

Chapter 2

Diffraction and Light Scattering

Light interacts with matter. This is the foundation for all the rich phenomena and useful applications associated with light and biological media. The interaction takes many forms. Light can be absorbed, or transmitted, or reflected, or scattered. All these processes can participate in interference phenomena in biology and medicine.

The interaction of light with matter (Fig. 2.1) is characterized as a scattering process that converts an incident (initial) optical mode into an outgoing (final) optical mode

$$(\vec{k}_i, \omega_i, \hat{p}_i) \rightarrow (\vec{k}_f, \omega_f, \hat{p}_f) \quad (2.1)$$

A single mode is an infinite plane wave with a frequency ω_i , a wave-vector \vec{k}_i and a polarization \hat{p}_i . The amplitude of the outgoing mode is related to the amplitude of the incident mode through the scattering function

$$E_f(\omega_f) = S_j^i(\theta, \phi, \omega_f - \omega_i) E_i(\omega_i) \quad (2.2)$$

where the scattering function $S_j^i(\theta, \phi; \omega_f - \omega_i)$ is a matrix connecting the input polarizations with the output polarizations of the light field, and θ and ϕ define the scattering direction relative to the incident direction. When the light scattering is elastic, the incident and final frequencies are equal. Inelastic light scattering can also occur, as when scatterers are in motion or when the scattered light is shifted to new frequencies, for instance through Doppler or Raman effects.

Elastic scattering is a general term that can be divided roughly into two separate contributions: geometric ray optics and diffraction. Geometric ray optics involves the reflection or refraction of light. Light rays are reflected or refracted by material density differences when the spatial scale of the inhomogeneity is larger than the wavelength of light. The reflection by a surface, or the bending of a light ray by a change in the index of refraction, is geometric ray optics. The other contribution to scattering is from diffraction. Diffraction is an essential wave phenomenon in which each part of an incident wavefront becomes the source for secondary

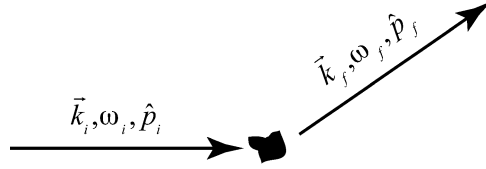


Fig. 2.1 The scattering process converts an incoming optical mode (defined by k -vector, frequency ω and polarization p) to an outgoing mode

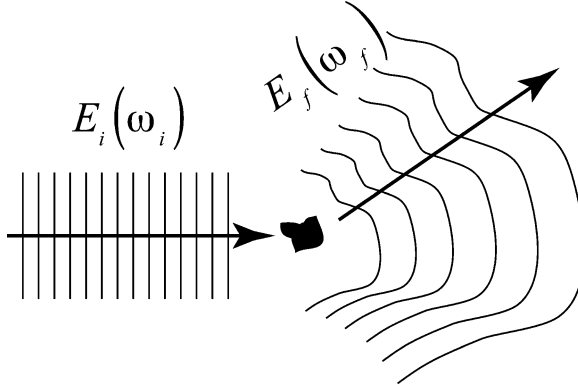


Fig. 2.2 Wave diffraction converts an incident wavefront into a diffracted wavefront

wavelets through Huygen's principle (Fig. 2.2). The secondary wavelets all interfere with each other to produce the observed diffracted light intensities. In this sense, diffraction is the result of wave interference.

2.1 Diffraction

All diffraction is a form of wavefront-splitting interferometry. The interference of partial waves that travel different paths from different parts of a wavefront leads to the complex and beautiful phenomena that fall under the topic of diffraction. Indeed, this summing up of the interference of parts of a wave is the basis for the diffraction integral, which is the main engine of diffraction applications.

2.1.1 Scalar Diffraction Theory

Although electromagnetic waves are vector waves that must satisfy Maxwell's equations at boundaries, a significant simplification is achieved in diffraction theory by treating the field amplitudes as scalar fields. This simplification is often very

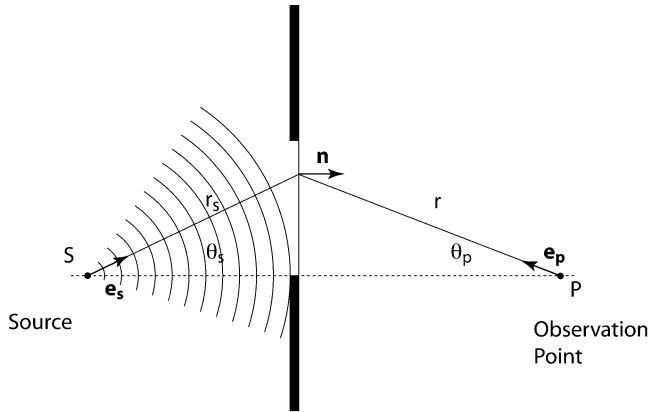


Fig. 2.3 Geometry for the Fresnel–Kirchhoff diffraction approximation

accurate, especially when the diffracting objects are weakly diffracting and produce small perturbations on an incident wave. This is the case if the diffracting objects have small variations in optical path length, and if these variations are on length scales large relative to a wavelength of light. Examples for which scalar diffraction theory fail include Mie scattering by spheres (discussed later in this chapter) because a sphere has a large variation of optical path length, and photonic crystals with photonic bandgaps, because refractive indices vary rapidly on the scale of a wavelength.

The starting point for scalar diffraction is the Helmholtz–Kirchhoff integral for a scalar field

$$\Psi(\mathbf{x}) = \frac{1}{4\pi} \int_{\text{ap}} \left[\Psi \nabla \left(\frac{e^{i kr}}{r} \right) - \frac{e^{i kr}}{r} \nabla \Psi \right] \cdot \mathbf{n} da \quad (2.3)$$

The simple configuration in Fig. 2.3 shows a source point S creating a wave that is incident on an aperture that diffracts the wave to an observation point P . The incident field is a spherical wave

$$\Psi_{\text{inc}} = A \frac{e^{i kr_s}}{r_s} \quad (2.4)$$

and the Helmholtz–Kirchhoff integral becomes

$$\Psi(\mathbf{x}_p) = \frac{1}{4\pi} \int_{\text{ap}} \left[\Psi_{\text{inc}} \nabla \left(\frac{e^{i kr}}{r} \right) - \frac{e^{i kr}}{r} \nabla \Psi_{\text{inc}} \right] \cdot \mathbf{n} da \quad (2.5)$$

After applying the gradient operator, this is

$$\Psi(\mathbf{x}_p) = \frac{1}{4\pi} \int_{\text{ap}} \left[\Psi_{\text{inc}} \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr} \mathbf{e}_p - \frac{e^{ikr}}{r} \left(\frac{ik}{r_s} - \frac{1}{r_s^2} \right) r_s \Psi_{\text{inc}} \mathbf{e}_s \right] \cdot \mathbf{n} da \quad (2.6)$$

In the limit of large distances, only the terms that are linearly inverse to distance contribute and the equation becomes

Kirchhoff diffraction integral:

$$\Psi(\mathbf{x}_p) = \frac{iAk}{4\pi} \int_{\text{ap}} \frac{1}{rr_s} e^{ik(r+r_s)} (\mathbf{e}_p - \mathbf{e}_s) \cdot \mathbf{n} da \quad (2.7)$$

Often, the incident wave is a plane wave of constant amplitude across the aperture, and this integral reduces to

$$\Psi(\mathbf{x}_p) = -iAk \int_{\text{ap}} \frac{1}{r} e^{ikr} \left[\frac{1}{2} (\cos \theta_s + \cos \theta_p) \right] da \quad (2.8)$$

where the term in square brackets is the Fresnel obliquity factor, which is approximately unity for small-angle forward diffraction.

One of the most important features of the Kirchhoff diffraction integral for interferometry is the imaginary number $-i$ in the prefactor. This has the important consequence that the diffracted wave has a $\pi/2$ phase advance relative to the incident wave. This phase shift plays important roles in many aspects of light scattering and diffraction. The mathematical source of this phase shift is the gradient in (2.5) on the dynamic phase of the wave. All extended sources of radiation experience this phase shift upon diffraction. However, point sources do not have this phase shift. This phase difference between extended and point sources is the key element in the origin of refractive index.

A final simplifying step in the development of scalar diffraction integrals takes the observation point far from a planar aperture, known as the Fraunhofer approximation. The geometry in this approximation is shown in Fig. 2.4. This leads to the Fraunhofer diffraction integral

Fraunhofer diffraction integral:

$$\Psi(\mathbf{x}_p) = -i \frac{A e^{ikR}}{\lambda R} \int_{\text{ap}} e^{-ik(x \sin \theta + y \sin \phi)} dx dy \quad (2.9)$$

where R is the (large) distance from the aperture to the observation point, and $x/R = \sin \theta$, $y/R = \sin \phi$.

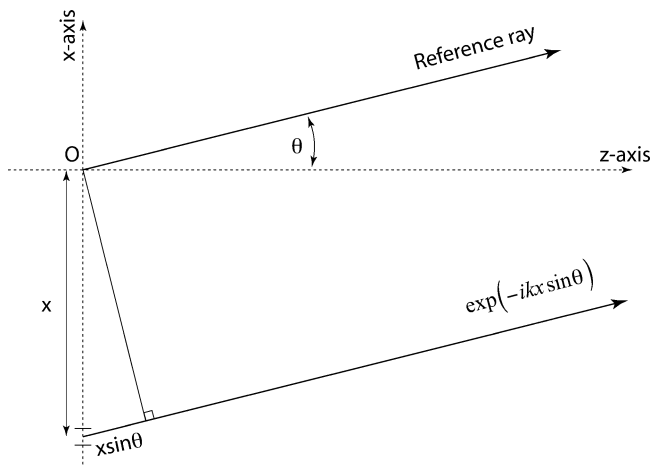


Fig. 2.4 Fraunhofer diffraction geometry. All rays are parallel and inclined at an angle θ . The reference ray from the origin defines zero phase. The path length difference is $x \sin \theta$, with x increasing vertically and θ positive in the counter-clockwise direction

There are an endless number of configurations of sources and apertures that appear in diffraction problems. Several of the most common will be presented as examples that can be extended to more complicated cases that may be encountered in biological applications.

2.1.2 Fraunhofer Diffraction from Apertures and Gratings

Some types of molecular biosensors are based on Fraunhofer diffraction from apertures and gratings. This section presents several basic examples of diffraction that are used in later chapters in this book. Some of the examples include the diffraction effects of thin bilayers, and the molecular responsivities of these diffraction structures are defined.

Example: Diffraction from a Single Square Slit

The single square slit is the simplest example, and is always the starting point to build up to more complicated diffraction patterns. The slit has a width equal to a , shown in Fig. 2.5. The Fraunhofer integral is

$$E_P = -i \frac{E_0 e^{ikR}}{\lambda R} \int_{-a/2}^{a/2} e^{-ikx \sin \theta} dx \quad (2.10)$$

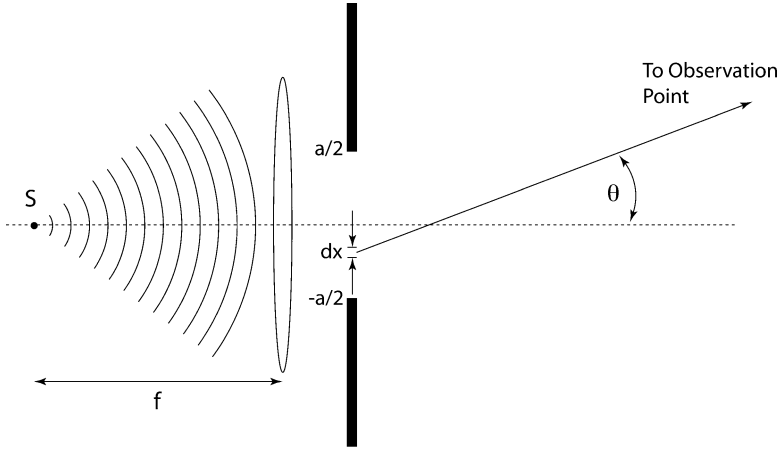


Fig. 2.5 Fraunhofer diffraction geometry for a single slit. The lens converts a point source to a plane wave

where E_0 is the field amplitude in the aperture. The Fraunhofer integral is evaluated to be

$$\begin{aligned}
 E_P &= -i \frac{E_0 e^{ikR}}{\lambda R} \int_{-a/2}^{a/2} e^{-ikx \sin \theta} dx \\
 &= i \frac{E_0 e^{ikR}}{\lambda R} \frac{2}{2ik \sin \theta} \left[e^{-(ika \sin \theta)/2} - e^{i(ka \sin \theta)/2} \right] \\
 &= -2i \frac{E_0 e^{ikR}}{\lambda R} \frac{1}{k \sin \theta} \sin \left(\frac{ka}{2} \sin \theta \right) = -ia \frac{E_0 e^{ikR}}{\lambda R} \text{sinc} \left(\frac{ka}{2} \sin \theta \right) \quad (2.11)
 \end{aligned}$$

giving the field

$$E_P = -iE_0 e^{ikR} \left(\frac{a}{\lambda R} \right) \text{sinc}(ka \sin \theta/2) \quad (2.12)$$

and intensity

$$I_P = P_0 \left(\frac{a}{\lambda R} \right)^2 \text{sinc}^2(ka \sin \theta/2) \quad (2.13)$$

Example: Diffraction from a Slit with a Partial Biolayer

Diffraction is one way to detect thin molecular films, such as biomolecules captured by high-affinity capture molecules (like antibodies on surfaces). As an example, consider a rectangular aperture that is half covered by a thin film of refractive index n and thickness d . We want to consider how the presence of the film changes the diffracted intensity in the far-field. The field (2.12) is the starting point for this solution. The answer can be written down by inspection as

$$E_P = -iE_0 e^{ikR} \frac{a}{2\lambda R} \text{sinc}\left(\frac{ka}{4} \sin \theta\right) \left[e^{i(ka \sin \theta)/4} + e^{ik(n-1)d} e^{-i(ka \sin \theta)/4} \right] \quad (2.14)$$

which is the sum of fields from two apertures of width $a/2$, one of which has the extra phase $\delta = k(n-1)d$. The diffracted intensity is

$$I_P = 2I_0 \left(\frac{a}{2\lambda R}\right)^2 \text{sinc}^2\left(\frac{ka}{4} \sin \theta\right) \left[1 + \cos\left(\frac{ka}{2} \sin \theta - \delta\right) \right] \quad (2.15)$$

which still has the basic sinc^2 envelope, with an extra term (the second term in the square brackets) that causes a small shift in the far-field diffraction. This small shift provides a measure of the thickness of the film and is one way to use optical interferometry (in this case diffraction) to detect bilayers in an optical biosensor.

The performance of diffraction-based optical biosensors is characterized by the responsivity that is defined as the intensity change per optical path length difference of the bilayer $h = (n-1)d$. The responsivity of this diffraction-based biosensor is

$$\mathbb{R}_\delta = \frac{dI_P}{d\delta} = -\frac{I_0}{2} \left(\frac{a}{\lambda R}\right)^2 \text{sinc}^2(ka \sin \theta/4) \sin\left(\frac{ka}{2} \sin \theta\right) \quad (2.16)$$

which continues to have the sinc^2 envelope, but now with the extra sine term at the end. The angular responsivity on the detection x -axis is shown in Fig. 2.6. The shift in the far-field diffraction caused by the bilayer leads to an asymmetric intensity change. If only the total intensity is measured, then the bilayer effect would not be detected. However, by placing a split detector on the detection plane, and differencing the left and right detected intensities, then the difference signal is linearly proportional to the thickness of the bilayer. This diffraction-based biosensor configuration is closely related to phase-contrast detection on bio-optical compact disks (BioCDs) [1].

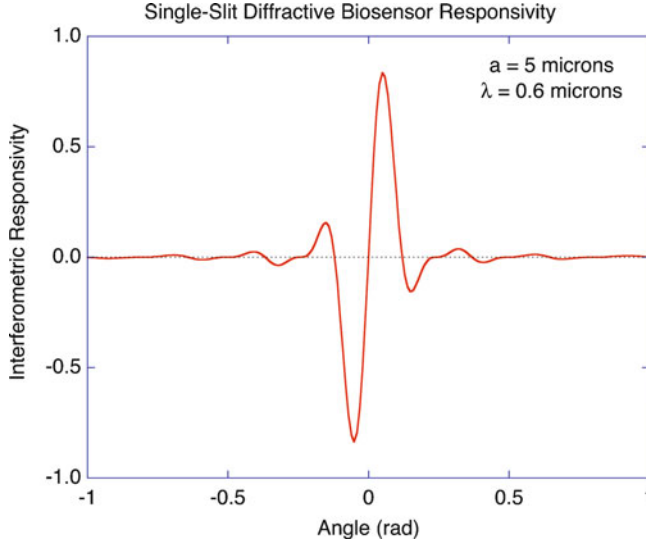


Fig. 2.6 Rectangular-aperture diffractive biosensor responsivity. A rectangular aperture is half-covered by a molecular film of thickness d and refractive index n . The interferometric responsivity is the change in the diffracted intensity per phase shift caused by the film. The film causes an asymmetric far-field diffraction pattern with a responsivity approaching unity

Example: Diffraction by a Circular Aperture

One of the most common apertures encountered in experimental optics is the circular aperture. The Fraunhofer diffraction integral is expressed in polar coordinates (r, Φ) on the detection plane

$$E(r, \Phi) = -i\sqrt{S_0} \frac{e^{ikR}}{\lambda R} \int_0^a \int_0^{2\pi} e^{-i(kr'r/R) \cos(\varphi - \Phi)} d\varphi r' dr' \quad (2.17)$$

integrated over (r', φ) on the aperture plane, where

$$R = \sqrt{L^2 + r^2} \quad (2.18)$$

and L is the distance to the screen (or is the focal length of a Fourier-transform lens). This integral is re-expressed using Bessel functions

$$J_m(u) = \frac{i^{-m}}{2\pi} \int_0^{2\pi} e^{i(mv + u \cos v)} dv \quad (2.19)$$

The diffraction integral becomes

$$E(r, \Phi) = -i\sqrt{S_0} \frac{e^{ikR}}{\lambda L} 2\pi \int_0^a J_0(kr'r/R) r' dr' \quad (2.20)$$

that is evaluated using

$$\int_0^u u' J_0(u') du' = u J_1(u) \quad (2.21)$$

to yield

$$\begin{aligned} E(r) &= -i\sqrt{S_0} \frac{e^{ikR}}{\lambda L} 2\pi a^2 \left(\frac{R}{kar} \right) J_1 \left(\frac{kar}{R} \right) = -i\sqrt{S_0} \frac{e^{ikR}}{\lambda L} \pi a^2 \left[\frac{2J_1 \left(\frac{kar}{R} \right)}{\left(\frac{kar}{R} \right)} \right] \\ &\approx -i\sqrt{S_0} \frac{e^{ikR}}{\lambda L} \pi a^2 \left[\frac{2J_1(ka \sin \theta)}{ka \sin \theta} \right] \end{aligned} \quad (2.22)$$

where $r/R = \sin \theta$, and the value of $J_1(x)/x = 1/2$ as x goes to zero. The angular intensity is approximately (for small angles θ)

$$I(\theta) = I_0 \left(\frac{2J_1(ka \sin \theta)}{ka \sin \theta} \right)^2 \quad (2.23)$$

that has an oscillatory behavior qualitatively similar to the sinc squared function of a rectangular aperture.

Example: Diffraction by Multiple Square Slits

Multiple slits form a diffraction grating. The diffraction from each slit is given by (2.12). This is modulated by a periodic part determined by the periodic spacing Λ of the multiple slits. For the periodic part, the total field is

$$E = E_p e^{-i\omega t} e^{ikr_1} \left[1 + e^{i\phi} + (e^{i\phi})^2 + (e^{i\phi})^3 + \dots + (e^{i\phi})^{N-1} \right] \quad (2.24)$$

where $\phi = k\Lambda \sin \theta$, and where $E_p = -iE_0 e^{iKR} (a/\lambda R) \text{sinc}(Ka \sin \theta/2)$ is the field from a single slit located at r_1 . The term in brackets is a geometric series with the value

$$\frac{(e^{iN\phi} - 1)}{(e^{i\phi} - 1)} = e^{i(N-1)\phi/2} \left(\frac{\sin N\phi/2}{\sin \phi/2} \right) \quad (2.25)$$

The total field is then

$$E = E_p e^{-i\omega t} e^{i[kr_1 + (N-1)\phi/2]} \left(\frac{\sin N\phi/2}{\sin \phi/2} \right) \quad (2.26)$$

If the array is referenced to its center, then the periodic factor is

$$E = E_P e^{ikR - i\omega t} \left(\frac{\sin N\phi/2}{\sin \phi/2} \right) \quad (2.27)$$

with a total field

$$E = E_P e^{ikR - i\omega t} \left(\frac{\sin N\phi/2}{\sin \phi/2} \right) \text{sinc}(ka \sin \theta/2) \quad (2.28)$$

and an intensity

$$I = I_0 \left(\frac{\sin^2 N\phi/2}{\sin^2 \phi/2} \right) \text{sinc}^2(ka \sin \theta/2) \quad (2.29)$$

Note that this is the product of the rectangular grating diffraction pattern and the diffraction pattern of N point sources placed in a regular array with a spacing Λ . Because diffraction is equivalent to a Fourier transform, the convolution of two aperture functions on the object plane becomes the product of the diffraction patterns in the far-field.

Example: Diffraction of a Gaussian Beam by Multiple Slits

A Gaussian beam with radius w_0 illuminating a multiple slit grating diffracts as a Gaussian beam. The field is

$$E_G(\theta) = -iE_0 \text{sinc}\left(\frac{ka \sin \theta}{2}\right) \sum_{m=-M}^M \exp\left(-2\left(\frac{\pi w_0}{\Lambda}\right)^2 \sin^2(\theta - \theta_m)\right) \quad (2.30)$$

for a periodicity Λ and a slit width $a < \Lambda$. The conditions on the diffraction orders are

$$\begin{aligned} \sin \theta_m &= m\lambda/\Lambda \\ M &= \text{trunc}(\Lambda/\lambda) \end{aligned} \quad (2.31)$$

The periodic part can be given its own function definition as

$$P_G(\theta, w_0/\Lambda) = \sum_{m=-M}^M \exp\left(-2\left(\frac{\pi w_0}{\Lambda}\right)^2 \sin^2(\theta - \theta_m)\right) \quad (2.32)$$

which is encountered any time a Gaussian beam is diffracted from a periodic grating. It replaces the periodic function in (2.29)

$$P_G(\theta, w_0/\Lambda) \approx \left(\frac{\sin^2 Nk\Lambda \sin \theta/2}{\sin^2 k\Lambda \sin \theta/2} \right) \quad (2.33)$$

where the effective number of slits is

$$N = \pi w_0/\Lambda \quad (2.34)$$

When there is no overlap between the diffraction orders, the intensity is

$$\begin{aligned} I_G(\theta) &= I_0 \operatorname{sinc}^2 \left(\frac{ka \sin \theta}{2} \right) \sum_{m=-M}^M \exp \left(- \left(\frac{2\pi w_0}{\Lambda} \right)^2 \sin^2(\theta - \theta_m) \right) \\ &= I_0 \operatorname{sinc}^2 \left(\frac{ka \sin \theta}{2} \right) [P_G(\theta, w_0/\Lambda)]^2 \end{aligned} \quad (2.35)$$

The term in the summation is a series of Gaussian beams with angular widths given by

$$\Delta\theta = \frac{\Lambda}{2\pi w_0} \quad (2.36)$$

Example: Diffraction by a Periodic Biolayer Grating

One class of diffraction-based biosensors uses a periodic grating of capture molecules on a surface. When exposed to a sample containing the target biomolecules, these bind in a stripe pattern, shown in Fig. 2.7. The thin nature of the biomolecular stripes imparts a periodic phase modulation on a reflected optical wave. If we assume a sinusoidal phase modulation, the transmitted field is

$$E_r = E_0 \exp[ik_0(n-1)d(1 - \cos(Kx + \phi))] \quad (2.37)$$

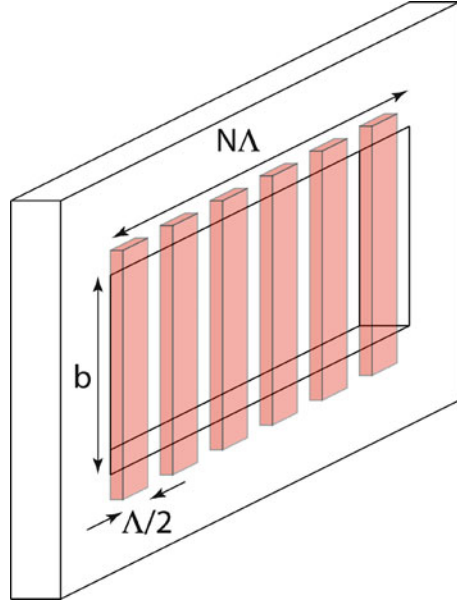
Using the Bessel function identity

$$\exp[i\delta \cos(Kx + \phi)] = \sum_{m=-\infty}^{\infty} J_m(\delta) \exp[im(Kx + \phi + \pi/2)] \quad (2.38)$$

the reflected field just after reflection (in the near field) can be written as

$$E_r = E_0 e^{i\delta_0} \sum_{m=-\infty}^{\infty} J_m(\delta) \exp[im(Kx + \phi + \pi/2)] \quad (2.39)$$

Fig. 2.7 A periodic square grating of N stripes on a rectangular aperture



where $\delta = k_0(n-1)d$. The Fraunhofer diffraction integral for the m th diffraction order selects out each of the Bessel functions as

$$E_m = E_0 e^{i\delta_0} e^{i(mKx+m\phi+m\pi/2)} J_m(k_0(n-1)d) \quad (2.40)$$

The Bessel function is expanded as

$$J_m(x) = \frac{x^m}{2^m m!} \quad (2.41)$$

and the first-order diffraction is then

$$E_{\pm 1} = E_0 e^{i\delta_0} e^{i(\pm Kx \pm \phi \pm \pi/2)} k_0(n-1)d/2 \quad (2.42)$$

with a diffraction efficiency given by

$$\eta = \frac{k_0^2(n-1)^2 d^2}{4} \quad (2.43)$$

for the sinusoidal grating.

If the grating is a square grating (commonly encountered with protein patterning using lithography) as shown in Fig. 2.7, then the diffracted intensity is given by (2.15) modulated with the periodic envelope function of (2.27)

$$I_P = 2I_0 \left(\frac{\Lambda}{2N\lambda R} \right)^2 \left(\frac{\sin(Nk_0\Lambda \sin \theta/2)}{\sin(k_0\Lambda \sin \theta/2)} \right)^2 \text{sinc}^2(k\Lambda \sin \theta/4) [1 + \cos(k\Lambda \sin \theta/2 - \delta)] \quad (2.44)$$

The diffraction efficiency of the square grating is

$$\eta = \left(\frac{2}{\pi} \right)^2 \frac{1 - \cos \delta}{1 + \cos \delta} \approx \left(\frac{2}{\pi} \right)^2 \frac{k_0^2 (n-1)^2 d^2}{4} \quad (2.45)$$

which can be compared with (2.43).

2.1.3 Linear vs. Quadratic Response and Detectability

An important issue for the detectability of small signals is whether linear sensitivity (when signal intensity is proportional to the quantity to be detected), or quadratic sensitivity (when signal intensity is proportional to the squared value of the quantity) gives the best ability to detect small quantities. The answer to this question is generally not fundamental, but depends on details of the noise sources, which in turn usually depend on intensity. For example, diffracted intensity depends on the square of the phase modulation induced by the biolayer, while interferometric intensity (in quadrature) is linear in the phase modulation. Because phase modulation caused by a biolayer is typically less than 1%, linear detection has a much larger absolute modulation caused by a biolayer than diffraction does. On the other hand, interferometric linear detection has a much higher background (lower contrast) that can cause more noise. This is a basic trade-off between the two types of biosensors: linear detection with high-background and low-contrast vs. quadratic detection with low-background and high-contrast. As a general rule, linear detection is more sensitive for detection of very sub-monolayer films in the weak-signal limit, while quadratic detection can have better signal-to-noise in the strong-signal limit.

To make these arguments more quantitative, consider three contributions to the noise

$$I_N^2 = C_{\text{RIN}}^2 I_B^2 \text{BW} + h\nu I_B \text{BW} + C_0 \text{BW} \quad (2.46)$$

where BW refers to the detection bandwidth, and I_B is the background intensity incident on the detector. The first term is relative intensity noise (RIN), the second term is shot noise and the third term is a system noise floor (usually electronic noise). It is important to remember that these noise contributions are frequency dependent, usually with a $1/f$ behavior at low detection frequencies, evolving into white noise at higher detection frequencies. The linear dependence on signal bandwidth is strictly true only for white noise. The signal is characterized by

$$\Delta I_S = I_0 C^{(L)} \Delta \phi + I_0 C^{(Q)} \Delta \phi^2 \quad (2.47)$$

where the first term is the linear dependence and the second term is quadratic, and the intensity I_0 is related to the average intensity on the detector. The signal-to-noise ratio is then

$$S/N = \frac{(\Delta I_S)^2}{I_N^2} = \frac{I_0^2 (C^{(L)} \Delta \phi + C^{(Q)} \Delta \phi^2)^2}{C_R^2 I_B^2 BW + h\nu I_B BW + C_0 BW} \quad (2.48)$$

which can be interpreted in different limits. If RIN is dominant, then the S/N ratio is independent of intensity. If shot noise is dominant, then the S/N ratio increases linearly with intensity. If the system noise floor dominates, then the S/N ratio increases quadratically with intensity. This general behavior of the S/N ratio was discussed and shown in Fig. 1.10.

The relative importance of linear sensitivity vs. quadratic sensitivity depends on the intensity dependence of the signal, and the relationship between I_0 and I_B . To illustrate this, consider the classic performance of the two-port interferometer. The background intensity is given by

$$I_B = \frac{I_0}{2} (1 + m \cos \phi) \quad (2.49)$$

and the signal is obtained as

$$\begin{aligned} I_B + \Delta I_S &= \frac{I_0}{2} (1 + m \cos(\phi + \Delta \phi)) \\ &= \frac{I_0}{2} + I_0 \frac{m}{2} [\cos \phi \cos \Delta \phi - \sin \phi \sin \Delta \phi] \\ &= \frac{I_0}{2} + I_0 \frac{m}{2} \cos \phi \left(1 - \frac{1}{2} \Delta \phi^2 \right) - I_0 \frac{m}{2} \sin \phi \Delta \phi \end{aligned} \quad (2.50)$$

The linear and quadratic coefficients are

$$\begin{aligned} C^{(L)} &= -\frac{m}{2} \sin \phi \\ C^{(Q)} &= -\frac{m}{4} \cos \phi \end{aligned} \quad (2.51)$$

The S/N ratio for the simple interferometer is

$$S/N = \frac{m^2 (\sin \phi \Delta \phi + \frac{1}{2} \cos \phi \Delta \phi^2)^2}{C_R^2 (1 + m \cos \phi)^2 BW + 2h\nu (1 + m \cos \phi) BW + C_0 BW} \quad (2.52)$$

When the detection is limited by RIN, then for optimized quadratic sensitivity to exceed optimized linear sensitivity the condition

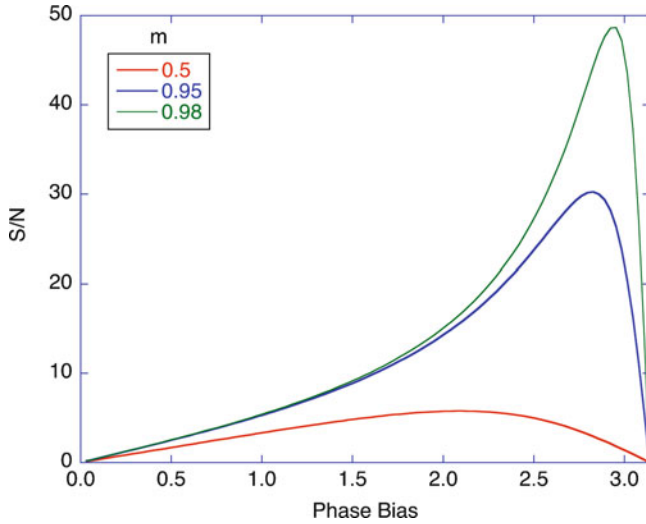


Fig. 2.8 S/N ratio for $\Delta\phi = 0.002$ and a 0.01% RIN for a two-wave interferometer as a function of phase bias for three contrasts of $m = 0.5, 0.95$ and 0.98 . A phase bias near π reduces the background noise while still allowing a (small) linear response

$$\frac{\frac{m}{2}\Delta\phi^2}{C_R(1-m)\sqrt{BW}} > \frac{m\Delta\phi}{C_R\sqrt{BW}} \quad (2.53)$$

should hold, which gives

$$\Delta\phi_{\text{quad}} > 2(1-m) \quad (2.54)$$

and a quadratic detection would be favored over linear only if the contrast m of the interferometer can be constructed to be larger than

$$m > 1 - \frac{k(n - n_m)d}{2} \quad (2.55)$$

or $m > 0.998$ for 1 nm of protein in a water ambient. Otherwise linear detection in quadrature is favored.

A different approach to this analysis sets a maximum practical contrast m , and considers what phase bias ϕ of the interferometer gives the largest S/N ratio. This yields phase biases that are between $\pi/2$ and π , but approaching π as m approaches unity, because of the suppression of the noise by suppressing the background intensity. Indeed, most diffraction-based biosensors have a slight structural bias that introduces a small linear dependence in addition to the quadratic dependence of diffraction on the magnitude of the grating. This small linear dependence can dominate the signal for small grating amplitudes, while giving a strong S/N ratio because of the low background. This is shown in Fig. 2.8 for a two-mode

interferometer for contrasts of $m = 0.50, 0.95$ and 0.98 . A phase bias near π reduces the background noise while still allowing a (small) linear response. Diffraction-based biosensors are discussed in Chap. 7.

2.2 Fourier Optics

Fourier optics provides a particularly useful point of view of diffraction. The foundation of Fourier optics is the recognition that the Fraunhofer integral is equivalent to a mathematical Fourier transform. Furthermore, a lens can perform the equivalent of a Fourier transform under appropriate object and screen distances relative to the focal length of the lens. This allows the far-field Fraunhofer condition to be realized on focal planes of lenses in compact optical systems. Fourier optics combines the power of Fourier analysis with the ease of use of lenses, leading to a fruitful approach to understand image formation as well as diffraction phenomena.

To begin, an object is considered to be composed of a superposition of multiple spatial frequencies with periodicities Λ and associated K -vectors. For quasi-planar objects, the K -vectors lie in the object plane. An example of a single spatial frequency is shown in Fig. 2.9, in which the object amplitude is

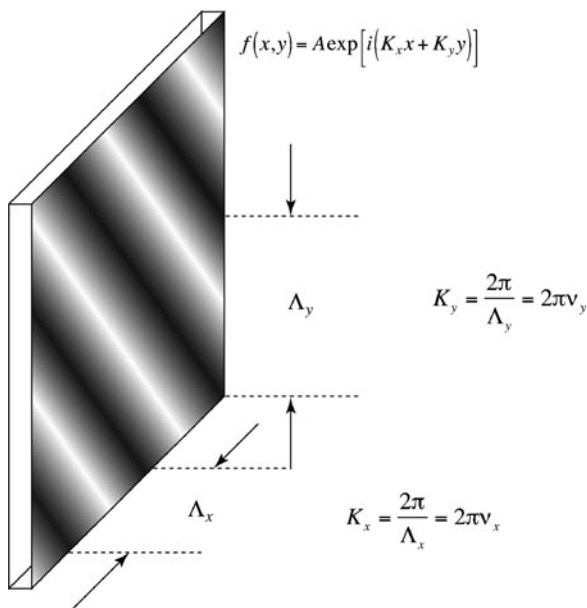


Fig. 2.9 A planar object consisting of a single spatial periodicity. The periodicities in the x and y directions are Λ_x and Λ_y , with the associated spatial frequencies v_x and v_y

$$f(x, y) = A \exp[i(K_x x + K_y y)] \quad (2.56)$$

where the components of the K -vector are

$$\begin{aligned} K_x &= \mathbf{K} \cdot \hat{\mathbf{x}} = K \cos \phi = \frac{2\pi}{\Lambda} \cos \phi = \frac{2\pi}{\Lambda_x} = 2\pi v_x \\ K_y &= \mathbf{K} \cdot \hat{\mathbf{y}} = K \sin \phi = \frac{2\pi}{\Lambda} \sin \phi = \frac{2\pi}{\Lambda_y} = 2\pi v_y \end{aligned} \quad (2.57)$$

The spatial frequencies that appear in the Fraunhofer integral and the Fourier integral are

$$v_x = \frac{1}{\Lambda_x} \quad v_y = \frac{1}{\Lambda_y} \quad (2.58)$$

To see the connection between spatial periodicities and the Fourier transform, consider a single harmonic grating

$$f(x, y) = A \exp[i(K_x x + K_y y)] = A \exp[i2\pi(v_x x + v_y y)] \quad (2.59)$$

The diffraction integral is

$$\begin{aligned} E_d(\theta_x, \theta_y) &= \iint f(x, y) \exp\left[-i\frac{2\pi}{\lambda}(\sin \theta_x x + \sin \theta_y y)\right] dx dy \\ &= \iint A \exp[i2\pi(v_x x + v_y y)] \exp\left[-i\frac{2\pi}{\lambda}(\sin \theta_x x + \sin \theta_y y)\right] dx dy \\ &= A\delta\left(\left(v_x - \frac{\sin \theta_x}{\lambda}\right), \left(v_y - \frac{\sin \theta_y}{\lambda}\right)\right) \end{aligned} \quad (2.60)$$

which defines a delta function at the diffraction angles given by

$$\begin{aligned} \sin \theta_x &= \lambda v_x \\ \sin \theta_y &= \lambda v_y \end{aligned} \quad (2.61)$$

This can be interpreted physically as a mapping of a unique scattering angle to a unique point on the observation (Fourier) plane. In other words, all rays leaving the object with a given angle converge to a single point on the Fourier plane. This viewpoint becomes especially easy to visualize when a lens performs the Fourier transform (see Fig. 2.12 below).

2.2.1 Fresnel Diffraction

If $f(x, y)$ is a combination of harmonics

$$f(x, y) = \iint F(v_x, v_y) \exp[i2\pi(v_x x + v_y y)] dv_x dv_y \quad (2.62)$$

then the transmitted field just after the screen is

$$E_s(x, y, z) = \iint F(v_x, v_y) \exp[i2\pi(v_x x + v_y y)] \exp[ik_z z] dv_x dv_y \quad (2.63)$$

where

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} = 2\pi \sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2} \quad (2.64)$$

For small angles this is approximately

$$k_z z = \frac{2\pi z}{\lambda} \sqrt{1 - \lambda^2 v_x^2 - \lambda^2 v_y^2} \approx \frac{2\pi z}{\lambda} - \pi \lambda z (v_x^2 + v_y^2) \quad (2.65)$$

The transmitted field can then be written as

$$\begin{aligned} E_s(x, y, z) &= \iint F(v_x, v_y) \exp[i2\pi(v_x x + v_y y)] \exp(ikz) \exp[-i\pi\lambda z(v_x^2 + v_y^2)] dv_x dv_y \\ &= \iint F(v_x, v_y) \exp[i2\pi(v_x x + v_y y)] H(v_x, v_y) dv_x dv_y \end{aligned} \quad (2.66)$$

where

$$H(v_x, v_y) = \exp(ikz) \exp[-i\pi\lambda z(v_x^2 + v_y^2)] \quad (2.67)$$

is known as the free-space propagator in the Fresnel approximation.

The Green's function solution (response to a delta function) for an impulse function on the object plane is the inverse Fourier transform of the free-space propagator

$$G(x, y, z) = \frac{1}{i\lambda z} \exp(ikz) \exp\left[ik \frac{x^2 + y^2}{2z}\right] \quad (2.68)$$

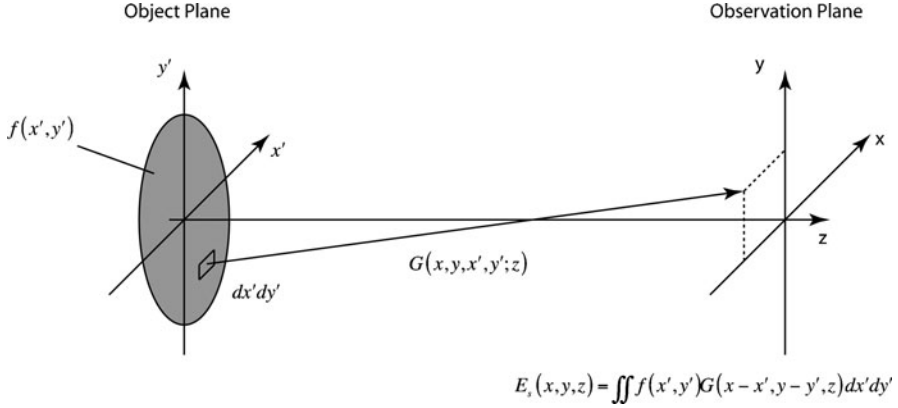


Fig. 2.10 Fresnel diffraction geometry. A delta function on the object plane has an associated Green's function. Convolution of the object function $f(x', y')$ with the Green's function gives the field at the observation plane

so that the general scattered field at z is a convolution of the object function over the Green's function

$$\begin{aligned} E_s(x, y, z) &= \iint f(x', y') G(x - x', y - y', z) dx' dy' \\ &= \frac{1}{i\lambda z} \exp(ikz) \iint f(x', y') \exp\left[ik \frac{(x - x')^2 + (y - y')^2}{2z}\right] dx' dy' \end{aligned} \quad (2.69)$$

This is the Fresnel integral that can be used for general scattering problems, including holography and lens-free imaging. Holographic reconstruction in digital holography uses the Fresnel integral explicitly to reconstruct three-dimensional aspects of an object. The scattering geometry with the object and observation plane are shown in Fig. 2.10.

2.2.2 Optical Fourier Transforms

The Fresnel regime converts to the Fraunhofer (far-field) regime when the distance to the observation plane is much larger than the size of the scattering object. The cross-over condition is captured quantitatively in terms of the Fresnel number

$$N_F = \frac{a^2}{\lambda L} \quad (2.70)$$

where a is the size of the scattering object and L is the distance to the detection plane. The scattering is in the Fresnel regime when $N_F > 1$ (but still $L \gg \lambda$), and in the Fraunhofer regime when $N_F < 1$. The cross-over is not a sharp threshold, so it is

best to satisfy $N_F \ll 1$ to ensure that the scattering is the Fraunhofer regime. As an example, a 10 μm object illuminated with visible light scatters to the Fraunhofer regime for an observation distance greater than a millimeter.

In the far-field regime at a large distance L from the object screen the Fresnel integral becomes

$$\begin{aligned} E_F(x, y) &= \frac{1}{i\lambda L} \exp(ikL) \exp\left[ik \frac{x^2 + y^2}{2L}\right] \iint f(x', y') \exp\left[-ik \frac{xx' + yy'}{L}\right] dx' dy' \\ &= \frac{1}{i\lambda L} \exp(ikL) \exp\left[ik \frac{x^2 + y^2}{2L}\right] F(v_x, v_y) \end{aligned} \quad (2.71)$$

in which $F(v_x, v_y)$ is the Fourier transform of $f(x, y)$, and where the spatial frequencies of the object are identified as

$$v_x = \frac{x}{\lambda L} \quad v_y = \frac{y}{\lambda L} \quad (2.72)$$

This result has the important consequence that the far-field diffraction pattern is proportional to the Fourier transform of the object modulation function. It is an easy way to do a Fourier transform – just diffract to the far-field. However, it is usually convenient to work with lenses, and these too perform an optical Fourier transform.

A single simple (thin) lens of focal length f can be viewed mathematically as imposing a quadratic phase profile on an incoming wave by the transmission function

$$t(x, y) = \exp\left[-i\pi \frac{(x^2 + y^2)}{\lambda f}\right] \quad (2.73)$$

The geometry of the Fourier lens is shown in Fig. 2.11 with the distance L from the object to the lens, and with the observation plane a distance f from the lens. The angle-to-point transformation that is achieved by a Fourier lens is illustrated in Fig. 2.12. Using (2.73) in (2.71) gives

$$E_F(x, y) = \frac{1}{i\lambda f} \exp[ik(L + f)] \exp\left[-i\pi \frac{(x^2 + y^2)(L - f)}{L\lambda f}\right] F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right) \quad (2.74)$$

This is proportional to the Fourier transform of $f(x, y)$, but with a quadratic phase factor. When the object distance $L = f$, then

$$E_F(x, y) = \frac{1}{i\lambda f} \exp(ik2f) F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right) \quad (2.75)$$

which is the desired Fourier transform with a simple phase factor.

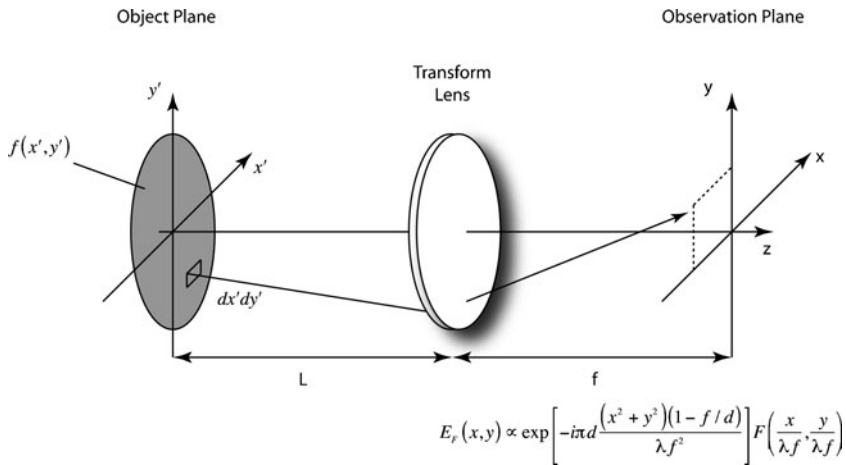


Fig. 2.11 Configuration of a Fourier lens. When the object, lens, and observation plane are in a $1f$ - $1f$ condition (both the object plane and the image plane are a distance f from the lens), then the field on the observation plane (now called the Fourier plane) is the Fourier transform of the object amplitude

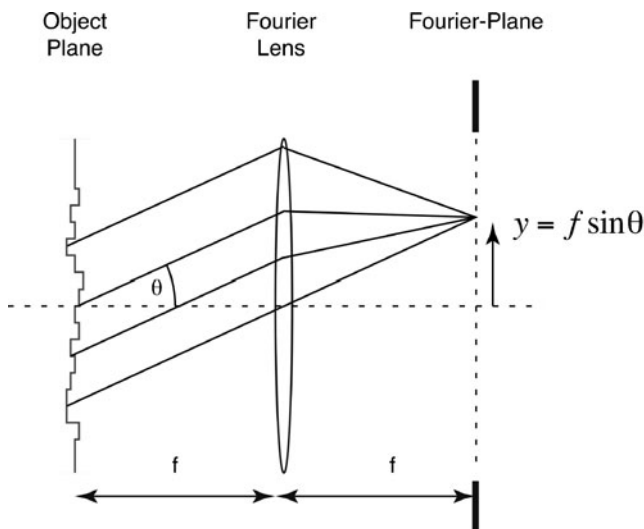


Fig. 2.12 A Fourier lens in a $1f$ - $1f$ configuration. All rays emitted from the object at a given angle are focused by the lens to a point (actually a point spread function) on the Fourier plane

2.2.3 Gaussian Beam Optics

In laser optics, one of the most common intensity distributions is the Gaussian beam. This is a beam that has the transverse intensity profile given by

$$I(x, y) = S_0 \exp(-(x^2 + y^2)/w_0^2) \quad (2.76)$$

with an integrated power

$$P = \iint I(x, y) dx dy = S_0 \pi w_0^2 \quad (2.77)$$

Because a Gaussian beam has a limited lateral extent, it self-diffracts into the far-field, even without any aperture to diffract off of. This is an important part of the principle of wave propagation. No matter how a field is spatially modulated, whether by a physical aperture, or simply because fields are inhomogeneous, these fields propagate as partial waves with relative phases and relative path lengths that cause interference and hence diffraction.

For a Gaussian beam at its focal plane, the dimensionless field amplitude is

$$E = \sqrt{I} = \sqrt{S_0} \exp(-(x^2 + y^2)/2w_0^2) \quad (2.78)$$

and the diffraction integral is

$$E_d = -i\sqrt{S_0} \int_{-\infty}^{\infty} \exp(-(x^2 + y^2)/2w_0^2) e^{ikx \sin \theta} e^{iky \sin \psi} dx dy \quad (2.79)$$

This has the diffraction field

$$E(\theta, \Psi) = -i\sqrt{S_0} \frac{kw_0^2}{R} \exp\left(-\frac{k^2 w_0^2}{2} (\sin^2 \theta + \sin^2 \psi)\right) \quad (2.80)$$

which is again a Gaussian, but now expressed as an angular distribution in angles θ and Ψ . When Fourier transformed using a lens of focal distance f , the field at the Fourier plane is

$$E(x', y') = -i\sqrt{S_0} \frac{2\pi w_0^2}{f\lambda} \exp\left(-\frac{2\pi^2 w_0^2}{(f\lambda)^2} (x'^2 + y'^2)\right) \quad (2.81)$$

illustrated in Fig. 2.13 as the focal waist of a Gaussian beam. The intensity at the Fourier plane is

$$I(x', y') = S_0 \left(\frac{2\pi}{f\lambda}\right)^2 w_0^4 \exp\left(-\left(\frac{2\pi}{f\lambda}\right)^2 w_0^2 (x'^2 + y'^2)\right) \quad (2.82)$$

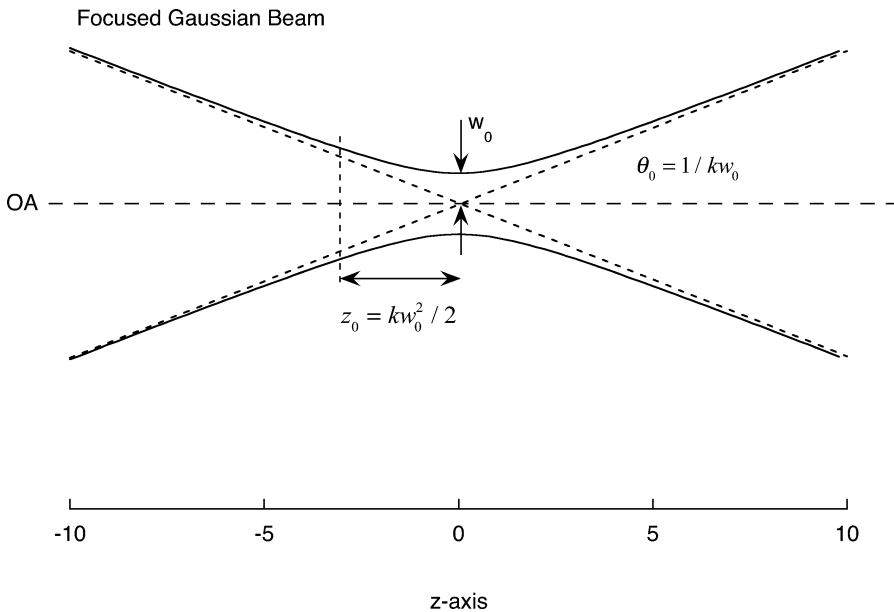


Fig. 2.13 Profile of a focused Gaussian beam. The beam waist w_0 occurs at the focal plane at $z = 0$. At distances large relative to the Rayleigh range z_0 the beam converges or diverges with angle θ_0 . The beam evolution is self diffraction

with a beam radius given by

$$w'_0 = \frac{f\lambda}{2\pi w_0} \quad (2.83)$$

This result is of special importance in Fourier optics. It expresses the inverse relationship between the widths of Fourier-transform pairs of Gaussians. Tightly focused beams are transformed into broad beams, and vice versa. This inverse relation also holds generally for smooth non-Gaussian beams, but with different numerical factors.

2.3 Dipoles and Rayleigh Scattering

The origin of all light-scattering phenomena is the interaction of the electric and magnetic fields of a light wave with electric and magnetic multipole moments of electrons in atoms and molecules. For optical interactions, the strongest interactions are usually through the electric dipole interaction, which tends to be stronger than the magnetic dipole interaction by the ratio $e^2/\hbar c = 1/137$. However, in cases in

which the electric dipole interaction is not allowed for symmetry reasons, magnetic dipole or other multipole moments can dominate. In this book, we consider electric dipole interaction exclusively.

All molecules have a molecular polarizability, that is, a tensor relation between the applied field and the structure and symmetry of the molecule

$$p^i = \varepsilon_m \alpha_j^i E^j \quad (2.84)$$

where α_j^i is the molecular polarizability tensor and ε_m is the isotropic dielectric constant of the surrounding medium. Because protein molecules lack general symmetry, the tensor polarizability is simplified through configurational averaging to a scalar property relating the induced dipole moment to the applied electric field.

$$\mathbf{p} = \varepsilon_m \alpha \mathbf{E} \quad (2.85)$$

Many protein molecules are globular in structure, such as the immunoglobulins, and to lowest approximation may be viewed as dielectric spheres. The polarizability of a dielectric sphere of radius a is

Polarizability of a dielectric sphere:

$$\alpha = \frac{(\varepsilon - \varepsilon_m)}{(\varepsilon + 2\varepsilon_m)} 4\pi\varepsilon_0 a^3 \quad (2.86)$$

where ε is the dielectric function of the sphere. For comparison, the polarizability of a metallic sphere is

Polarizability of a metallic sphere:

$$\alpha = 4\pi\varepsilon_0 a^3 \quad (2.87)$$

Given an induced dipole moment p , the resulting dipole field is

Dipole field:

$$\mathbf{E}(\mathbf{r}) = \frac{3(\mathbf{p} \cdot \mathbf{r})\mathbf{r} - r^2\mathbf{p}}{4\pi\varepsilon_m\varepsilon_0 r^5} \quad (2.88)$$

which falls off as the inverse cube of the distance from the dipole. The dipole field of the dipole induced by an incident field is the origin of molecular scattering and ultimately of the refractive index of a layer of biological molecules.

The first approximation for understanding molecular scattering is scattering in the Rayleigh limit. The incident electric and magnetic fields in a dielectric medium with $k_m = n_m k_0$ are

$$\begin{aligned} \mathbf{E}_{\text{inc}} &= \varepsilon_0 E_0 \mathbf{e}^{ik_m \hat{n}_0 \cdot \mathbf{x}} \\ \mathbf{B}_{\text{inc}} &= \frac{\hat{n}_0}{v_m} \times \mathbf{E}_{\text{inc}} \end{aligned} \quad (2.89)$$

where \mathbf{n}_0 is the unit vector in the direction of the incident field. The scattered fields from the induced dipole are

$$\begin{aligned}\mathbf{E}_{\text{sc}} &= k_m^2 \frac{e^{ikr}}{4\pi\epsilon_m\epsilon_0 r} [(\mathbf{n} \times \mathbf{p}) \times \mathbf{n}] \\ \mathbf{B}_{\text{sc}} &= \frac{1}{v_m} \mathbf{n} \times \mathbf{E}_{\text{sc}}\end{aligned}\quad (2.90)$$

where \mathbf{n} is the unit vector in the direction of the scattered wave. The Poynting vector is

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \quad (2.91)$$

and the differential scattering cross-section is constructed from the Poynting vector along the scattering direction

$$\frac{d\sigma}{d\Omega}(\theta, \phi) = \frac{r^2 |\mathbf{S}(\theta, \phi)|^2}{\frac{1}{2} v \epsilon \epsilon_0 |\mathbf{E}_0|^2} = \frac{k_m^4}{(4\pi\epsilon\epsilon_0)^2 E_0^2} |\mathbf{p}|^2 \frac{1}{2} (1 + \cos \theta) \quad (2.92)$$

where θ is measured relative to the incident k -vector direction.

Example: Small Dielectric Sphere in Vacuum

In the case of scattering by a small dielectric sphere, the induced dipole is

$$\mathbf{p} = 4\pi\epsilon_0 \left(\frac{\epsilon - 1}{\epsilon + 2} \right) a^3 \mathbf{E}_{\text{inc}} \quad (2.93)$$

with a scattering cross-section

$$\frac{d\sigma}{d\Omega} = k^4 a^6 \left| \frac{\epsilon - 1}{\epsilon + 2} \right|^2 |\boldsymbol{\epsilon}^* \cdot \boldsymbol{\epsilon}_0|^2 \quad (2.94)$$

The scattering cross-section is decomposed into expressions for each incident polarization relative to the scattering plane

$$\begin{aligned}\frac{d\sigma_{\parallel}}{d\Omega} &= \frac{k^4 a^6}{2} \left| \frac{\epsilon - 1}{\epsilon + 2} \right|^2 \cos^2 \theta \\ \frac{d\sigma_{\perp}}{d\Omega} &= \frac{k^4 a^6}{2} \left| \frac{\epsilon - 1}{\epsilon + 2} \right|^2\end{aligned}\quad (2.95)$$

When these are averaged over polarization, they become

Rayleigh differential cross-section:

$$\frac{d\sigma}{d\Omega} = k^4 a^6 \left| \frac{\varepsilon - 1}{\varepsilon + 2} \right|^2 \frac{1}{2} (1 + \cos^2 \theta) \quad (2.96)$$

with the familiar angular distribution of scattered light with symmetry between forward-scattering and back-scattering. The differential cross-section is integrated over all scattering angles to obtain the total cross-section

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \frac{8\pi}{3} k^4 a^6 \left| \frac{\varepsilon - 1}{\varepsilon + 2} \right|^2 \quad (2.97)$$

The scattering cross-section depends on the sixth power of the radius and on the fourth power of the frequency. The cross-section has units of area, as seen when it is expressed as

Rayleigh total cross-section:

$$\sigma = \left(\frac{8}{3} (ka)^4 \left| \frac{\varepsilon - 1}{\varepsilon + 2} \right|^2 \right) \pi a^2 \quad (2.98)$$

The effective scattering area of a dielectric sphere is proportional to the cross-sectional area of the sphere, but reduced by the factor in parentheses in front. For biological molecules with a radius of 1 nm and a dielectric constant of 2 the reduction is approximately 2×10^{-10} . The effective cross-section for such a molecule is $\sigma = 10^{-23} \text{ cm}^2$.

Example: Small Dielectric Sphere in Water

There are many applications in which the dielectric sphere is in a water environment. It is important in this case to be clear about the vacuum k -vector k_0 , and the medium k -vector, which is $k_m = n_m k_0$. In addition, the dielectric water medium modifies the polarization as

$$\alpha = \frac{(\varepsilon - \varepsilon_m)}{(\varepsilon + 2\varepsilon_m)} 4\pi\varepsilon_0 a^3 \quad (2.99)$$

Therefore, the differential cross-section in water is

$$\frac{d\sigma}{d\Omega} = \varepsilon_m^2 k_0^4 a^6 \left| \frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m} \right|^2 \frac{1}{2} (1 + \cos^2 \theta) \quad (2.100)$$

Table 2.1 Rayleigh scattering

In air	In medium
$\alpha = \frac{(\varepsilon-1)}{(\varepsilon+2)} 4\pi\varepsilon_0 a^3$	$\alpha = \frac{(\varepsilon-\varepsilon_m)}{(\varepsilon+2\varepsilon_m)} 4\pi\varepsilon_0 a^3$
$\mathbf{p} = \left(\frac{\varepsilon-1}{\varepsilon+2}\right) 4\pi\varepsilon_0 a^3 \mathbf{E}_{\text{inc}}$	$\mathbf{p} = \left(\frac{\varepsilon-\varepsilon_m}{\varepsilon+2\varepsilon_m}\right) 4\pi\varepsilon_m \varepsilon_0 a^3 \mathbf{E}_{\text{inc}}$
$\frac{d\sigma}{d\Omega} = k^4 a^6 \left \frac{\varepsilon-1}{\varepsilon+2}\right ^2 \frac{1}{2} (1 + \cos^2\theta)$	$\frac{d\sigma}{d\Omega} = \varepsilon_m^2 k_0^4 a^6 \left \frac{\varepsilon-\varepsilon_m}{\varepsilon+2\varepsilon_m}\right ^2 \frac{1}{2} (1 + \cos^2\theta)$
$\sigma = \left(\frac{8}{3}(ka)^4 \left \frac{\varepsilon-1}{\varepsilon+2}\right ^2\right) \pi a^2$	$\sigma = \left(\frac{8}{3}(k_m a)^4 \left \frac{\varepsilon-\varepsilon_m}{\varepsilon+2\varepsilon_m}\right ^2\right) \pi a^2$

with the corresponding total cross-section (real dielectric function: no absorption)

$$\sigma = \left(\frac{8}{3}(k_m a)^4 \left|\frac{\varepsilon - \varepsilon_m}{\varepsilon + 2\varepsilon_m}\right|^2\right) \pi a^2 \quad (2.101)$$

Selected formulas for Rayleigh scattering are given in Table 2.1.

2.4 Refractive Index of a Dilute Molecular Film

Refractive index is a macroscopic property of a material that relates the phase velocity of a light wave to frequency and wavelength through the expression

$$v_p = \frac{\omega}{k} = \frac{c}{n(\omega)} \quad (2.102)$$

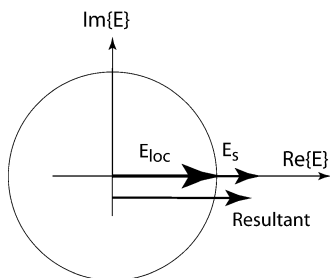
In this section, we approach the derivation of the refractive index using concepts from diffraction theory.

An apparent paradox arises in relation to the phase shifts experienced upon scattering from a dipole and the phase shifts that are detected in the far-field that are attributed to, or interpreted as, refractive index. The central paradox is that the scattered wave from a dipole is in phase with the incident field. If the scattered field and incident field have the same phase, then why is a phase shift detected in the far-field? The answer to this paradox comes from diffraction theory (scalar Kirchhoff diffraction suffices) due to the difference between a continuous field distribution compared with the discrete scatterer. An outline of the physics is shown in Fig. 2.14. Two phasor diagrams are shown, one in the near-field and one in the far-field. In the near-field, the local field and the scattered field are in phase. However, as the local field propagates to the far-field, it acquires a $\pi/2$ phase shift, while the scattered field does not. Therefore, in the far-field, the scattered field is in phase quadrature with the continuous field, which is manifested as a phase shift. Mathematically, integrating the diffraction integral over a finite-size continuous distribution yields a factor of $-i$, while a discrete delta-function scatterer simply produces a spherical wave that propagates to the far-field without any phase shift.

Scattering and Phase Quadrature

Near Field: in-phase

$$E_s \sim fE_{loc}$$



Far Field: in-quadrature

$$E_{far} = fE_{loc} \frac{e^{(ikr - i\omega t)}}{r} + -i \frac{1}{r} \int e^{(ikr - i\omega t)} E_{loc} d^2x$$

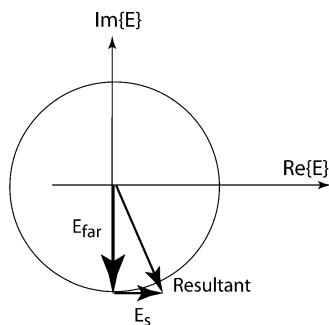


Fig. 2.14 Light scattering from a molecule and the $\pi/2$ phase shift. In the near field, the scattered wave is in-phase with the incident field. In the far-field, a $\pi/2$ phase shift has occurred that places the scattered wave in phase-quadrature with the original wave, representing a phase modulation related to the refractive index

2.4.1 Phase Shift of a Single Molecule in a Focused Gaussian Beam

A single molecule in a tightly focused Gaussian beam will scatter light and impart a small phase shift on the direct wave in the far-field. This phase shift is a direct consequence of diffraction and interference. An interesting question is whether this single-molecule phase shift might be detectable experimentally. The induced dipole from (2.86) on the optic axis and on the focal plane of the Gaussian beam is

$$\mathbf{p} = \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) a^3 \mathbf{E}_0 \quad (2.103)$$

where the field on the optic axis is $E_0 = \sqrt{S_0}$. The local scattered electric field is then

$$\mathbf{E}_{sc} = k^2 \frac{e^{ikr}}{r} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) a^3 \sqrt{S_0} \quad (2.104)$$

This field is transformed to the Fourier plane using a lens of focal length f and diameter $D > \frac{1}{2\pi} \frac{f\lambda}{w_0}$. The power collected by this lens is

$$P_{\text{scat}} = \frac{\pi}{4} D^2 \frac{k^4}{f^2} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right)^2 a^6 S_0 \quad (2.105)$$

and the field at the Fourier plane is

$$E_{\text{scat}} = \frac{k^2}{f} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) a^3 \sqrt{S_0} \quad (2.106)$$

The interfering fields (the original field plus the scattered field) at the Fourier plane are now

$$\begin{aligned} E &= E_{\text{scat}} + E_0 \\ &= \frac{k\sqrt{S_0}}{f} \left[ka^3 \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) - iw_0^2 \exp \left(-\frac{2\pi^2 w_0^2}{(f\lambda)^2} (x'^2 + y'^2) \right) \right] \end{aligned} \quad (2.107)$$

The phase shift on the optic axis of the detected Gaussian field is given by

$$\tan \phi \approx \frac{ka^3}{w_0^2} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \quad (2.108)$$

For a molecule of radius $a = 3$ nm with $\varepsilon = 1.43^2 = 2$ at a wavelength of 500 nm and a beam radius of 0.5 μm , this phase shift is about 1×10^{-7} radians.

To estimate how detectable this phase shift is, consider a situation when the $\pi/2$ phase shift between the Gaussian field and the scattered field is shifted to zero to place it in the condition of constructive interference. Then the intensity at the Fourier plane is

$$\begin{aligned} I &= \frac{k^2 S_0}{f^2} \left[w_0^4 \exp \left(-\frac{4\pi^2 w_0^2}{(f\lambda)^2} (x'^2 + y'^2) \right) \right. \\ &\quad \left. + 2w_0^2 ka^3 \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \exp \left(-\frac{2\pi^2 w_0^2}{(f\lambda)^2} (x'^2 + y'^2) \right) \right] \end{aligned} \quad (2.109)$$

When this is integrated over the Fourier plane, it yields

$$\int I da = S_0 \left[\pi w_0^2 + 4\pi ka^3 \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \right] \quad (2.110)$$

and the relative modulation is

$$\frac{\Delta I}{I} = \frac{4ka^3}{w_0^2} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \quad (2.111)$$

which is four times the value of the phase shift along the optic axis. One factor of two is from the interference cross terms, and the other is from the integration over the Gaussian profile. Therefore, for typical values of the parameters, the relative intensity modulation from a single macromolecules is about one part per million.

The approximate shot noise that this corresponds to is a photon number of 1×10^{12} , which is about 400 nJ. If this energy is detected in a millisecond detection time, the power at the detector would be about 400 μ W, which is an easy power level to achieve with conventional photodiodes. Therefore, a single macromolecule could be detected interferometrically under shot-noise-limited detection conditions.

2.4.2 Phase Shift from a Dilute Collection of Molecules

The effective phase shift of a collection of molecules in a Gaussian beam can be obtained by extending the single-molecule result. The scattered field for a molecule off the optic axis is

$$E'_{\text{scat}} = \frac{k^2}{f} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) a^3 \sqrt{S_0} \exp(-r^2/2w_0^2) \exp(ikxx'/f) \exp(ikyy'/f) \quad (2.112)$$

where the primed coordinates are in the Fourier (detection) plane, and the unprimed coordinates are in the focal plane. The contribution of a distribution of molecules is obtained using the integration

$$\int_0^\infty 2\pi r e^{-r^2/2w_0^2} dr = 2\pi w_0^2 \quad (2.113)$$

to give, for a surface density $N_A = N/\pi w_0^2$, the effective number of scatterers

$$\frac{N}{\pi w_0^2} 2\pi w_0^2 = 2N \quad (2.114)$$

The phase is then

$$\tan \phi \approx 2N \frac{ka^3}{w_0^2} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \quad (2.115)$$

This is turned into an effective index of refraction for a layer with a thickness equal to $2a$ as

$$\phi \approx 2N \frac{ka^3}{w_0^2} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) = k(n_{\text{av}} - 1)2a \quad (2.116)$$

or

$$n_{\text{av}} = 1 + N \frac{a^2}{w_0^2} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) = 1 + \pi N_{\text{A}} a^2 \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \quad (2.117)$$

where N_{A} is the area number density (particles per area). This is restated in terms of the volume fraction f_v using

$$f_v = \frac{(4/3)\pi a^3}{2a} N_{\text{A}} = \frac{2}{3} \pi a^2 N_{\text{A}} \quad (2.118)$$

to give

$$n_{\text{av}} = 1 + \frac{3}{2} f_v \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) \quad (2.119)$$

which is consistent with the Maxwell Garnett effective medium in the Rayleigh limit, which is described next.

2.5 Local Fields and Effective Medium Approaches

Effective medium theories seek to convert a distribution of discrete and stochastic elements into an effective homogeneous medium that behaves with appropriately averaged properties. For instance, the optical effects of a dilute dispersion of particles on a surface (such as biomolecules in optical biosensors) can be treated as if the particles constituted a thin layer having an average refractive index. These approaches can never capture the “full” physics, especially when fluctuations dominate the optical behavior, but they give a heuristic approach, usually with analytic formulas that can be evaluated and compared against experiment.

2.5.1 Local Fields and Depolarization

The scattering from a dilute collection of scatterers as a function of increasing scatterer density is straightforward until multiple scattering becomes significant. In the case of discrete scatterers, the scattered field from each is added to the total field self consistently as a new incident field that is scattered by the collection. To treat this problem explicitly using self-consistent approaches is numerically exhaustive, such as through the discrete dipole approximation [2]. A simpler approach, that works well when the medium is dense, is a mean field approach that uses the mean

local fields of neighboring induced dipoles as a depolarization field experienced by an individual scatterer.

In the case of a spherical dielectric particle, the polarizability is the well-known result

$$\alpha = 4\pi\epsilon_0 a^3 \frac{(\epsilon - \epsilon_m)}{(\epsilon + 2\epsilon_m)} \quad (2.120)$$

The uniform field inside the particle is reduced from the field nearby as

$$\mathbf{E}_{\text{in}} = \mathbf{E}_{\text{near}} + \frac{\mathbf{P}}{3\epsilon_0} \quad (2.121)$$

If the material has high symmetry or is isotropic, then the average near-field is zero, and the induced dipole moment of a molecule at the center is

$$\mathbf{p} = \alpha\epsilon_m(\mathbf{E} + \mathbf{E}_{\text{in}}) \quad (2.122)$$

For a collection of dipoles with number density N , the polarizability is

$$\mathbf{P} = N\alpha\epsilon_m \left(\mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \right) = \chi\epsilon_0 \mathbf{E} \quad (2.123)$$

Solving for χ yields

$$\chi = \frac{1}{\epsilon_0} \frac{N\alpha\epsilon_m}{1 - \frac{N\alpha\epsilon_m}{3\epsilon_0}} \quad (2.124)$$

Using the relationship $\epsilon = 1 + \chi$ yields the Clausius–Mossotti relation

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{\epsilon_m}{3\epsilon_0} \sum_j N_j \alpha_j \quad (2.125)$$

This expression includes the local fields generated by nearby molecules. The Clausius–Mossotti relationship is most accurate for dilute systems such as gases. However, it still holds as a heuristic relationship even for denser liquids and solids.

2.5.2 Effective Medium Models

There are many different approaches to effective medium approximations. Each is equivalent in the dilute Rayleigh limit, but each is slightly different in the case of mixtures of different materials. Furthermore, different approximations relate to

different situations. For instance, small spherical inclusions distributed inside a homogeneous medium are best approximated by Maxwell Garnett formula, while two uniformly mixed phases are best approximated by the Bruggeman formula.

2.5.2.1 Maxwell Garnett

When the collection of scatterers is no longer dilute, but is not sufficiently dense or symmetric to follow the Clausius–Mossotti relation, then a working alternative is the Maxwell Garnett approximation. The Maxwell Garnett model assumes a collection of small particles (inclusions) of material with bulk index n with *radius much smaller than a wavelength* embedded in a medium with *a thickness much larger than the diameters of the spheres*.

The average field is

$$\langle \mathbf{E} \rangle = (1 - f) \langle \mathbf{E}_m \rangle + \sum_k f_k \langle \mathbf{E}_k \rangle \quad (2.126)$$

where $\langle \mathbf{E}_m \rangle$ is the average field in the medium. The volume fraction of each particle sums to the total volume fraction

$$\sum_k f_k = f \quad (2.127)$$

The average polarization is

$$\langle \mathbf{P} \rangle = (1 - f) \langle \mathbf{P}_m \rangle + \sum_k f_k \langle \mathbf{P}_k \rangle \quad (2.128)$$

which is related to the average susceptibility tensor by

$$\langle \mathbf{P} \rangle = \epsilon_0 \vec{\chi}_{av} \langle \mathbf{E} \rangle \quad (2.129)$$

The total self-consistent field now satisfies

$$(1 - f)(\vec{\epsilon}_{av} - \vec{\epsilon}_m) \langle \mathbf{E}_m \rangle + \sum_k f_k (\vec{\epsilon}_{av} - \vec{\epsilon}) \langle \mathbf{E}_k \rangle = 0 \quad (2.130)$$

The next assumption is the central assumption of the Maxwell Garnett model. It states that the uniform field inside the k th particle is related to the external uniform field in the medium through the tensor relation

$$\mathbf{E}_k = \vec{\lambda}_k \mathbf{E}_m \quad (2.131)$$

Table 2.2 Geometric factors

Prolate spheroids			Oblate spheroids		
b/a	L ₁	L ₂ = L ₃	a/b	L ₁	L ₂ = L ₃
0 (needle)	0	0.5	0 (disk)	1.0	0.0
0.2	0.056	0.472	0.2	0.750	0.125
0.4	0.134	0.433	0.4	0.588	0.206
0.6	0.210	0.395	0.6	0.478	0.261
0.8	0.276	0.362	0.8	0.396	0.302
1.0	0.333	0.333	1.0	0.333	0.333

for the tensor $\vec{\lambda}$. The principal components of the $\vec{\lambda}$ tensor along the principal axes of the ellipsoid are

$$\lambda_j = \frac{\epsilon_m}{\epsilon_m + L_j(\epsilon - \epsilon_m)} \quad (2.132)$$

and the anisotropy factors L_j for spheroids are given in Table 2.2.

The self-consistent field equation is now

$$(1-f)(\vec{\epsilon}_{av} - \vec{\epsilon}_m) + \sum_k f_k (\vec{\epsilon}_{av} - \vec{\epsilon}) \vec{\lambda}_k = 0 \quad (2.133)$$

which is still a tensor expression with a sum over the individual particles in the medium.

For a collection of randomly distributed spheroids, the tensor equation is converted to a scalar equation for ϵ_{av} by integrating over all the particle orientations. This gives

$$\epsilon_{av} = \frac{(1-f)\epsilon_m + f\beta\epsilon}{1-f + f\beta} \quad (2.134)$$

which is the final Maxwell Garnett equation.

The most important special cases for Maxwell Garnett are for spheres

$$\beta = \frac{3\epsilon_m}{\epsilon + 2\epsilon_m} \quad (2.135)$$

$$\epsilon_{av} = \epsilon_m \left[1 + \frac{3f \left(\frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \right)}{1 - f \left(\frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \right)} \right] = \epsilon_m \left[\frac{\epsilon(1+2f) + 2(1-f)\epsilon_m}{\epsilon(1-f) + (2+f)\epsilon_m} \right] \quad (2.136)$$

An alternative expression for the Maxwell Garnett result is

$$\frac{\epsilon_{av} - \epsilon_m}{\epsilon_{av} + 2\epsilon_m} = f \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \quad (2.137)$$

which captures the relationship of this model to the Clausius–Mossotti relation.

2.5.2.2 Bruggemann

The Maxwell Garnett formula is not symmetric with respect to the inclusions and the matrix. It specifically applies to spheroids in a homogeneous medium. However, in many situations, it is difficult to identify which material is the inclusion and which is the matrix. In such a case, a slightly different effective medium model can be used that is symmetric between the two phases. This is the Bruggeman model that applies for a two-phase mixture

$$f \frac{\varepsilon - \varepsilon_{av}}{\varepsilon + 2\varepsilon_{av}} = (f - 1) \frac{\varepsilon_m - \varepsilon_{av}}{\varepsilon_m + 2\varepsilon_{av}} \quad (2.138)$$

For dilute systems (of either phase) this gives the same result as Maxwell Garnett to lowest order. This expression is easily extended to multi-component systems by

$$\sum_j f_j \frac{\varepsilon_j - \varepsilon_{av}}{\varepsilon_j + 2\varepsilon_{av}} = 0 \quad (2.139)$$

where the f_j are the individual fractions of the different components.

2.6 Mie Scattering

Many approaches to optical biosensors use nanoparticles and gold and glass beads, which can be large and outside of the Rayleigh limit. When their size approaches an appreciable fraction of a wavelength of light, they enter into the Mie, or resonant scattering, regime. Of special interest are spherical particles and their light-scattering properties.

2.6.1 Spherical Particles

The relationship between the incident fields and the scattered fields is

$$\begin{pmatrix} E_{\parallel s} \\ E_{\perp s} \end{pmatrix} = \frac{e^{ik(r-z)}}{-ikr} \begin{pmatrix} S_2 & 0 \\ 0 & S_1 \end{pmatrix} \begin{pmatrix} E_{\parallel i} \\ E_{\perp i} \end{pmatrix} \quad (2.140)$$

where parallel \parallel and perpendicular \perp relate to the scattering plane. The scattering matrix elements are

$$S_1 = \sum_n \frac{2n+1}{n(n+1)} (a_n \pi_n + b_n \tau_n) \quad (2.141)$$

$$S_2 = \sum_n \frac{2n+1}{n(n+1)} (a_n \tau_n + b_n \pi_n) \quad (2.142)$$

and the series are terminated after a sufficient number of terms. In the forward direction $S_1(0^\circ) = S_2(0^\circ) = S(0^\circ)$.

The angle-dependent functions are

$$\pi_n = \frac{P_n^1}{\sin \theta} \quad \tau_n = \frac{dP_n^1}{d\theta} \quad (2.143)$$

where

$$P_n^1 = -\frac{dP_n}{d\theta} \quad (2.144)$$

is the derivative of the Legendre polynomials P_n .

The Mie coefficients are

$$a_n = \frac{m\psi_n(mx)\psi'_n(x) - \psi_n(x)\psi'_n(mx)}{m\psi_n(mx)\xi'_n(x) - \xi_n(x)\psi'_n(mx)} \quad (2.145)$$

$$b_n = \frac{\psi_n(mx)\psi'_n(x) - m\psi_n(x)\psi'_n(mx)}{\psi_n(mx)\xi'_n(x) - m\xi_n(x)\psi'_n(mx)} \quad (2.146)$$

where $m = n/n_m$ is the relative refractive index of the particle relative to the surrounding medium, and the functions are defined as

$$\psi_n(\rho) = \rho j_n(\rho) \quad \xi_n(\rho) = \rho h_n^{(1)}(\rho) \quad (2.147)$$

The spherical Bessel functions are

$$j_n(\rho) = \sqrt{\frac{\pi}{2\rho}} J_{n+1/2}(\rho) \quad y_n(\rho) = \sqrt{\frac{\pi}{2\rho}} Y_{n+1/2}(\rho) \quad (2.148)$$

These are combined into the spherical Hankle functions

$$h_n^{(1)}(\rho) = j_n(\rho) + iy_n(\rho) \quad (2.149)$$

The scattering cross-section is

$$C_{\text{sca}} = \frac{2\pi}{k^2} \sum_n (2n+1) (|a_n|^2 + |b_n|^2) \quad (2.150)$$

and the extinction cross-section is

$$C_{\text{ext}} = \frac{2\pi}{k^2} \sum_n (2n+1) \text{Re}(a_n + b_n) \quad (2.151)$$

For a particle that is small relative to the wavelength, the scattering matrix elements are

Rayleigh particle scattering coefficients:

$$\begin{aligned} S_1 &= -ix^3 \frac{m^2-1}{m^2+2} \\ S_2 &= -ix^3 \frac{m^2-1}{m^2+2} \cos \theta \end{aligned} \quad (2.152)$$

where $x = n_m k_0 a = k_m a$. These may also be expressed in terms of polarizability α as

$$S_1 = \frac{-ik_m^3 \alpha}{4\pi\epsilon_0} \quad (2.153)$$

using the polarizability

$$\alpha = 4\pi\epsilon_0 a^3 \frac{\epsilon_{\text{sph}} - \epsilon_m}{\epsilon_{\text{sph}} + 2\epsilon_m} \quad (2.154)$$

Note that the scattered far-field of the Rayleigh particle is purely real and positive if the particle dielectric constant is purely real. The exciting local field (assume a focused Gaussian beam) is also real, but acquires a $\pi/2$ phase shift upon free-space propagation to the far-field. Therefore, a nonabsorbing Rayleigh particle in a nonabsorbing medium induces a phase shift on the combined scattered and original wave in the far-field.

2.6.2 Effective Refractive Index of a Dilute Plane of Particles

The effective refractive index of a dilute collection of Mie scatterers detected in transmission is obtained by generalizing (2.107) to a collection of particles. This is a mean scattered field approach, and differs in its predictions from effective medium models like Maxwell Garnett. The mean scattered field approach is better suited to describe effective scattering properties of particles on surfaces. The total field in the far-field of a single particle on the optic axis is

$$E(\theta) = \frac{k\sqrt{S_0}}{f} \left[\frac{S(\theta)}{-ik^2} - iw_0^2 \exp(-k^2 w_0^2 \theta^2 / 2) \right] \quad (2.155)$$

When there is an area density of spherical scatterers, the integral over the particle density weighted by the Gaussian incident field is

$$N_A \int_0^\infty e^{-r^2/2w_0^2} 2\pi r dr = 2\pi N_A w_0^2 \quad (2.156)$$

then the total far-field becomes

$$E(\theta) = \frac{k\sqrt{S_0}}{f} \left[2\pi N_A w_0^2 \frac{S(\theta)}{-ik^2} - iw_0^2 \exp\left(-\frac{2\pi^2 w_0^2}{\lambda^2} \theta^2\right) \right] \quad (2.157)$$

The field on the optic axis ($\theta = 0^\circ$) is

$$E(0) = \frac{k\sqrt{S_0}}{f} (-iw_0^2) \left[1 - \frac{2\pi N_A}{k^2} S(0) \right] \quad (2.158)$$

If this is compared with the extinction of a wave in a homogeneous medium with refractive index n_{av}

$$E(0) = E_0 e^{ikd(n_{av}-1)} \approx E_0 (1 + ikd(n_{av} - 1)) \quad (2.159)$$

the average refractive index is

$$n_{av} = 1 + i \frac{2\pi \mathbb{N}_V}{k^3} S(0) \quad (2.160)$$

where $\mathbb{N}_V = N_A/d$ is the equivalent volume density. The real and imaginary parts of the effective refractive index $n = n' + i\kappa$ are

$$\begin{aligned} n'_{av} &= 1 - \frac{2\pi \mathbb{N}_V}{k^3} \text{Im}\{S(0)\} \\ \kappa_{av} &= \frac{2\pi \mathbb{N}_V}{k^3} \text{Re}\{S(0)\} \end{aligned} \quad (2.161)$$

In terms of the volume fraction

$$f_V = \frac{4\pi}{3} \mathbb{N}_V a^3 \quad (2.162)$$

these are

$$\begin{aligned} n'_{av} &= 1 - \frac{3f_V}{2k^3 a^3} \text{Im}\{S(0)\} \\ \kappa_{av} &= \frac{3f_V}{2k^3 a^3} \text{Re}\{S(0)\} \end{aligned} \quad (2.163)$$

In the Rayleigh limit, with

$$S(0) = -ik^3 a^3 \frac{m^2 - 1}{m^2 + 2} \quad (2.164)$$

in a medium with index n_m and relative index $m = n/n_m$, these are

$$\begin{aligned} n'_{av} &= n_m + \frac{3}{2} f_v \operatorname{Re} \left\{ \frac{m^2 - 1}{m^2 + 2} \right\} \\ \kappa_{av} &= \frac{3}{2} f_v \operatorname{Im} \left\{ \frac{m^2 - 1}{m^2 + 2} \right\} \end{aligned} \quad (2.165)$$

The effective index is linear in the volume fraction f_v in this limit. For denser dispersions, local screening modifies these values and more general approaches must be used, such as generalized Mie theory [3]. Equation (2.165) is consistent with Maxwell Garnett in the dilute limit.

2.7 Nanoparticle Light-Scattering

Nanoparticles have become essential tools for biomedical research. They are used in two main applications: (1) light scattering or emitting labels to image molecular and cellular processes and (2) drug delivery vehicles. In the labeling application, the distinct signature of the light scattering provides specific identification of the nanoparticle density and distribution within a biological sample. In the drug application, light scattering can track the transport and clearing of the drug delivery vehicles.

There are many varieties of biomedical nanoparticles. The simplest are nanobeads that are dielectric spheres that scatter light through Rayleigh or Mie scattering. These beads can be transparent, or dyed with dye molecules (chromophores) having identifiable absorption spectra. Beads also can be fluorescent emitters that are detected using fluorescence interference microscopy. Among the brightest nanoparticle emitters are semiconductor quantum dots. The emission wavelengths of quantum dots are tuned by controlling the size of the semiconductor particle using quantum confinement effects to shift the electron quantum states. Gold and silver nanoparticles scatter light resonantly (and hence strongly) through surface plasmon oscillations that are excited by the incident light fields. The optical properties of these nanoparticles are tuned by selecting a wide variety of sizes and shapes, from nanorods to nanostars.

Luminescent Quantum Dot

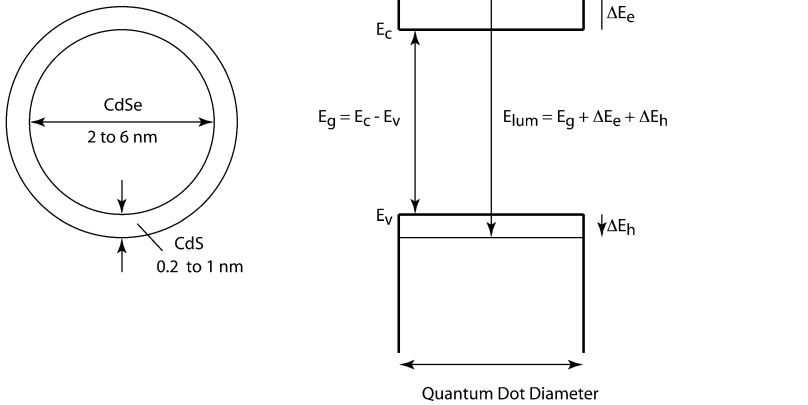


Fig. 2.15 Luminescent quantum dot structure. A CdSe nanocrystal is encased in a CdS shell to prevent nonradiative surface recombination. The photon energy of the luminescence is determined by the bandgap and by quantum confinement of the electrons in the conduction band and the holes in the valence band

2.7.1 Quantum Dots

Quantum dots are semiconductor nanocrystals typically 2–10 nm in diameter. Semiconductors have distinct and strong optical properties because of the bandgap between valence states occupied by holes and conduction states occupied by electrons. When an electron in the conduction band falls across the bandgap to fill a hole in the valence band, light is emitted with a photon energy equal to the energy that the electron loses in the process. For large nanocrystals, the energy is approximately equal to the bandgap between the top of the valence band and the bottom of the conduction band

$$\hbar\omega_g = \frac{hc}{\lambda_g} = E_g = E_c - E_v \quad (2.166)$$

One of the important properties of quantum dots is the size dependence of the emission wavelength. This is caused by quantum confinement effects on the electrons and holes. In quantum mechanics, when a particle is confined to a finite volume, there is a quantum energy associated with the confinement. The stronger the confinement, the larger the confinement energy. Therefore, as the quantum dots get smaller, the emission wavelength shifts to the blue (higher energy).

The structure of a luminescent quantum dot is illustrated in Fig. 2.15 for a CdSe nanocrystal encased in a CdS shell. The shell reduces surface recombination that

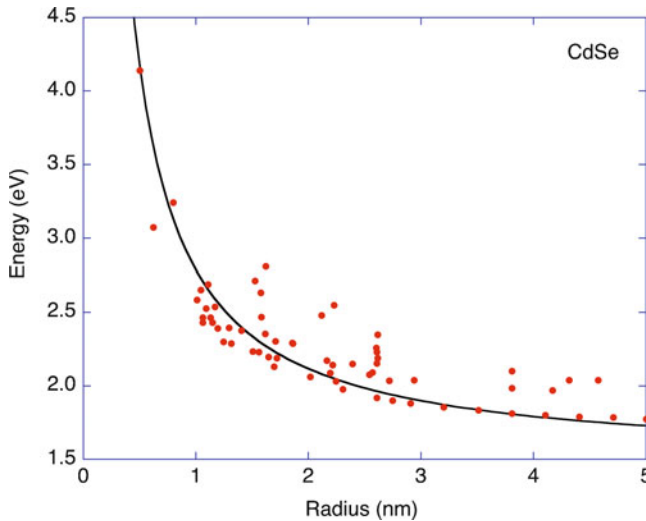


Fig. 2.16 CdSe luminescence energy from assorted experiments as a function of the quantum dot radius. Reproduced with permission from [5]

lowers the luminescence efficiency of the quantum dot. The emitted photon energy is determined by the bandgap of the nanocrystal plus the quantum confinement.

For a spherical quantum dot of radius R , the emission energy of the photon is

$$E \approx E_g + \frac{\hbar^2 \pi^2}{2R^2} \left(\frac{1}{m_e^*} + \frac{1}{m_h^*} \right) - \frac{1.786e^2}{4\pi\epsilon_0\epsilon R} - 0.248E_{\text{Ry}} \quad (2.167)$$

where the first term is the crystalline bandgap, the second term is the quantum confinement energy for the electrons and holes, and the third and fourth terms are a binding energy between the electron and hole caused by the Coulomb interaction between the particles [4]. The bound state of the electron and the hole is called an exciton, and the binding energy is the solid-state analog of the hydrogen atom. In (2.167) m_e^* is the electron effective mass, m_h^* is the hole effective mass, ϵ is the dielectric constant, and E_{Ry} is the Rydberg energy of the exciton. The emission energy of CdSe quantum dots as a function of radius is shown in Fig. 2.16 compared with the numerical result of (2.167). Examples of absorption and emission for CdSe quantum dots are shown in Fig. 2.17.

2.7.2 Gold and Silver Nanoparticles

Gold and silver nanoparticles are among the most commonly used particles in labeling applications of light scattering and interferometry. These particles show

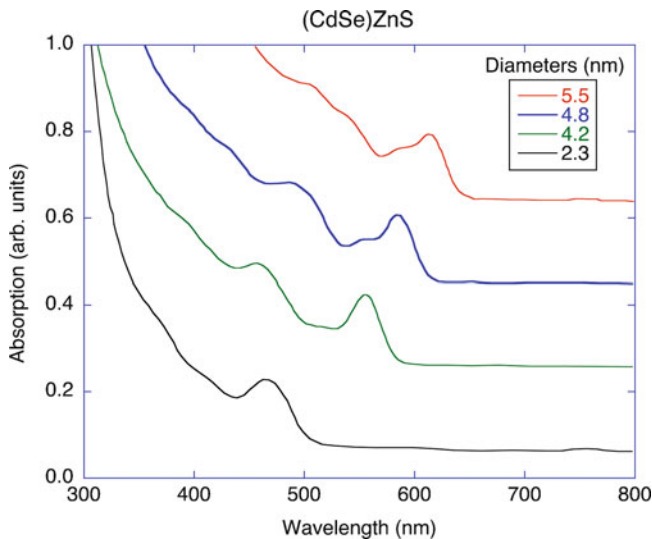


Fig. 2.17 Absorbance of CdSe quantum dots clad with ZnS as a function of wavelength for increasing diameters of (a) 2.30, (b) 4.20, (c) 4.80, and (d) 5.50 nm. (Reproduced with permission from [6])

strong plasmonic resonances with plasmonically enhanced absorption and light scattering. In the Rayleigh limit, the particles contribute to enhanced effective-medium properties, while in the large Mie-scattering limit, they provide single-particle detectability and finite-particle statistics in applications such as imaging and laser scanning.

The plasmonic enhancements of gold and silver arise from the nearly free-electron behavior of these noble metals. The dielectric functions for a free-electron gas, including the bulk dielectric function $\epsilon_b = 6 + i1.6$, is

$$\begin{aligned}\epsilon_1 &= \text{Re}(\epsilon_b) - \frac{\omega_p^2}{\omega^2 + \gamma^2} \\ \epsilon_2 &= \text{Im}(\epsilon_b) + i\omega_p^2 \frac{\gamma/\omega}{\omega^2 + \gamma^2}\end{aligned}\quad (2.168)$$

For gold, the plasmon energy is approximately $\hbar\omega_p = 7.5$ eV, and the damping rate is approximately $\hbar\gamma = 0.25$ eV. The free-electron functions for gold are shown in Fig. 2.18. The real part of the dielectric function has a zero near 400 nm. This is not near $\omega_p = 7.5$ eV at a wavelength of 165 nm because of the background dielectric function arising from interband transitions among the gold orbitals. There is also a significant imaginary component caused by damping (scattering) of the free electrons and from the interband transitions.

Free-electron models capture the rough behavior of real noble metals, but fail to capture the contributions from interband absorption and the associated effects on

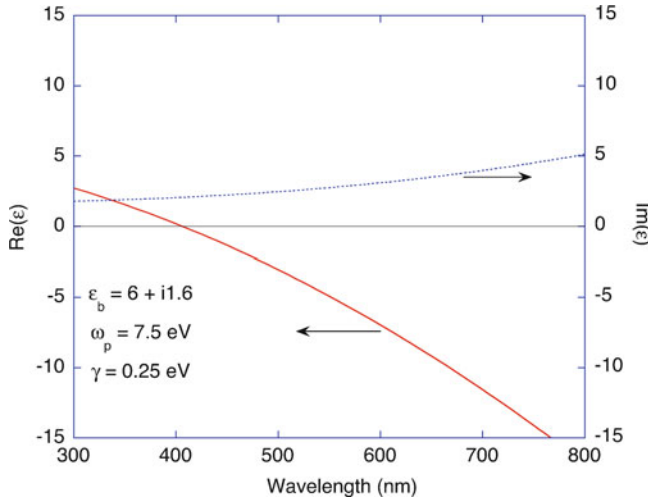


Fig. 2.18 Real and imaginary components of the dielectric functions of gold assuming a free-electron model

the refractive index. A more accurate parameterization for gold has been provided in [7], including the role of the finite size of gold particles. As the particle size decreases, the electron scattering by the surface increases, and the overall damping of the plasmon resonance increases. This is parameterized as

$$\gamma_p(d) = \frac{1}{1/\gamma_p(\infty) + 1/Ad} \quad (2.169)$$

where A is a scattering parameter and d is the particle diameter in nanometers. The total dielectric function can be expressed as

$$\begin{aligned} \varepsilon(\lambda) = & \varepsilon_\infty - \frac{1}{\lambda_p^2(1/\lambda^2 + i/\gamma_p\lambda)} \\ & + \sum_{i=1,2} \frac{A_i}{\lambda_i} \left[\frac{e^{i\phi_i}}{1/\lambda_i - 1/\lambda - i/\gamma_i} + \frac{e^{-i\phi_i}}{1/\lambda_i + 1/\lambda + i/\gamma_i} \right] \end{aligned} \quad (2.170)$$

The first term is the plasmon contribution, while the second term is the contribution from interband absorption. The parameters that best fit the data in [8] are given in Table 2.3.

The real and the imaginary parts of the dielectric functions are plotted in Fig. 2.19. Strong finite-size effects on the imaginary part occur for particle radii smaller than 40 nm. However, the real part is only significantly affected for radii smaller than 10 nm.

Table 2.3 Parameters in the parameterization of the gold dielectric function by Etchegoin [7]

Data parameters	Value
ϵ_∞	1.53
λ_p	145 nm
$\gamma_p(\infty)$	17,000 nm
A	170
A_1	0.94
ϕ_1	$-\pi/4$
λ_1	468 nm
γ_1	2,300 nm
A_2	1.36
ϕ_2	$-\pi/4$
λ_2	331 nm
γ_2	940 nm

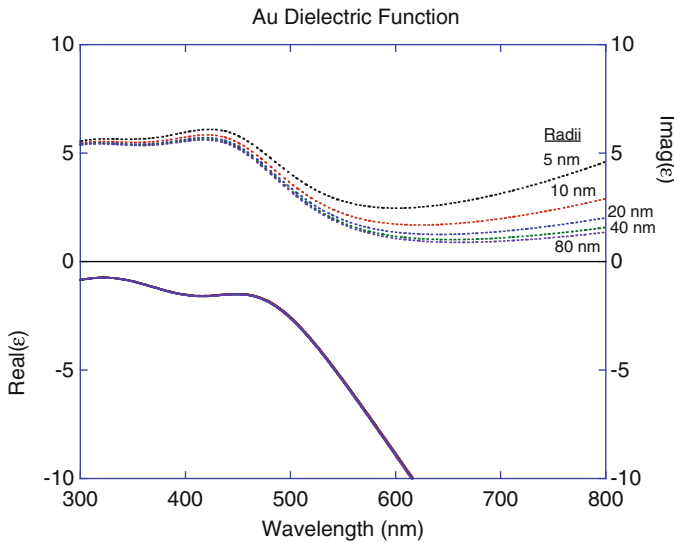


Fig. 2.19 Dielectric function of gold parameterized [7] and fit to the data [8]. The surface plasmon resonance of a spherical nanoparticle in vacuum occurs when $\text{Re}(\epsilon) = -2$

The dipole moment of a sphere was given in (2.93). When the sphere is embedded in a medium with dielectric constant ϵ_m , it is

$$\mathbf{p} = 4\pi\epsilon_m\epsilon_0 \left(\frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \right) a^3 \mathbf{E}_{\text{inc}} \quad (2.171)$$

Therefore, when $\epsilon_{\text{gold}} = -2\epsilon_m$ a spherical nanoparticle has a plasma resonance that occurs in both the absorption and the scattering cross-sections. The absorption and scattering efficiencies of gold nanoparticles are calculated numerically using

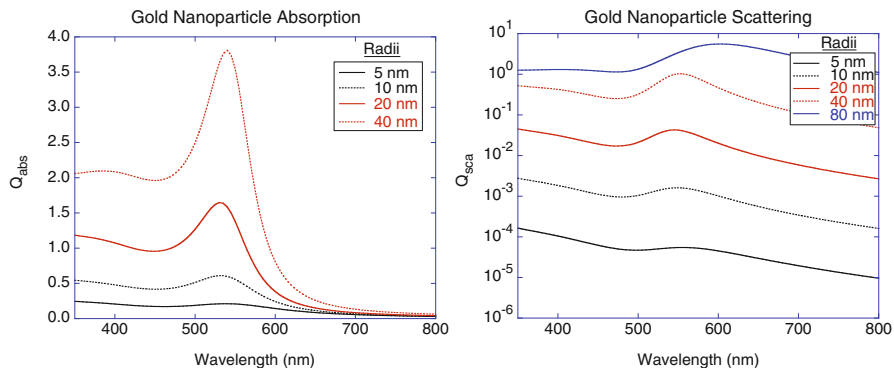


Fig. 2.20 Numerical simulations of the absorption and scattering efficiencies of gold nanoparticles calculated using Mie theory and the gold dielectric function of (2.170)

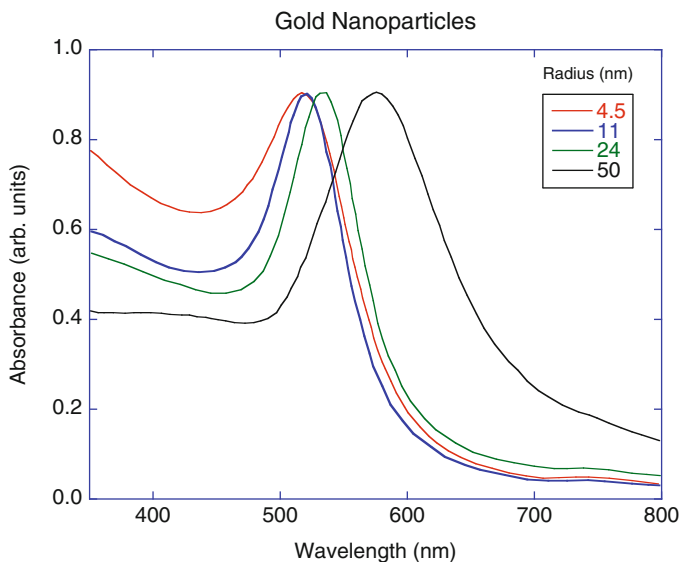


Fig. 2.21 Experimental absorbance of gold nanoparticles in solution reprinted with permission from [9] as a function of particle radius

Mie theory and (2.170) and shown in Fig. 2.20. A particle with a radius of 40 nm shows a strong absorption resonance with enhanced scattering. Smaller particles have broader resonances because of increased surface scattering of the free electrons, and smaller efficiencies because they move into the Rayleigh limit. The particles begin to move out of the Rayleigh limit and into the Mie regime as the particle radii become larger than about 50 nm. Experimental normalized absorbance of gold nanoparticles in solution is shown in Fig. 2.21 from [9] for diameters from 9 to 99 nm. The resonance shifts to longer wavelengths with increasing size.

Selected Bibliography

- van de Hulst, H.C.: *Light Scattering by Small Particles*. Dover, New York (1981) (The old classic on light scattering)
- Bohren, C.F., Huffman, D.R.: *Absorption and Scattering of Light by Small Particles*. Wiley, New York (1998) (This is more up-to-date, and includes computer codes for scattering)
- Berne, B.J., Pecora, R.: *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics*. Dover, New York (2000) (This is a treasure trove of dynamic light scattering, but with many older references)

References

1. Zhao, M., Cho, W., Regnier, F., Nolte, D.: Differential phase-contrast BioCD biosensor. *Appl. Opt.* **46**, 6196–6209 (2007)
2. Felidj, N., Aubard, J., Levi, G.: Discrete dipole approximation for ultraviolet-visible extinction spectra simulation of silver and gold colloids. *J. Chem. Phys.* **111**(3), 1195–1208 (1999)
3. Gouesbet, G., Grehan, G.: Generalized Lorenz–Mie theory for assemblies of spheres and aggregates. *J. Opt. A: Pure Appl. Opt.* **1**(6), 706–712 (1999)
4. Brus, L.: Electronic wave-functions in semiconductor clusters – experiment and theory. *J. Phys. Chem.* **90**(12), 2555–2560 (1986)
5. Tomasulo, A., Ramakrishna, M.V.: Quantum confinement effects in semiconductor clusters II. *J. Chem. Phys.* **105**(9), 3612–3626 (1996)
6. Dabbousi, B.O., RodriguezViejo, J., Mikulec, F.V., Heine, J.R., Mattoussi, H., Ober, R., Jensen, K.F., Bawendi, M.G.: (CdSe)ZnS core-shell quantum dots: synthesis and characterization of a size series of highly luminescent nanocrystallites. *J. Phys. Chem. B* **101**(46), 9463–9475 (1997)
7. Etchegoin, P.G., Le Ru, E.C., Meyer, M.: An analytic model for the optical properties of gold, vol. 125, p. 164705, 2006
8. Johnson, P.B., Christy, R.W.: Optical-constants of noble-metals. *Phys. Rev. B* **6**(12), 4370–4379 (1972)
9. Link, S., El-Sayed, M.A.: Spectral properties and relaxation dynamics of surface plasmon electronic oscillations in gold and silver nanodots and nanorods. *J. Phys. Chem. B* **103**(40), 8410–8426 (1999)

<http://www.springer.com/978-1-4614-0889-5>

Optical Interferometry for Biology and Medicine

Nolte, D.D.

2012, XII, 354 p., Hardcover

ISBN: 978-1-4614-0889-5