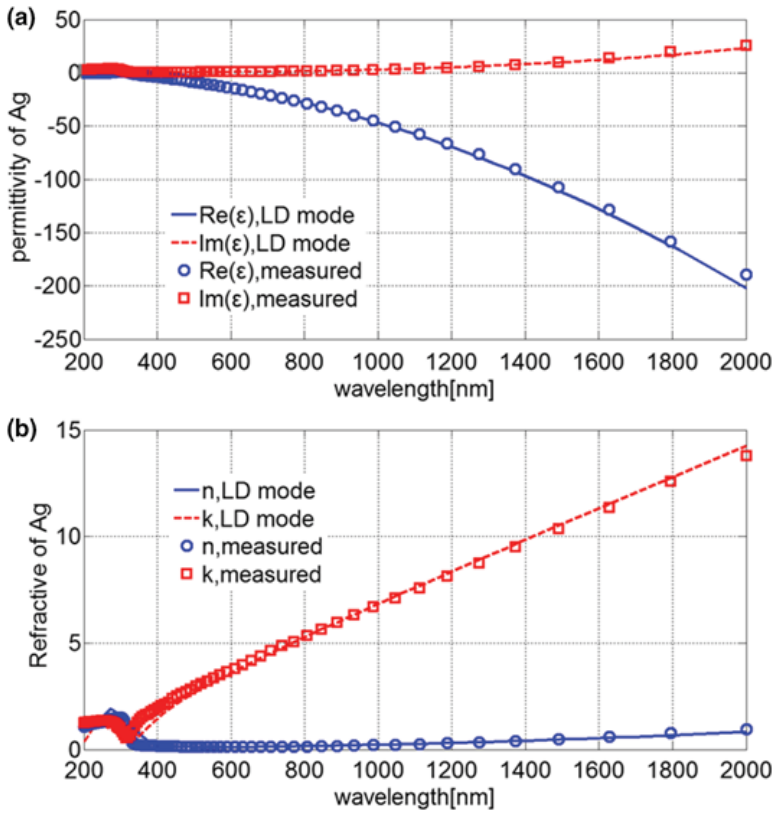

Figure 1.11 (a) Permittivity and (b) corresponding refractive index of metal silver. The Drude–Lorentz model in (a) considers the contribution of both free and bound electrons. The experimental data come from Ref. [10].

The reflection coefficient in Fresnel equations is presented by $(n_1 - n_2)/(n_1 + n_2)$, where n_1 is the real part and n_2 is the imaginary part of the refractive index. In the conditions of $n_2 >> n_1$ at the wavelengths longer than the corresponding plasma frequency of the interband transition, the reflection coefficient approaches an absolute value of unity, which indicates that most of the parts of light are reflected at the interface between a dielectric medium and silver metal.

The dielectric functions for the metals of Au, Cu, Ni, and Cr are compared in Fig. 1.12. The imaginary parts of the dielectric permittivity of these metals remain positive at optical wavelengths ranging from 200 nm to 900 nm, whereas the real parts of dielectric functions are distinctly negative. An increasing absolute value of the real part indicates that most of the light reflects at the dielectric–metallic interface. The imaginary part of metals Ni and Cr increase prominently after a wavelength of 500 nm while their real parts have a faint value. The reducing reflection coefficient means that the light will be absorbed in those wavelengths.

Following a similar methodology as the deducing of the Lorentz model, the effective permeability can be obtained as

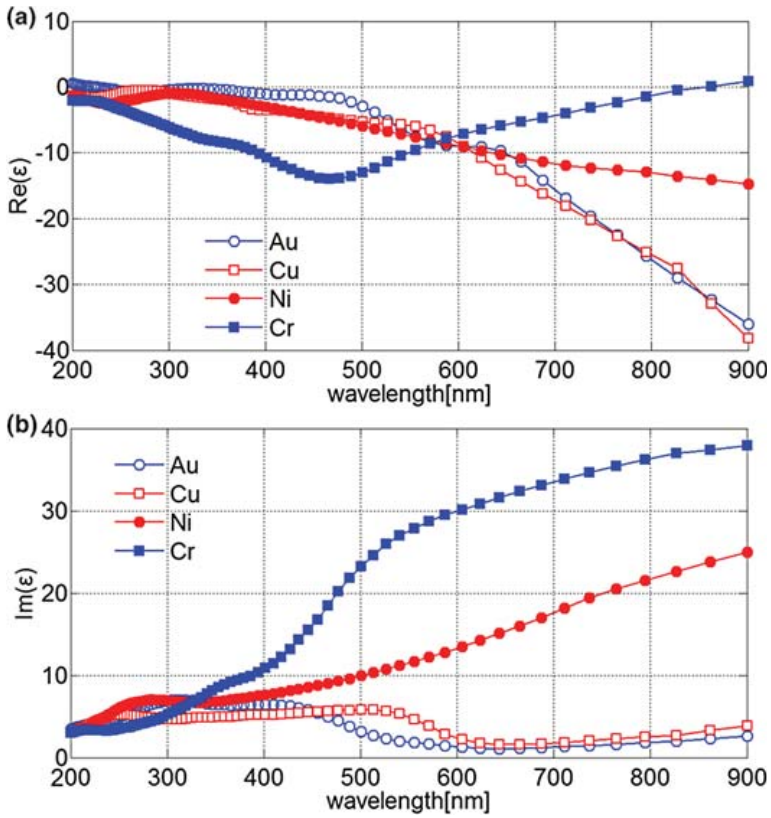

Figure 1.12 (a) Real part and (b) imaginary part of the dielectric functions for metals Au, Cu, Ni, and Cr at wavelengths ranging from 200 nm to 900 nm. Fitting lines are shown for clarity. The experimental data come from Ref. [10].

$$\mu_r(\omega) = 1 + \frac{\omega_{pm}^2}{\omega_{om}^2 - \omega^2 - i\gamma_m\omega}, \tag{1.59}$$

where ω_{pm}, ω_{om}, and γ_m are the plasma frequency, resonant frequency, and the damping coefficient in the magnetization of a magnetic dipole. For multiple orders of magnetic resonance, Eq. (1.59) is extended into a multiple form as

$$\mu_r(\omega) = 1 + \sum_{k=1}^{M} \frac{\omega_{pmk}^2}{\omega_{omk}^2 - \omega^2 - i\gamma_{mk}\omega}. \tag{1.60}$$

The resonant components ω_{pmk} correspond with k orders of resonances of magnetic dipoles. The relative permeability of metals and dielectric media is often assumed to be unity. However, the effective permeability for composite structures made of dielectric or metal materials can be tuned to values not accessible in natural materials.[13] A large imaginary component of permeability helps one obtain composite structures with a negative refractive index.[14] The extension of relative permeability allows for arbitrary spectral dispersion.

1.4 Optical Properties of Dielectric Materials

Both dielectric and metallic materials constitute materials used to design nanostructures in plasmonic optics. In conventional geometrical optics, the vast majority of functional parts are made from dielectric materials, such as crystalline and glassy materials. The characteristics of dielectric materials enable them to be useful for various band energies.

1.4.1 Dielectric function of dielectric media

Figure 1.13 shows the dielectric functions of silicon and germanium. Likewise, two resonances for these semiconductors appear indistinctly at the visible wavelength region, with positive imaginary values, and with two near resonances crossing zero points at a wavelength of ~300 nm. In the conditions of wavelengths approaching 900 nm, the real parts of the dielectric functions approach asymptotic values, and the imaginary values approach zero limits.[9]

Photonic bandgap structures are made of semiconductor materials with periodic structures to modulate the refractive index, where the dimensional sizes are comparable to the wavelength of light. Such electromagnetic bandgap structures that work in the microwave or visible range are called

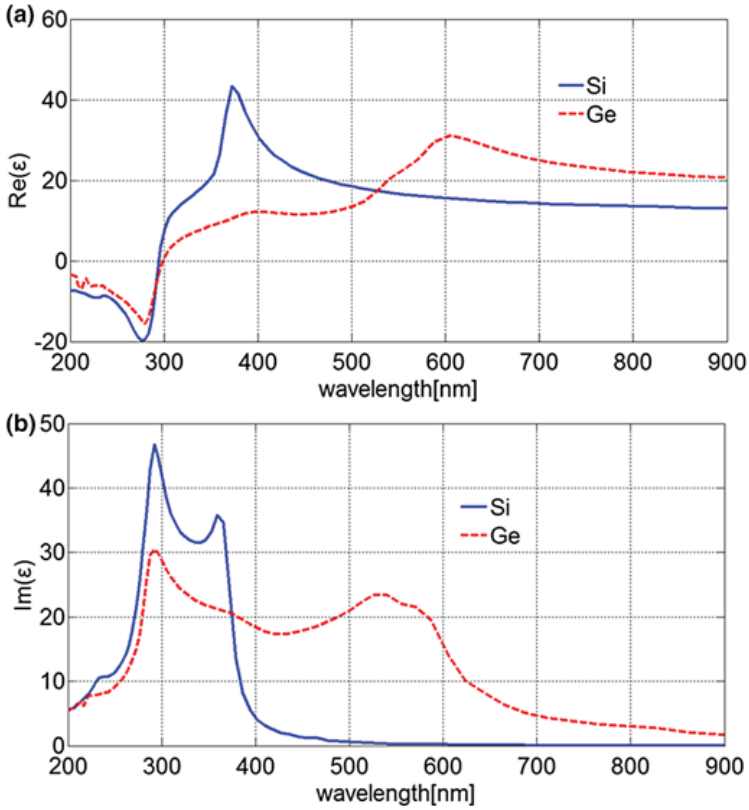

Figure 1.13 Dielectric functions for semiconductor materials of silicon and germanium. The experimental data comes from Ref. [10].

Table 1.4 Bandgap and optical properties of frequently used dielectric materials.

Material	InAs	Ge	Si	InP	GaAs	TiO_2	ZnO	SiO_2	Al_2O_3
Bandgap (eV)	0.35	0.66	1.12	1.34	1.424	3.0	3.44	~8.5	~8.9
λ (μm)	3.50	1.87	1.10	0.92	0.87	0.41	0.36	~0.15	~0.14
Carrier density (cm^{-3})	10^{15}	2×10^{13}	10^{10}	1.3×10^{7}	2.1×10^{6}	—	—	—	—

semiconductor crystals. The photonic bandgap structures exhibit a periodic potential by designing allowed and forbidden energy bands, which influence the motion equations of electrons.

In order to obtain periodic potentials, these bandgap semiconductor crystals are constructed by materials with a sharp contrast of dielectric function in the constituted cell. For instance, a nanostructure combines a dielectric material with a metal. The prominent superiority of such artificial structures has wide applications, for example, to enable light propagation in specified directions, to localize light in particular zones, or to completely prohibit light propagation within a certain bandgap. Table 1.4 lists the bandgap and optical properties of a few frequently used dielectric materials.[10,16]

There are no absolute conductors or dielectrics: the behavior of metals is primarily conductive due to the domination of the intraband transitions at a low frequency. However, metals behave like dielectrics at frequencies greater than the plasma frequency of the ultraviolet wavelength.

1.4.2 Kramers–Kronig relation

The dielectric functions and complex refractive index are the inherent definitions of optical properties of media. Unfortunately, they are not measured directly in experiments. What we can obtain physically are the transmissivity and reflectivity of a medium. The Kramers–Kronig formula relates the complex optical properties with the physical observables, which are useful tools to obtain the dielectric functions.[17]

When the real part ε_1 (or the imaginary part ε_2) of a permittivity function has been obtained over a frequency region (ω_l, ω_h), the complex counterpart will be determined from the Kramers–Kronig formula[18]

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi} P_c \int_{\omega_l}^{\omega_h} \frac{\tilde{\omega} \times \varepsilon_2(\tilde{\omega})}{\tilde{\omega}^2 - \omega^2} d\tilde{\omega}, \tag{1.61a}$$

$$\varepsilon_2(\omega) = -\frac{2\omega}{\pi} P_c \int_{\omega_l}^{\omega_h} \frac{\varepsilon_1(\tilde{\omega})}{\tilde{\omega}^2 - \omega^2} d\tilde{\omega}, \tag{1.61b}$$

where P_c is the Cauchy principal value. For the refractive index, the real part and its imaginary part are related with the Kramers–Kronig relations[18]

$$n_1(\omega) = 1 + \frac{2}{\pi} P_c \int_{\omega_l}^{\omega_h} \frac{\tilde{\omega} \times n_2(\tilde{\omega}) - \omega n_2(\omega)}{\tilde{\omega}^2 - \omega^2} d\tilde{\omega}, \tag{1.62a}$$

$$n_2(\omega) = -\frac{2\omega}{\pi} P_c \int_{\omega_l}^{\omega_h} \frac{n_1(\tilde{\omega}) - n_1(\omega)}{\tilde{\omega}^2 - \omega^2} d\tilde{\omega}. \tag{1.62b}$$

The relation between the reflectivity $R(\omega)$ and its phase difference $\Delta\varphi(\omega)$ are also defined by the Kramers–Kronig form

$$\ln R(\omega) = \frac{2}{\pi} P_c \int_{f_l}^{f_h} \frac{\tilde{\omega} \times \Delta(\tilde{\omega}) - \omega \times \Delta\varphi(\omega)}{\tilde{\omega}^2 - \omega^2} d\tilde{\omega}, \tag{1.63a}$$

$$\Delta\varphi(\omega) = -\frac{2\omega}{\pi} P_c \int_{\omega_l}^{\omega_h} \frac{\ln R(\tilde{\omega}) - \ln R(\omega)}{\tilde{\omega}^2 - \omega^2} d\tilde{\omega}. \tag{1.63b}$$

The Kramers–Kronig theory is widely useful to quantify the optical characteristics of natural and artificial materials, which enable us to measure optical properties.

1.4.3 Obtaining optical functions from physical observables

The complex dielectric functions (permittivity and permeability) are linked to each other with the refractive index, as expressed in Eqs. (1.26) and (1.27), given the permeability is zero. At a semiconductor-metal interface, the reflectivity and the associated phase difference in normal incidences is calculated by[12]

$$R = \frac{(n_0 - n_1)^2 + n_2^2}{(n_0 + n_1)^2 + n_2^2}, \tag{1.64a}$$

$$\Delta\varphi = \arctan \frac{2n_0 n_2}{(n_0^2 - n_1^2 - n_2^2)}, \tag{1.64b}$$

where n_0 refers to the refractive index of the incident medium, and $n_1 + jn_2$ is the complex refractive index of the substrate material. In general, at the boundary of free space and a nonmagnetic substrate material, the reflection coefficient is presented through Snell's law:[3]

$$R \times e^{j\omega\theta} = \frac{n_1 - 1 + jn_2}{n_1 + 1 + jn_2}, \tag{1.65}$$

where θ refers to a series of incident angles that depend on the wave frequency. The optical functions are then related to the physical observables by

$$n_1 = \frac{1 - R^2}{1 + R^2 - 2R \times \cos\theta}, \tag{1.66a}$$

$$n_2 = \frac{2R \times \sin\theta}{1 + R^2 - 2R \times \cos\theta}. \tag{1.66b}$$

In Eqs. (1.64)–(1.66), optical parameters are the function of optical frequency. The optical refractive index can be determined in Eq. (1.66) by the reflectivity over a frequency region.

The reflection coefficient and the corresponding phase difference can be calculated from the frequency-dependent reflectivity. In order to calculate the complex refractive index $n_1 + jn_2$, two independent measurements could be made to relate the complex dielectric function to physical observables. One convenient method measures the reflectivity in different conditions, for example, different incident angles of light with the same wavelength or various wavelengths with the same incident angles. Both the real and imaginary parts of a complex dielectric function can be related with the Kramers–Kronig relations.

1.5 Effective Medium Approach for Composite Nanostructures

1.5.1 Effective medium theory

In plasmonic optics, composite nanostructures are often made from two or multiple materials in either an arbitrary fashion or ordered patterns. Among them, the host materials are often electromagnetically continuous media, while a small number of inclusions or particles are incorporated in host materials. For the composite nanostructures made from metallic and dielectric components, the overall optical properties may be significantly different from those of constituent host materials, and also differ greatly from its inclusions. The electromagnetic properties are determined by the electrical dimensional sizes of the inclusions and the interparticle distance. It is hard to develop a universal method to analyze the optical properties of such arbitrary nanostructure materials.[11]

In general, composite nanostructures are designed based on one of two strategies: (1) random metal–dielectric composites or (2) well-structured building blocks. In both of these cases, the dimensional sizes of the constituted structures are designed intentionally smaller than the wavelength. In such a condition, the composite nanostructure is assumed to be an effective continuum medium. The overall effective optical effects can be described by the previously known dielectric functions of individual components and their volume fractions. There are several analytical approaches to derive the effective electromagnetic response of the composited materials, which are often known as the effective medium approach.[19,20]

In the effective medium theory, the electromagnetic responses of the composites are assumed to be electric dipoles, and the collective responses of the electric dipoles take on an overall dielectric response of the matter. In this case, the diffuse field is neglected because it is much smaller compared to the coherent field. Several models of the effective medium theory have been proposed, e.g., the Maxwell–Garnett theory, the Bruggerman model, and the asymptotic multiscale theory.[21]

1.5.2 Topologies of metal–dielectric composites

The effective medium approach attempts to describe the heterogeneous medium using effective dielectric functions (effective permittivity and permeability) or optical functions (effective refractive index). Without losing generality in the following sections, the media made of metal–dielectric components are assumed to be composed of isotropic inclusions.

Two typical topologies are presented as shown in Fig. 1.14. When the relative concentration of inclusion particles embedded in the host material is small and the inclusions have well-defined shapes, this type of composite topology is called Maxwell–Garnett geometry.[21] When the percentage compositions of two constituent materials are almost identical to each other, as shown in Fig. 1.14(b), the two compositions play comparable roles in its optical parameters. This type of heterogeneous topology is called Bruggeman geometry.[22] Corresponding to these topology geometries are two effective medium approaches that are widely used to obtain the effective parameters, i.e., Maxwell–Garnett theory (MGT) and Bruggeman effective medium theory (EMT).

1.5.3 Lorentz cavity model

As illustrated in Fig. 1.15, a Lorentz cavity model is used to evaluate the local field surrounding the particle. In this model, an individual particle immersed in a uniform medium is excited by an applied field. The external electric field

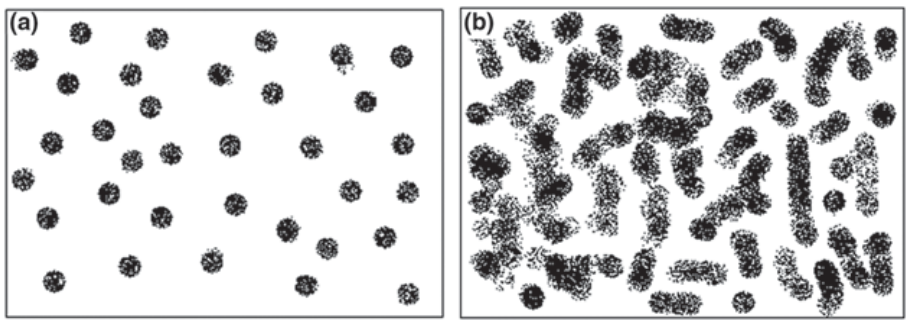

Figure 1.14 Typical topology geometries of metal–dielectric composites: (a) Maxwell–Garnett geometry and (b) Bruggeman geometry. The bright and dark areas represent the host dielectric components and the embedded metal inclusions, respectively.

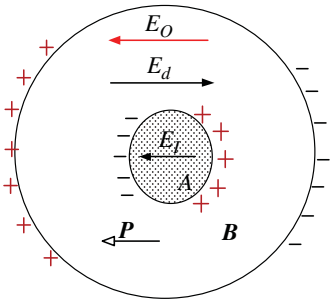

Figure 1.15 Lorentz cavity model to calculate the local field of the central particle.

E induces the redistribution of surface charges of the cavity model, as marked in the two-sided ordinate. The charge redistribution results in an additional depolarized field $\mathbf{E}_d$, which results in the polarized surface charges of the central particle as labeled in bilateral symmetry. The surface charges lead to the polarization field $\mathbf{E}_I$ inside the particle.

The local field $\mathbf{E}_L$ at the site of the particle refers to a microscopic field, which acts on the particle and fluctuates sharply in the Lorentz cavity. Thus, the local field $\mathbf{E}_L$ is decomposed into four components[11]

$$\mathbf{E}_L = \mathbf{E} + \mathbf{E}_d + \mathbf{E}_I + \mathbf{E}_p. \tag{1.67}$$

It is notable that the term $\mathbf{E}_p$ denotes the field induced by other particles dispersed throughout the whole system. When the particles are randomly distributed, such as the conditions of particles dispersing in a liquid or gas environment, the term $\mathbf{E}_p$ would vanish due to the counterbalance of the symmetric lattice. In the familiar cubic crystal lattice, the $\mathbf{E}_p$ term is also zero due to the lattice symmetry.

The depolarized local field $\mathbf{E}_d$ relates the macroscopic polarization **P** by the relation $\mathbf{E}_p = -\mathbf{P}/\varepsilon_0$, where ε_0 is the permittivity of the free space in the aforementioned cavity. It is assumed that the particles are immersed in the host medium of vacuum. Both **E** and $\mathbf{E}_d$ are field values averaged within the entire composite medium. For the electric fields inside the particle, the total charge is integrated over its normal surface. Because of the structural lattice symmetry, the total field $\mathbf{E}_I$, ascribed to all of surface charges accumulated on the particles, is parallel with the external field. The total polarization field is estimated with a magnitude of[11,24]

$$\mathbf{E}_I = \mathbf{P}/3\varepsilon_0. \tag{1.68}$$

Therefore, the local field upon the particle is

$$\mathbf{E}_L = \mathbf{E} + \mathbf{P}/3\varepsilon_0 - \mathbf{P}/\varepsilon_0. \tag{1.69}$$

Following the treatment derived from the Drude model, the polarization **P** is related to the electric dipole moment. The macroscopic polarization is expressed by

$$\mathbf{P} = f\alpha(\mathbf{E} + \mathbf{P}/3\varepsilon_0), \tag{1.70}$$

where f is the volume density of the particles, and α is the polarizability of the individual particle. A combination of the constitutive equation $\mathbf{P} = \varepsilon_0(\varepsilon_r - 1)\mathbf{E}$ and Eq. (1.70) leads to the particle polarizability formulated with the relative dielectric permittivity

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = f\frac{\alpha}{3\varepsilon_0}, \tag{1.71a}$$

$$\alpha = \frac{3\varepsilon_0}{f}\frac{\varepsilon_r - 1}{\varepsilon_r + 2}, \tag{1.71b}$$

where $\varepsilon = \varepsilon_0\varepsilon_r$ is the permittivity of the individual particle. This notion is called the Clausius–Mossotti theory, which calculates the microscopic polarizability of an individual particle from the observable dielectric function of bulk materials. It provides a distinct connection between the electric response of a microscopic particle and the macroscopic behavior of a bulk material.

If the particle polarizability volume $\alpha' = 4\pi\varepsilon_0\alpha(\mathrm{m}^3)$ is used in Eq. (1.71), the Clausius–Mossotti formula is rewritten in the form of

$$\frac{\varepsilon - \varepsilon_0}{\varepsilon + 2\varepsilon_0} = \frac{4\pi}{3}V\alpha'. \tag{1.72}$$

The following section adopts the particle polarizability α, as in Eqs. (1.71).

1.5.4 Maxwell–Garnett theory

In a metal–dielectric composite medium with the Maxwell–Garnett geometry, in which the volume fraction follows that shown in Fig. 1.16(a), the metal particles and the dielectric constituent are viewed as the inclusions and the host medium, respectively. When the metallic particles are spherical particles and the host medium is an isotropic dielectric substance, the Clausius–Mossotti relation of Eq. (1.71) becomes

$$\frac{f\alpha}{3\varepsilon_0\varepsilon_h} = \frac{\varepsilon - \varepsilon_h}{\varepsilon + 2\varepsilon_h}. \tag{1.73a}$$

When considering a composite inclusion with permittivity ε_i, the parameter ε in Eq. (1.71a) represents the effective permittivity of the composite inclusion. In these conditions, the polarizability in Eq. (1.69b) becomes

$$\alpha = 3\varepsilon_0\varepsilon_h f_i\frac{\varepsilon_i - \varepsilon_h}{\varepsilon_i + 2\varepsilon_h}, \tag{1.73b}$$

where f_i is the volume-filling fraction of particles in the composite, and ε_h and ε_i are the relative permittivity of the host medium and the composite particle,

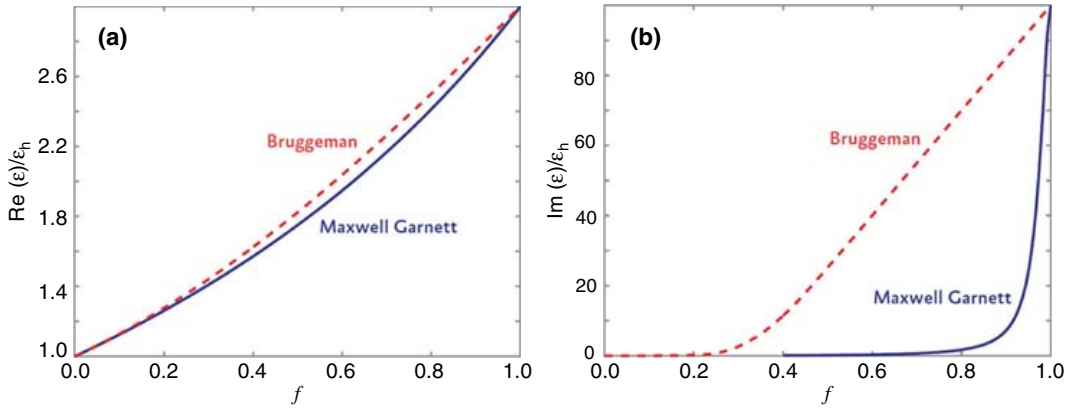

Figure 1.16 Effective dielectric functions as a function of the volume fraction for composite materials as predicted by the Maxwell–Garnett and Bruggeman formula: (a) dielectric function of a composite for permittivity ratio $\varepsilon_i/\varepsilon_h = 3$, and (b) imaginary part of the permittivity for $\varepsilon_i/\varepsilon_h = 5 - j100$. Both feature predictions using the Maxwell–Garnett (solid) and the Bruggeman (dashed) formula.

respectively. Equation (1.73b) is substituted into Eq. (1.73a), and the Clausius–Mossotti relation is rewritten as

$$\frac{\varepsilon - \varepsilon_h}{\varepsilon + 2\varepsilon_h} = f_i \frac{\varepsilon_i - \varepsilon_h}{\varepsilon_i + 2\varepsilon_h}. \tag{1.74}$$

The equivalent dielectric function of a metal–dielectric composite medium is obtained as

$$\varepsilon = \varepsilon_h \frac{1 + 2\Gamma_i}{1 - \Gamma_i}, \quad \Gamma_i = f_i \frac{\varepsilon_i - \varepsilon_h}{\varepsilon_i + 2\varepsilon_h}. \tag{1.75}$$

Equation (1.75) formulates the equivalent dielectric function of a composite medium by the individual dielectric functions of the inclusions and the host material, i.e., the Maxwell–Garnett theory. In the MGT, the embedded particles and the host solid material are treated in an unsymmetrical manner.

When the difference of the dielectric functions between the two composite materials are obvious, this asymmetry becomes particularly sharp. For a dilute composite material with the volume-filling factor less than a unit away $f_i << 1$, the equivalent dielectric function of the composite materials appears in the function expansion of the Taylor series as

$$\varepsilon = \varepsilon_h + 3\Gamma_i \varepsilon_h + O(3f_i^2). \tag{1.76}$$

It indicates that the equivalent dielectric function of a dilute composite material is a linear function of the volume density and the dielectric function of their individual inclusions. Equation (1.75) is still valid in the limiting conditions: when the constituted phase of inclusions vanishes $f_i \to 0$, the estimated dielectric function approaches that of the host material by $\varepsilon \to \varepsilon_h$.

On the other hand, when the host medium phase vanishes $f_i \rightarrow 1$, it is predicted to be the dielectric function of inclusions by $\varepsilon \rightarrow \varepsilon_i$.

For a single metallic, spherical particle embedded in a host medium, the equivalent dielectric function in Eq. (1.75) satisfies the resonance requisition $\varepsilon_i = -2\varepsilon_h$ for a surface plasmon resonance, which means that the dielectric function of the inclusion materials equals the negative double value of the surrounding medium.

1.5.5 Bruggeman medium theory

The Maxwell–Garnett theory only predicts the approximation of equivalent dielectric function in the dilute two-phase composite. In fact, with the increasing of the volume fraction of inclusions, the deviation between the effective properties predicted by the Maxwell–Garnett formula and by Lord Rayleigh become obvious.[24] Bruggeman proposed a widely known mean-field theory to evaluate the effective dielectric function of composite media. In the Bruggeman theory, the inclusions and the surrounding medium are weighted symmetrically using the volume fraction of the individual components. For a composite medium made of two inclusions, the formula is expressed as[22]

$$\frac{\varepsilon - \varepsilon_h}{\varepsilon + 2\varepsilon_h} = f_1 \frac{\varepsilon_1 - \varepsilon_h}{\varepsilon_1 + 2\varepsilon_h} + f_2 \frac{\varepsilon_2 - \varepsilon_h}{\varepsilon_2 + 2\varepsilon_h}, \tag{1.77}$$

where ε_j ($j = 1, 2$) and ε_h are the dielectric functions of the dispersed spherical particles and the environment medium, and f_j ($j = 1, 2$) is the volume-filling fraction of i-order phase inclusions. Because the two-phase inclusions appear in a symmetric manner, the total sum of the volume-filling fractions equals one unit. It can obviously be generalized to systems that contain more than two phases by adding more terms:

$$f_i = n_i \Big/ \sum_m n_m, \quad \sum_i f_i = 1, \tag{1.78}$$

where m represents the total phase of inclusions. Furthermore, when $\varepsilon = \varepsilon_h$, there is always a valid expression

$$f_1 \frac{\varepsilon_1 - \varepsilon}{\varepsilon_1 + 2\varepsilon} + f_2 \frac{\varepsilon_2 - \varepsilon}{\varepsilon_2 + 2\varepsilon} = 0. \tag{1.79}$$

The equivalent dielectric function to this quadratic equation is

$$\varepsilon = \Gamma_2 \pm \sqrt{[\Gamma_2]^2 + \frac{\varepsilon_1 \varepsilon_2}{2}}, \tag{1.80a}$$

$$\Gamma_2 = \frac{3(\varepsilon_1 f_1 + \varepsilon_2 f_2) - (\varepsilon_1 + \varepsilon_2)}{4}. \tag{1.80b}$$

The +/− sign in Eq. (1.80a) is determined to ensure that the imaginary part of the obtained dielectric function is positive.[11]

Based on Eq. (1.79) for an effective medium expression, it is reasonable to get the general theory for a composite material with multiple components in a symmetrical manner

$$\sum_i f_i \frac{\varepsilon_i - \varepsilon}{\varepsilon_i + 2\varepsilon} = 0, \quad \sum_i f_i = 1. \tag{1.81}$$

Equation (1.81) is the popular Bruggeman form of the EMT. This formula can be expanded to any number of components in a variety of composite materials because each component is treated equally in the mixture. For two-phase composite materials, the Taylor series expansion of the Bruggeman approach in the case of low-volume fractions ($f_i << 1$) becomes

$$\varepsilon = \varepsilon_h + 3\Gamma_i \varepsilon_h + O(9f_i^2). \tag{1.82}$$

The second order of volume fraction $O(9V_i^2)$ differs from the expansion in the MGT in Eq. (1.76).

Figure 1.16 illustrates the effective dielectric function of a composite material made of two components predicted by the Maxwell Garnett and the Bruggeman formula. For a low-volume fraction of $f_i < 0.33$, both the Maxwell–Garnett and the Bruggeman formula present nearly the same curves. However, when the volume fraction increases, the predictions become strongly different. Especially for the imaginary part, the deviations increase sharply. For the whole region, the Maxwell–Garnett prediction remains lower than that of the Bruggeman estimate. The deviation of the imaginary part of the relative dielectric functions suddenly becomes obvious after the volume fraction increases over the threshold value $f_i \approx 0.33$.

When the shape effect of the inclusions are taken into account, the term $\varepsilon_i + 2\varepsilon$ is replaced by $\varepsilon_i + (d-1)\varepsilon$ in the Maxwell–Garnett and the Bruggeman formula for a more-general d-dimensional inclusion. For a two-phase d-dimensional particle, the Maxwell–Garnett formula becomes the generalized formula

$$\frac{\varepsilon - \varepsilon_h}{\varepsilon + (d-1)\varepsilon_h} = f_i \frac{\varepsilon_i - \varepsilon_h}{\varepsilon_i + (d-1)\varepsilon_h}, \tag{1.83}$$

and the Bruggeman expression becomes

$$f_1 \frac{\varepsilon_1 - \varepsilon}{\varepsilon_1 + (d-1)\varepsilon} + f_2 \frac{\varepsilon_2 - \varepsilon}{\varepsilon_2 + (d-1)\varepsilon} = 0, \tag{1.84}$$

where d refers to the dimensions of the inclusions determined by the shape and the orientation of the nanoparticles with respect to the external electric field. The value 2 is usually taken corresponding to spherical inclusions. For 2D

metal–dielectric composite films, the resonance peak in the MGT will be $\omega_p/\sqrt{1+(n-1)\varepsilon_h}$, and the inflexion of the effective dielectric function in the Bruggeman prediction will be 0.5.

The aforementioned volume–weight approaches for the mean value of permittivity are just analytical approximations for the effective dielectric function of composite nanostructures. The term "effective" indicates that they are equivalent in physical concept. It is a simple implemention for low-dimensional composite materials. A more-general approach for effective medium analysis is the Bergman theory or the Bergman–Milton theory, which is often called the spectral representation method. Further details about this approach can be found in textbooks[23] and the original papers.[25]

References

1. J. Aukong, *Electromagnetic Wave Theory*, John Wiley & Sons, New York (1986).
2. P. Šolín, *Partial Differential Equations and the Finite Element Method*, John Wiley & Sons, New York (2006).
3. N. Johnson et al, *Electromagnetic Characterization of Complex Nano-structured Materials: An Overview*, ECONAM CSA project, http://econam.metamorphose-vi.org/index.php?option=com_docman&task=doc_download&gid=263&Itemid=24 (2011).
4. J. M. Lourtioz, *Photonic Crystals: Towards Nanoscale Photonic Devices*, p. 122, Springer, Berlin (2005).
5. J. E. Conrad, *Electromagnetic Waves and Radiating Systems*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ (1968).
6. A. vander Vorst, A. Rosen, and Y. Kotsuka, *RF/Microwave Interaction with Biological Tissues*, John Wiley & Sons, New York (2006)
7. S. G. Rodrigo, F. J. García-Vidal, and L. Martín-Moreno, "Influence of material properties on extraordinary optical transmission through holes arrays," *Physical Review B* **77**(7), 075401 (2008).
8. M. S. Dresselhaus, "*Solid State Physics Part II: Optical Properties of Solids*," lecture notes, Massachusetts Institute of Technology, Cambridge, MA (1999).
9. R. J. Van Overstraeten and R. P. Mertens, "Heavy doping effects in silicon," *Solid State Electron*. **30**(11), 1077–1087 (1987).
10. E. D. Palik, *Handbook of Optical Constants of Solids*, Academic Press Limited, New York (1998).
11. W. Cai and V. M. Shalaev, *Optical Metamaterials: Fundamentals and Applications*, pp. 26–37, Springer, New York (2010).
12. D. Rakić, A. B. Djurišić, J. M. Elazar, and M. L. Majewski, "Optical properties of metallic films for vertical-cavity optoelectronic devices," *Appl. Optics* **37**(22), 5271–5283 (1998).

13. J. B. Pendry, A. J. Holden, and D. J. Robbins, "Magnetism from Conductors and Enhanced Nonlinear Phenomena," *IEEE Trans. Microwave Theory Techniques* **47**(11), 2075–2084 (1999).
14. M. Dalarsson, Z. Jakšić, and P. Tassin, "Structures Containing Left-Handed Metamaterials with Refractive Index Gradient: Exact Analytical Versus Numerical Treatment," Microwave Review, (2009).
15. S. C. Jain and D. J. Roulston, "A simple expression for band gap narrowing (BGN) in heavily doped Si, Ge, GaAs and Ge_xSi_{1-x} strained layers," *Solid-State Electronics* **34**(5), 453–465 (1991).
16. C. Kittel, *Introduction to Solid State Physics*, Wiley, New York (2004).
17. V. Lucarini, "Kramers-Kronig relations and sum rules in optical harmonic generation processes: theory and applications," Ph.D. dissertation, University of Joensuu, Finland (2003).
18. V. Lucarini, J. J. Saarinen, K.-E. Peiponen, and E. M. Vartiainen, *Kramer-Kronig Relations in Optical Materials Research*, Springer, Berlin (2005).
19. Wikipedia, "Effective medium approximations: for the Maxwell Garnett equation and Bruggeman's model," http://en.wikipedia.org/wiki/Effective_medium_approximations.
20. A. A. Maradudin, *Structure Surfaces as Optical Metamaterials*, Cambridge University Press, Cambridge, UK (2011).
21. C. C. Tuck, *Effective Medium Theory*, Oxford University Press, Oxford, UK (1999).
22. D. E. Aspnes, "Local-field effects and effective medium theory: A microscopic perspective,." *Am. J. Phys.* **50**(8) (1982).
23. J. C. M. Garnett, "Colours in metal glasses and in metallic films," *Phil. Trans. R. Soc. Lond.* **203**, 237–288 (1905).
24. G. W. Milton, *The Theory of Composites*, Cambridge University Press, Cambridge, UK (2002).
25. D. J. Bergman, "The dielectric constant of a composite material—A problem in classical physics," *Phys. Rep.* **43**(9), 377–407 (1978).