Chapter 8
Fundamentals of Thermal Radiation

Radiation is one of the fundamental modes of heat transfer. However, the concepts of thermal radiation are much more complicated and, hence, very difficult to perceive. The main features of radiation that are distinct from conduction and convection are as follows: (a) radiation can transfer energy with and without an intervening medium; (b) the radiant heat flux is not proportional to the temperature gradient; (c) radiation emission is wavelength dependent, and the radiative properties of materials depend on the wavelength and the temperature; (d) the radiant energy exchange and the radiative properties depend on the direction and orientation [1, 2].

The dual theory explains the nature of radiation as either electromagnetic waves or a collection of particles, called photons. Although radiation can travel in vacuum, it originates from matter. All forms of matter emit radiation through complicated mechanisms (e.g., molecular vibration in gases, and electron and lattice vibrations in solids). In most solids and some liquids, radiation emitted from the interior is strongly absorbed by adjoining molecules. Therefore, radiation from or to these materials is often treated as *surface phenomena,* while radiation in gases and some semitransparent solids or liquids has to be treated as *volumetric phenomena.* Nevertheless, one must treat solids or liquids volumetrically as a medium to understand the mechanisms of reflection and emission, to predict the radiative properties of thin films and small particles, and to calculate radiation heat transfer between objects placed in close vicinity. *Thermal radiation* refers to a type of radiation where the emission is directly related to the temperature of the body (or surface).

There are numerous engineering applications where radiation heat transfer is important, such as solar energy, combustion, furnaces, high-temperature materials processing and manufacturing, and insulation in space and cryogenic systems. Even at room temperature, radiative heat transfer may be of the same order of magnitude as convective heat transfer. The study of thermal radiation went along with the study of light phenomena and led to some major breakthroughs in modern physics. It is instructive to give a brief survey of major historical developments related to thermal radiation.

© Springer Nature Switzerland AG 2020

Z. M. Zhang, *Nano/Microscale Heat Transfer*, Mechanical Engineering Series,
https://doi.org/10.1007/978-3-030-45039-7_8

Quantitative understanding of light phenomena began in the seventeenth century with the discoveries of Snell's law of refraction, Fermat's least-time principle of light path, Huygens' principle of constructing the wavefront from secondary waves, and Newton's prism that helped him prove white light consists of many different types of rays. In the dawn of the nineteenth century, Sir Frederick Herschel (1738–1822), a German-born English astronomer, discovered infrared radiation [3]. His original objective was to find a suitable color for a glass filter, which could transmit the most of light but the least amount of heat, for use in solar observations. By moving a thermometer along the spectrum of solar radiation that passed through a prism, Herschel accidently found that the temperature of the thermometer would rise even though it was placed beyond the red end of the visible light. He published several papers in *Philosophical Transactions of the Royal Society of London* in 1800 and called the unknown radiation *invisible light* or *heat-making rays.* Young's double-slit experiment in 1801 demonstrated the interference phenomenon and the wave nature of light. It was followed by extensive studies on polarization and reflection led by French physicist Augustin-Jean Fresnel (1788–1827) who contributed significantly to the establishment of the wave theory of light. In 1803, radiation beyond the violet end of the visible spectrum via chemical effects was also discovered. The ultraviolet, visible, and infrared spectra were thus associated with chemical, luminous, and heating effects, respectively. Yet, the common nature of the different types of radiation was not known until the late nineteenth century.

One of the obstacles of accurately measuring infrared radiation (or heat radiation, as it was called in those days) was the lack of sensitive detectors. In the earlier years, measurements were performed using mercury-in-glass thermometers with blackened bulbs. In 1829, Italian physicists Leopoldo Nobili (1784–1835) and Macedonio Melloni (1798–1854) invented the thermopile made by connecting a number of thermocouples in series that is much more sensitive and faster than the thermometer. Melloni used the device to study the infrared radiation from hot objects and the sun. Gustav Kirchhoff (1824–1887), a German physicist, contributed greatly to the fundamental understanding of spectroscopy and thermal emission by heated objects. In 1862, he coined the term "black body" radiation and established Kirchhoff's law, which states that the emissivity of a surface equals its absorptivity at thermal equilibrium.

Many famous physicists and mathematicians have contributed to electromagnetism. The complete equations of electromagnetic waves were established in 1873 by Scottish physicist James Clerk Maxwell (1831–1879), and later confirmed experimentally by German physicist Heinrich Hertz (1857–1894) through the discovery of radio waves due to electrical vibrations. Before the existence of electrons was proved, Dutch physicist Hendrik Lorentz (1853–1928) proposed that light waves were due to oscillations of electric charges in the atom. His electron theory could explain the phenomenon discovered by his mentee Pieter Zeeman (1865–1943) that the lines in the spectrum can split into several lines under a strong magnetic field (known as the Zeeman effect). They shared the Nobel Prize in Physics in 1902 for their research into the influence of magnetism upon radiation phenomena. The electromagnetic wave theory has played a central role in radio, radar, television, microwave technology, telecommunication, thermal radiation, and physical optics. Albert Einstein arrived

at the famous formula $E = mc^2$ in 1905, after connecting the relativity principle with the Maxwell equations.

In 1881, Samuel Langley (1834–1906), the American astronomer, physicist, and aeronautics pioneer, invented a highly sensitive device called a *bolometer* for detecting thermal radiation. The bolometer used two platinum strips, connected in a Wheatstone bridge circuit with a sensitive galvanometer, to read the imbalance of the bridge caused by the exposure of one of the strips to radiation. Langley was the first to make an accurate map of the solar spectrum up to a wavelength of 2.8 μm. The Stefan–Boltzmann law of blackbody radiation is buit upon the empirical relation obtained by Slovenian physicist Joseph Stefan (1835–1893) in 1879, through careful experimental observation. The theoretical proof was provided by Austrian physicist Ludwig Boltzmann (1844–1906) in 1884, based on the thermodynamic relations of a Carnot cycle with radiation as a working fluid using the concept of radiation pressure. In the late nineteenth century, German physicist Wilhelm Wien (1864–1928) derived the displacement law in 1893 by considering a piston moving within a mirrored empty cylinder filled with thermal radiation. Wien also derived a spectral distribution of blackbody radiation, called Wien's formula, which is applicable to the short-wavelength region of the blackbody spectrum but deviates from experiments toward long wavelengths. Wien received the Nobel Prize in 1911 "for his discoveries regarding the laws governing the radiation of heat." In 1900, Lord Rayleigh (1842–1919), British physicist and Nobel Laureate in Physics in 1904, used the equipartition theorem to show that the blackbody emission should be directly proportional to temperature but inversely proportional to the fourth power of wavelength. Sir James Jeans (1877–1946), a British physicist, astronomer, and mathematician, derived a more complete expression in 1905. The Rayleigh–Jeans formula agreed with experiments at sufficiently high temperatures and long wavelengths, where Wien's formula failed, but disagreed with experiments at short wavelengths. It is noteworthy that Rayleigh has made great contributions to light scattering and wave phenomena, such as the discovery of Rayleigh scattering by small objects that explains why the sky is blue and the sunset glows red and orange. Rayleigh also predicted the existence of *surface waves,* sometimes called *Rayleigh waves,* which propagate along the interface between two different media. The amplitude of the wave, however, diminishes in each media as the distance from the interface increases.

In an effort to obtain a better agreement with measurements at long wavelengths, German physicist Max Planck (1858–1947) in 1900 used the maximum entropy principle, based on Boltzmann's entropy expression, to derive an equation, known as Planck's law, which agrees with experiments in the whole spectral region. Planck obtained his expression independently of Rayleigh's work, while the complete derivation of Rayleigh–Jeans formula was obtained several years later. In his book *The Theory of Heat Radiation,* Planck [4] showed that his formula would reduce to Wien's formula at small λT and Rayleigh–Jeans formula at large λT. In his derivation, Planck used a bold assumption that is controversial to classical electrodynamics. His hypothesis was that the energy of linear oscillators is not infinitely divisible but must assume discrete values that are multiples of $h\nu$, where h is a universal constant and ν is the frequency of the oscillator. This concept would have been easily accepted

for a system consisting of particles, like atoms or gas molecules, but not for oscillators that radiate electromagnetic energy. Planck's work opened the door to quantum mechanics. The idea of quantization of radiation was further developed by Einstein, who applied it to explain the photoelectric effect in 1905. Planck was awarded the Nobel Prize in Physics in 1918 for the discovery of energy quanta. In 1924, Indian mathematical physicist Satyendra Nath Bose (1894–1974) modified the Boltzmann statistics of ideal molecular gases, by introducing the concept of different quantum states at each energy level (degeneracy) using the phase space while treating the light quanta as indistinguishable. Subsequently, Bose was able to statistically derive Planck's distribution function without using the semi-classical oscillator concept. With the help of Einstein, Bose's work was published in *Zeitschrift für Physik* in 1924. Einstein further extended Bose's theory to atoms and predicted the existence of a phenomenon, known as Bose–Einstein condensate, as discussed in Chap. 3. It is clear that the journey of questing for the truth in understanding thermal radiation has led to important discoveries in modern physics.

This chapter contains an introduction to the electromagnetic wave theory, blackbody radiation, plane wave reflection, and refraction at the boundary between two semi-infinite media, evanescent waves and total internal reflection, and various models used to study the optical properties of different materials. A brief description of the typical experimental methods used to measure the spectral radiative properties is also presented. The materials covered in the following sections are intended to provide a sound background for more in-depth studies on the applications of thermal radiation to micro/nanosystems in subsequent chapters.

8.1 Electromagnetic Waves

In this section, we will study macroscopic Maxwell's equations and electromagnetic (EM) waves in isotropic media from dielectric to dissipative (lossy) to magnetic media. The concepts of polarization, absorption, and evanescent waves are introduced using complex wavevectors. Poynting's theorem describes the energy balance for EM waves including transfer, storage, and dissipation. The complex dielectric function is defined from which the complex refractive index for nonmagnetic materials can be calculated.

8.1.1 Maxwell's Equations

The propagation of electromagnetic waves in any media is governed by a set of equations, first stated together by Maxwell. The macroscopic Maxwell equations can be written in the differential forms as follows [5–9]:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{8.1}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{8.2}$$

$$\nabla \cdot \mathbf{D} = \rho_{\mathrm{e}} \tag{8.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{8.4}$$

Here, **E** is the electric field, **H** is the magnetic field, **J** is the electric current density (or charge flux according to the definition in Chap. 4), **D** is the electric displacement, **B** is the magnetic flux density (also called magnetic induction), and ρ_{e} is the charge density. In the SI units, **E** is in V/m, **H** in C/m s, **J** in A/m^2, **D** in C/m^2, **B** in Wb/m^2, and ρ_{e} in C/m^3. Note that 1 T (T) = 1 Wb/m^2 and 1 Wb (Wb) = 1 V s. The charge conservation or continuity equation, $\nabla \cdot \mathbf{J} + \partial \rho_{\mathrm{e}} / \partial t = 0$, is implicitly included in the Maxwell equations, because it can be obtained by taking the divergence of Eq. (8.2) and then applying Eq. (8.3). The constitutive relations for a linear isotropic medium are

$$\mathbf{D} = \varepsilon_{\mathrm{m}} \mathbf{E} \tag{8.5}$$

$$\mathbf{B} = \mu_{\mathrm{m}} \mathbf{H} \tag{8.6}$$

where ε_{m} is the electric permittivity in F/m and μ_{m} the magnetic permeability of the medium in N/A^2. Note that the farad (F) is the SI unit of capacitance: 1 F = 1 C/V. The permittivity and permeability values of free space (vacuum) are $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m and $\mu_0 = 4\pi \times 10^{-7}$ N/A^2, respectively. For anisotropic media, μ_{m} and ε_{m} are dyadic tensors. The microscopic form of Ohm's law gives

$$\mathbf{J} = \sigma \mathbf{E} \tag{8.7}$$

where σ is the electric conductivity in A/V m.

A brief discussion on the physical interpretation of Maxwell's equations is given next. Equation (8.1) is an expression of Faraday's law of induction, which states that a time-varying magnetic field produces an electric field in a coil. In other words, through any closed electric field line, there is a time-varying magnetic field. Combining Eq. (8.1) with Green's theorem, Eq. (B.71), we see that the integral of the electric field around a closed loop is equal to the negative of the integral of the time derivative of the magnetic induction, over the area enclosed by the loop. Equation (8.2) is the general Ampere's law, which includes Maxwell's displacement current ($\partial \mathbf{D} / \partial t$). It states that through any closed magnetic field line, there is an electric current density **J** or a displacement current or both. Conversely, circulating magnetic fields are produced by passing an electrical current through a conductor or changing electric fields or both. Equation (8.3) is Gauss's law, which implies that the electric field

diverges from electric charges. Using Gauss's theorem, Eq. (B.70), it can be seen from Eq. (8.3) that the integral of the electric field over a closed surface is proportional to the electric charges enclosed by that surface. If there are no electric charges inside a closed surface, there is no net electric field penetrating the surface. Equation (8.4) is an analogy to Gauss's law for the magnetic field. However, isolated magnetic poles (i.e., magnetic monopoles) have not been observed, so the integration of a magnetic field over any closed surface is zero.

The interpretations given in the preceding paragraph are straightforward since all variables and coefficients are considered as real quantities. However, Maxwell's equations are most useful when all quantities are expressed in complex variables. The material properties, such as ε_m and μ_m, are generally complex and frequency dependent. To facilitate understanding, we will start with simple cases first and then generalize the theory for more realistic problems.

8.1.2 The Wave Equation

Sometimes called free charge density, ρ_e in Eq. (8.3) should be treated as excess charges or net charges per unit volume. Because the number of electrons equals the number of protons in the nuclei, in most media, we can assume $\rho_e = 0$. For a nonconductive material, $\sigma = 0$. We further assume that ε_m and μ_m are both real and independent of position, time, and the field strength. This is true for a nondissipative (lossless), homogeneous, and linear material. If $\mu_m = \mu_0$, the material is said to be nonmagnetic. Therefore, a nonconductive and nonmagnetic material is a dielectric for which only ε_m is needed to characterize its electromagnetic behavior. Materials with both ε_m and μ_m being real but $\mu_m \neq \mu_0$ are sometimes called general dielectrics or dielectric-magnetic media. Substituting the constitutive relations into Maxwell's equations and then combining Eqs. (8.1) and (8.2), we obtain

$$\nabla^2 \mathbf{E} = \mu_m \varepsilon_m \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{8.8}$$

where the vector identity given in Eq. (B.64), $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E}$, has been employed. Equation (8.8) is the *wave equation,* which can also be written in terms of the magnetic field. The wave equation has infinite number of solutions (see Problem 8.1). The solution of Eq. (8.8) for a monochromatic plane wave can be written as

$$\mathbf{E} = \mathbf{E}_0 e^{-i(\omega t - \mathbf{k}\cdot\mathbf{r})} \tag{8.9}$$

where $\mathbf{E}_0$ is the amplitude vector, ω is the angular frequency, $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ is the position vector, and $\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}}$ is the *wavevector,* which points toward the direction of propagation. In order for Eq. (8.9) to be a solution of Eq. (8.8), the magnitude of $\mathbf{k}$ must be $k = \omega\sqrt{\mu_m \varepsilon_m}$. The complex form of the electric field is

used in Eq. (8.9) to facilitate mathematical manipulation. The actual electric field may be expressed as the real part of Eq. (8.9), viz.,

$$\mathrm{Re}(\mathbf{E}) = \mathrm{Re}(\mathbf{E}_0)\cos\phi + \mathrm{Im}(\mathbf{E}_0)\sin\phi \tag{8.10}$$

where Re or Im stands for taking the real part or the imaginary part, and $\phi = \omega t - \mathbf{k}\cdot\mathbf{r}$ is the phase. Equation (8.9) is a time-harmonic solution at a fixed frequency. Because any time-space-dependent function can be expressed as a Fourier series of many frequency components, we can integrate Eq. (8.9) over all frequencies to obtain the total electric field at any time and position. Therefore, understanding the nature of Eq. (8.9) is very important to the study of electromagnetic wave phenomena.

When Eq. (8.9) is substituted into Maxwell's equations, a time derivative $\partial/\partial t$ can be replaced by a multiplication of $-\mathrm{i}\omega$ and the operator ∇ can be replaced by $\mathrm{i}\mathbf{k}$. Hence, the first two Maxwell equations can be written as

$$\mathbf{k}\times\mathbf{E} = \omega\mu_\mathrm{m}\mathbf{H} \tag{8.11a}$$

and

$$\mathbf{k}\times\mathbf{H} = -\omega\varepsilon_\mathrm{m}\mathbf{E} \tag{8.11b}$$

The two equations suggest that **E**, **H**, and **k** are orthogonal and form a right-handed triplet, when both ε_m and μ_m are positive. On the surface normal to the wavevector **k**, the electric or magnetic field is a function of time only, because $\mathbf{k}\cdot\mathbf{r} = \text{const}$. This surface is called a *wavefront*. In the **k**-direction, the wavefront travels at the speed given by

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_\mathrm{m}\varepsilon_\mathrm{m}}} \tag{8.12}$$

which is called *phase speed* and it is the smallest speed at which the wavefront propagates. The phase velocity is the phase speed times the unit wavevector [8].

Figure 8.1 illustrates a plane wave, propagating in the positive x-direction, whose electric field is parallel to the y-direction and magnetic field parallel to the z-direction.

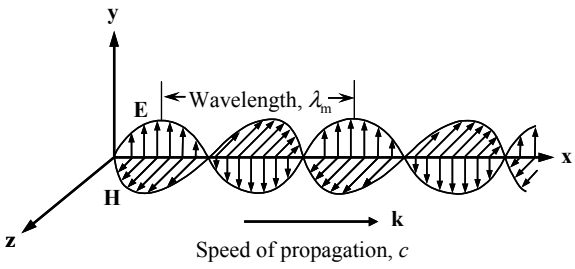

Fig. 8.1 Illustration of a linearly polarized electromagnetic wave

In such cases, $k = k_x$ and $\mathbf{k} \cdot \mathbf{r} = kx$. The wavefront is perpendicular to the x-direction. It can be seen clearly that the wavevector is related to the wavelength λ_m in the medium by $k = 2\pi/\lambda_m$.

In free space, the speed of an electromagnetic wave is given by $c_0 = 1/\sqrt{\mu_0 \varepsilon_0}$. The speed of light in a vacuum was instated as an exact number, $c_0 = 299, 792, 458$ m/s, by the General Conference on Weights and Measures (abbreviated as CGPM for Conférence Générale des Poids et Mesures) in 1983. The SI base unit meter has since been defined as the distance that light travels in a vacuum during a time interval of $1/299, 792, 458$ s. The reference on constant, units, and uncertainty can be found from the web page of the National Institute of Standards and Technology (NIST) for detailed discussions about the fundamental physical constants and the base SI units [10]. For most calculations, it suffices to use $c_0 = 2.998 \times 10^8$ m/s. The refractive index of a medium is given as $n = \sqrt{\frac{\mu_m \varepsilon_m}{\mu_0 \varepsilon_0}} = \frac{c_0}{c}$. Therefore, $c = c_0/n$ and $\lambda_m = \lambda/n$, where λ is the wavelength in vacuum. For nonmagnetic materials $\mu_m/\mu_0 = 1$; thus, $n = \sqrt{\varepsilon_m/\varepsilon_0}$.

Notice that n of a medium is a function of frequency (or wavelength) and is, in general, temperature dependent. For polychromatic light, the phase speed usually depends on wavelength because $n = n(\lambda)$ in a dispersive medium. In a vacuum, the energy propagation velocity is the same as the phase velocity. For polychromatic waves in a dispersive medium, the group velocity $\mathbf{v}_g$ determines the direction and speed of energy flow and is defined as

$$\mathbf{v}_g = \nabla_k \omega = \frac{\mathrm{d}\omega}{\mathrm{d}\mathbf{k}} = \frac{\partial \omega}{\partial k_x}\hat{\mathbf{x}} + \frac{\partial \omega}{\partial k_y}\hat{\mathbf{y}} + \frac{\partial \omega}{\partial k_z}\hat{\mathbf{z}} \tag{8.13}$$

which is the gradient of ω in the k-space. In a homogeneous and isotropic medium, $v_g = c_0\left(n + \omega \frac{dn}{d\omega}\right)^{-1}$ and the direction of the group velocity will be the same as that of the wavevector $\mathbf{k}$. In a nondispersive medium, where n is not a function of frequency, it is clear that $v_g = c = c_0/n$. A group front can also be defined based on the constant-amplitude surface of the wave group. In general, it is not parallel to the wavefront, when light is refracted from a nondispersive medium to a dispersive medium; furthermore, the energy flow direction is not necessarily perpendicular to the group front [11]. Notice that the wave equation is also applicable to other types of waves such as acoustic waves, which are matter waves with a longitudinal and two transverse modes, as mentioned in Chap. 5.

8.1.3 Polarization

A simple transverse wave will oscillate perpendicular to the wavevector. Because electromagnetic waves have two field vectors that can change their directions during propagation, the polarization behavior may be complicated. It is important to

understand the nature of polarization in order to fully characterize an electromagnetic wave. There are two equivalent ways to interpret a complex vector $\mathbf{A}$. The first method considers it as a vector whose components are complex, i.e.,

$$\mathbf{A} = A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}} \tag{8.14a}$$

where A_x, A_y, and A_z are complex numbers:

$$A_x = A_x' + \mathrm{i}A_x'',\ A_y = A_y' + \mathrm{i}A_y'',\ \text{and}\ A_z = A_z' + \mathrm{i}A_z'' \tag{8.14b}$$

The second method decomposes it into two real vectors such that

$$\mathbf{A} = \mathbf{A}' + \mathrm{i}\mathbf{A}'' \tag{8.15a}$$

where $\mathbf{A}'$ and $\mathbf{A}''$ are the real and imaginary parts of the complex vector, given by

$$\mathbf{A}' = A_x'\hat{\mathbf{x}} + A_y'\hat{\mathbf{y}} + A_z'\hat{\mathbf{z}}\ \text{and}\ \mathbf{A}'' = A_x''\hat{\mathbf{x}} + A_y''\hat{\mathbf{y}} + A_z''\hat{\mathbf{z}} \tag{8.15b}$$

In either case, a complex vector has six real scalar terms.

For the time being, let us assume all the material properties to have real values and $\mathbf{k}$ to be a real vector. Both $\mathbf{E}$ and $\mathbf{H}$ are complex, according to Eq. (8.9). To ensure that $\mathbf{k}\cdot\mathbf{E} = 0$ at any time and location, both $\mathrm{Re}(\mathbf{E}_0)$ and $\mathrm{Im}(\mathbf{E}_0)$ must be perpendicular to $\mathbf{k}$. The same is true for the magnetic vector. Because $\mathbf{H}$ can be obtained from Eq. (8.11a), the state of polarization can be based on how the electric field varies in time and along the $\mathbf{k}$-direction in space. In order to study the time dependence of the electric field, rewrite Eq. (8.10) as

$$\mathrm{Re}(\mathbf{E}) = \mathbf{a}\cos(\omega t) + \mathbf{b}\sin(\omega t) \tag{8.16}$$

where $\mathbf{a} = \mathrm{Re}(\mathbf{E}_0 e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}})$ and $\mathbf{b} = \mathrm{Im}(\mathbf{E}_0 e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}})$ are both real vectors and perpendicular to $\mathbf{k}$. In general, the electric field will vary with time in an ellipse, called the *vibration ellipse*, as shown in Fig. 8.2. If $\mathbf{a}$ and $\mathbf{b}$ are parallel or, equivalently, $\mathrm{Re}(\mathbf{E}_0)$ and

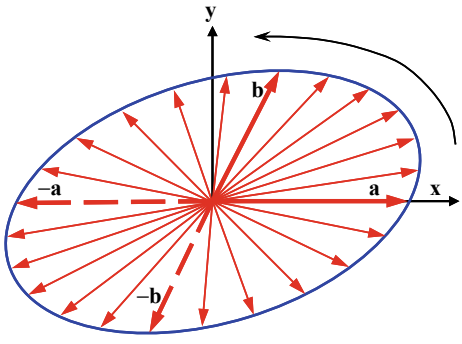

Fig. 8.2 Illustration of polarization by the vibration ellipse, for a plane wave propagating in the positive z-direction (out of the paper). The electric field vector is plotted at an increment of $\omega\Delta t = \pi/12$

$\mathrm{Im}(\mathbf{E}_0)$ are parallel to each other, then the electric field will not change its directions. The wave is said to be *linearly polarized*, and either **a** or **b** specifies the direction of polarization. An example of a linearly polarized wave is the wave shown in Fig. 8.1. When $\mathbf{a} \perp \mathbf{b}$ and $|\mathbf{a}| = |\mathbf{b}|$, the vibration ellipse is a circle and the wave is said to be *circularly polarized*. In general, a monochromatic wave described by Eq. (8.10) is *elliptically polarized.* For circularly or elliptically polarized light, if $\mathbf{a} \times \mathbf{b}$ is in the same direction as **k**, the vibration ellipse will rotate counterclockwise (left handed) when viewed toward the light source, and if $\mathbf{a} \times \mathbf{b}$ is opposite to the direction of propagation, the vibration ellipse will rotate clockwise (right handed). Similarly, one can consider the polarization of the electric field at a fixed time and observe the vibration ellipse along the direction of propagation as an exercise (see Problem 8.2).

Because of the random nature of thermal radiation, the Fourier component does not vary with time exactly following $e^{-i\omega t}$ but with some fluctuations in the amplitude. The polarization may become completely random, which is said to be *unpolarized*, *randomly polarized*, or *completely uncorrelated.* In any case, the electric field can be decomposed into the two orthogonal directions on the vibration ellipse. This is particularly useful for calculating energy transfer. The polarization status can be fully described by the four Stokes parameters and the response of an optical element can be modeled using the Mueller matrix formulation [6, 7]. For coherent monochromatic light, however, the Jones vector and Jones matrix formulation can provide additional phase information since it is based on the transformation of the electric field rather than the amplitude [12].

8.1.4 Energy Flux and Density

The energy conservation for an electromagnetic field can be obtained from Maxwell's equations, according to English physicist John Poynting (1852–1914). To derive Poynting's theorem, one can dot multiply Eqs. (8.1) and (8.2) by –**H** and **E**, respectively, and then add up each side. Using the vector identity in Eq. (B.63), we get $\nabla \cdot (\mathbf{E} \times \mathbf{H}) = (\nabla \times \mathbf{E}) \cdot \mathbf{H} - (\nabla \times \mathbf{H}) \cdot \mathbf{E}$. After some simplifications, we obtain

$$-\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \frac{\partial}{\partial t}\left(\frac{1}{2}\varepsilon_{\mathrm{m}}\mathbf{E} \cdot \mathbf{E} + \frac{1}{2}\mu_{\mathrm{m}}\mathbf{H} \cdot \mathbf{H}\right) + \mathbf{E} \cdot \mathbf{J} \tag{8.17}$$

The left-hand term represents the energy flow into a differential control volume, the first term on the right is the rate of change of the stored energy (associated with the electric and magnetic fields), and the last term is the dissipated electromagnetic work or Joule heating. The *Poynting vector* is defined as

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \tag{8.18a}$$

The Poynting vector is essentially the energy flux, which gives both the direction and the rate of energy flow per unit projected surface area. Equations (8.17) and (8.18a)

can be easily extended to the complex field notation. Although it is easy to write the Poynting vector (which is always real) as $\mathbf{S} = \mathrm{Re}(\mathbf{E}) \times \mathrm{Re}(\mathbf{H})$, it is not very helpful because one would have to evaluate the real parts of $\mathbf{E}$ and $\mathbf{H}$ individually. Besides, the frequency of oscillation is usually too high to be measured. For harmonic fields, the time-averaged Poynting vector can be expressed as

$$\langle \mathbf{S} \rangle = \frac{1}{2}\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*) \tag{8.18b}$$

where * signifies the complex conjugate. Similarly, the time-averaged energy density for time-harmonic fields becomes [5]

$$\langle u \rangle = \frac{1}{4}\varepsilon_\mathrm{m}\mathbf{E} \cdot \mathbf{E}^* + \frac{1}{4}\mu_\mathrm{m}\mathbf{H} \cdot \mathbf{H}^* \tag{8.19}$$

For an absorbing or dissipative medium, different approaches exist regarding the definition and determination of the electromagnetic energy density especially when magnetic materials are involved [13, 14].

Example 8.1 Prove that Eq. (8.18b) is the time-averaged Poynting vector for time-harmonic fields.

Solution Let $\mathbf{E} = \mathbf{E}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t}$ and $\mathbf{H} = \mathbf{H}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t}$, where $\mathbf{E}(\mathbf{r})$ and $\mathbf{H}(\mathbf{r})$ are complex vectors. Integrating the Poynting vector over a period T, we have

$$\begin{aligned}
\langle \mathbf{S} \rangle &= \frac{1}{T}\int_T \mathrm{Re}(\mathbf{E}) \times \mathrm{Re}(\mathbf{H})\mathrm{d}t \\
&= \frac{1}{4T}\int_T \left[\mathbf{E}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t} + \mathbf{E}^*(\mathbf{r})\mathrm{e}^{\mathrm{i}\omega t}\right] \times \left[\mathbf{H}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t} + \mathbf{H}^*(\mathbf{r})\mathrm{e}^{\mathrm{i}\omega t}\right]\mathrm{d}t \\
&= \frac{1}{4}\left(\mathbf{E} \times \mathbf{H}^* + \mathbf{E}^* \times \mathbf{H}\right) = \frac{1}{2}\mathrm{Re}\left(\mathbf{E} \times \mathbf{H}^*\right)
\end{aligned}$$

8.1.5 Dielectric Function

The conductivity is large at low frequencies for metals, due to free electrons. Even for good conductors, however, the electrons are not completely free but will be scattered by defects and phonons. At high frequencies, the current density $\mathbf{J}$ and the electric field $\mathbf{E}$ are not in phase anymore, suggesting that the conductivity should be a complex number. For insulators such as crystalline or amorphous dielectrics, electromagnetic waves can interact with bound electrons or lattice vibrations to transfer energy to the medium. At optical frequencies, the distinction between a conductor and an insulator becomes ambiguous unless the optical response over a wide frequency

region is considered. It is well known that a good conductor is highly reflective in a broad spectral region from the near infrared all the way to radio frequencies. Nevertheless, a dielectric material can also be highly reflective in certain frequency bands, especially in the mid-infrared region. At certain frequencies or in a narrow frequency band, the dielectric function $\varepsilon' + \mathrm{i}\varepsilon''$ may appear to be very similar for a metal and a dielectric material.

Let us consider a nonmagnetic material whose conductivity is σ. The wave equation for $\sigma \neq 0$ and $\mu_\mathrm{m} = \mu_0$ has the following form:

$$\nabla^2 \mathbf{E} = \mu_0 \sigma \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \varepsilon_\mathrm{m} \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{8.20}$$

Suppose Eq. (8.9) is a solution of this equation. We can substitute $\partial \mathbf{E}/\partial t = -\mathrm{i}\omega \mathbf{E}$, $\partial^2 \mathbf{E}/\partial t^2 = -\omega^2 \mathbf{E}$, and $\nabla^2 \mathbf{E} = -k^2 \mathbf{E}$ into Eq. (8.20) to obtain

$$k^2 = \mathrm{i}\omega\mu_0\sigma + \omega^2 \mu_0 \varepsilon_\mathrm{m} \tag{8.21}$$

Therefore, the wavevector becomes complex: $\mathbf{k} = \mathbf{k}' + \mathrm{i}\mathbf{k}''$, where $\mathbf{k}' = k_x' \hat{\mathbf{x}} + k_y' \hat{\mathbf{y}} + k_z' \hat{\mathbf{z}}$ and $\mathbf{k}'' = k_x'' \hat{\mathbf{x}} + k_y'' \hat{\mathbf{y}} + k_z'' \hat{\mathbf{z}}$ are real vectors. Note that Eq. (8.21) tells us the value of $k^2 = \mathbf{k} \cdot \mathbf{k} = k_x^2 + k_y^2 + k_z^2$, where each wavevector component may be complex, but does not specify the individual components. The *complex dielectric function* is defined as

$$\varepsilon = \varepsilon' + \mathrm{i}\varepsilon'' = \frac{\varepsilon_\mathrm{m}}{\varepsilon_0} + \mathrm{i}\frac{\sigma}{\omega\varepsilon_0} \tag{8.22}$$

For a nonmagnetic material, the *complex refractive index* $\tilde{n} = n + \mathrm{i}\kappa$ is related to the complex dielectric function by $\varepsilon = (n + \mathrm{i}\kappa)^2$. The imaginary part κ of the complex refractive index is called the *extinction coefficient*. By definition, we have

$$\varepsilon' = n^2 - \kappa^2 \quad \text{and} \quad \varepsilon'' = 2n\kappa \tag{8.23}$$

The refractive index n and the extinction coefficient κ are also called *optical constants*, although none of them are constant over a large wavelength region for real materials [15]. The dielectric function is also called relative permittivity, with respect to the permittivity of vacuum ε_0. One can consider the σ/ω term in Eq. (8.22) as the imaginary part of the permittivity. Some texts used $\varepsilon = \varepsilon' - \mathrm{i}\varepsilon''$ for the dielectric function and $\tilde{n} = n - \mathrm{i}\kappa$ for the complex refractive index. In doing so, Eq. (8.9) must be revised to $\mathbf{E} = \mathbf{E}_0 \mathrm{e}^{\mathrm{i}(\omega t - \mathbf{k}\cdot\mathbf{r})}$. In either convention, ε'' and σ must be nonnegative for a passive medium. Equation (8.21) can be rewritten as

$$k = \tilde{n}\omega/c_0 \tag{8.24}$$

For simplicity, we will remove the tilde and simply use n for the complex refractive index, where it can be clearly understood from the context.

By substituting i**k** for ∇ and $-\mathrm{i}\omega$ for $\partial/\partial t$, we can rewrite Maxwell's curl equations as

$$\mathbf{k} \times \mathbf{E} = \omega\mu_0\mathbf{H} \tag{8.25}$$

and

$$\mathbf{k} \times \mathbf{H} = -\omega\varepsilon_0\varepsilon\mathbf{E} \tag{8.26}$$

Similar to the definition of the complex dielectric function, one may choose to define a complex conductivity that satisfies Ohm's law at high frequencies, $\mathbf{J} = \tilde{\sigma}\mathbf{E}$, where

$$\tilde{\sigma} = \sigma' + \mathrm{i}\sigma'' = \sigma - \mathrm{i}\omega\varepsilon_\mathrm{m} \tag{8.27}$$

Note that we have assumed that σ is the real part of $\tilde{\sigma}$. Therefore,

$$\sigma'' = -\omega\varepsilon_0\varepsilon' \text{ and } \varepsilon'' = \sigma'/\omega\varepsilon_0 \tag{8.28}$$

Equation (8.26) can be recast in terms of the complex conductivity as

$$\mathbf{k} \times \mathbf{H} = -\mathrm{i}\tilde{\sigma}\mathbf{E} \tag{8.29}$$

In the subsequent discussion, we will omit the tilde above σ, when the context is sufficiently clear. The complex conductivity and the complex dielectric function are related to each other. For a linear, isotropic, and homogeneous nonmagnetic material, only two frequency-dependent functions are needed to fully characterize the electromagnetic response. The function pairs often found in the literature are (n, κ), $\left(\sigma', \varepsilon'\right)$, $(\varepsilon', \varepsilon'')$, and (σ', σ''). The principle of causality, which states that the effect cannot precede the cause, or no output before an input, imposes additional restrictions on the frequency dependence of the optical properties so that the real and imaginary parts are not completely independent, but related, to each other. In general, the relative permeability, which is complex and frequency dependent, can be expressed as

$$\mu = \mu' + \mathrm{i}\mu'' = \mu_\mathrm{m}/\mu_0 \tag{8.30}$$

The complex refractive index for magnetic materials should be defined as follows:

$$n = \sqrt{\varepsilon\mu} \tag{8.31}$$

The amplitude of the complex wavevector is $k = n\omega/c_0$, same as Eq. (8.24). One can verify that Eq. (8.9) is a solution of the wave equation. The relative permittivity ε and permeability μ will be used to formulate the general equations, later in this chapter. In most sections of this chapter, we deal with nonmagnetic materials, such as metals, dielectrics, and semiconductors. However, we will devote the discussion

of the optical properties of magnetic materials to Sect. 8.4.6, because of the emerging interest in *metamaterials*, which are synthesized materials with magnetic responses at microwave and higher frequencies (see Problem 8.6, for example).

8.1.6 Propagating and Evanescent Waves

In an absorbing nonmagnetic medium, the electric and magnetic fields will attenuate exponentially. As an example, consider a wave that propagates in the positive x-direction, with its electric field polarized in the y-direction. Then,

$$\mathbf{E} = \hat{\mathbf{y}} E_0 \mathrm{e}^{-\mathrm{i}(\omega t - k'x)} \mathrm{e}^{-k''x} \tag{8.32}$$

where $k' = \omega n/c_0$ and $k'' = \omega\kappa/c_0$ are the real and imaginary parts of the wavevector, respectively, that is, $\mathbf{k} = (k' + \mathrm{i}k'')\hat{\mathbf{x}}$. Equation (8.32) suggests that the amplitude of the electric field will decay exponentially according to $\mathrm{e}^{-(2\pi\kappa/\lambda)x}$. The magnetic field can be obtained from Eq. (8.25) as

$$\mathbf{H} = \hat{\mathbf{z}} \frac{n + \mathrm{i}\kappa}{\mu_0 c_0} E_0 \mathrm{e}^{-\mathrm{i}(\omega t - k'x)} \mathrm{e}^{-k''x} \tag{8.33}$$

By substituting Eqs. (8.32) and (8.33) into Eq. (8.18a), we obtain the time-averaged energy flux in the x-direction as

$$\langle S \rangle = \frac{n}{2\mu_0 c_0} E_0^2 \mathrm{e}^{-2k''x} = \frac{n}{2\mu_0 c_0} E_0^2 \mathrm{e}^{-a_\lambda x} \tag{8.34}$$

where $a_\lambda = 4\pi\kappa/\lambda$ is called the *absorption coefficient*. The inverse of a is called the *radiation penetration depth* (or photon mean free path) given by

$$\delta_\lambda = \frac{1}{a_\lambda} = \frac{\lambda}{4\pi\kappa} \tag{8.35}$$

It is the distance through which the radiation power is attenuated by a factor of e^{-1} ($\approx$37%). (See Problem 8.6 for some typical values of the penetration depth in various materials at different wavelengths and temperatures.)

When $\mathbf{k}$ is complex, the plane normal to $\mathbf{k}'$ is the constant-phase plane and the plane normal to $\mathbf{k}''$ is the constant-amplitude plane because

$$\mathbf{E} = \mathbf{E}_0 \mathrm{e}^{-\mathrm{i}(\omega t - \mathbf{k}'\cdot\mathbf{r})} \mathrm{e}^{-\mathbf{k}''\cdot\mathbf{r}} \tag{8.36}$$

When $\mathbf{k}' \times \mathbf{k}'' = 0$, the wave is said to be *homogeneous*; otherwise, the constant-phase planes will not be parallel to the constant-amplitude planes, and the wave is said to be *inhomogeneous*. A typical homogeneous wave is given in Eq. (8.32). Now let us

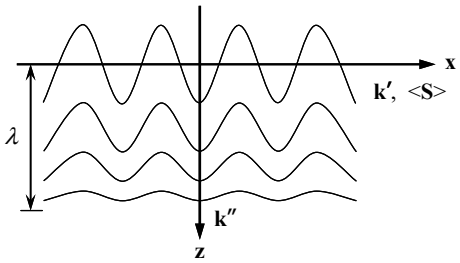

Fig. 8.3 Schematic of an evanescent wave near the $z = 0$ surface

use an example to illustrate an inhomogeneous wave. Consider a wave, defined in the $z \geq 0$ half plane filled with vacuum, with a wavevector $\mathbf{k} = 2\omega/c_0\hat{\mathbf{x}} + \mathrm{i}\sqrt{3}\omega/c_0\hat{\mathbf{z}}$. The electric field is linearly polarized in the y-direction; thus, $\mathbf{E} = \hat{\mathbf{y}}E_0\mathrm{e}^{-\mathrm{i}(\omega t - \mathbf{k}\cdot\mathbf{r})}$. It can be shown that $\mathbf{k}\cdot\mathbf{k} = k^2 = \omega^2/c_0^2$; hence, $\mathbf{k}$ is indeed a valid wavevector in vacuum. The electric field can be written as

$$\mathbf{E} = \hat{\mathbf{y}}E_0\mathrm{e}^{-\mathrm{i}(\omega t - k_x x)}\mathrm{e}^{-\eta z} \tag{8.37}$$

Here, $k_x = 2k = 4\pi/\lambda$, and $\eta = \mathrm{Im}(k_z) = 2\sqrt{3}\pi/\lambda$. Clearly, the wave has a constant phase for any constant-x plane and a constant amplitude for any constant-z plane. Furthermore, the amplitude decays exponentially toward the positive z-direction and becomes negligible, when $z > \lambda$, as shown schematically in Fig. 8.3. Such a wave is called an *evanescent wave*, which exists in waveguides and is important for near-field optics and nanoscale radiation heat transfer. It can be shown that the time-averaged Poynting vector is parallel to the x-direction so that no energy is transported toward the z-direction (see Problem 8.7).

8.2 Blackbody Radiation: The Photon Gas

This section deals with Planck's law of blackbody radiation, which is the foundation of far-field radiation heat transfer analysis. After the discussion of radiation thermometry and radiance temperature, we will study radiation (or photon) entropy and pressure. Photon entropy and exergy may be important for analyzing and designing advanced energy harvesting systems from solar radiation to near-field thermal radiative devices. The limitation of Planck's law is also addressed.

8.2.1 Planck's Law

Consider an enclosure of volume V, whose walls are at a uniform temperature T, as shown in Fig. 8.4a. The enclosure may contain a medium (such as a molecular gas) or

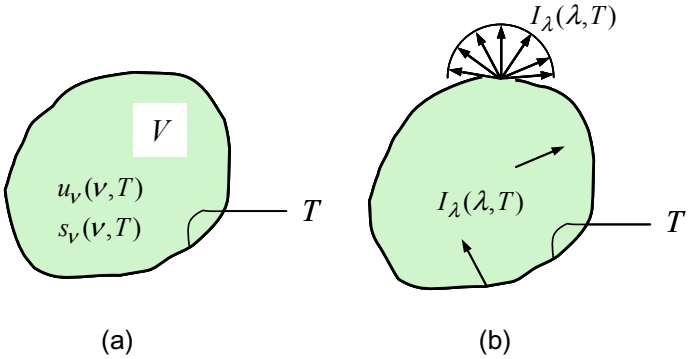

Fig. 8.4 An isothermal enclosure (blackbody cavity): **a** without an opening and **b** with a small opening on the wall that has little effect on the equilibrium distribution

may be evacuated (vacuum). Inside the enclosure, there exist electromagnetic fields, which may be viewed either as many transverse waves at different frequencies or as a large number of quanta with different energies. The particle theory treats radiation as a collection of photons. The energy and the momentum of each photon are related to the frequency and the speed of light, by $\varepsilon = h\nu$ and $p = h\nu/c$, respectively. We are interested in finding the equilibrium distribution of photons with respect to photon energy or frequency or momentum. Photons obey Bose–Einstein statistics, without requiring the total number be conserved. The number of photons $\mathrm{d}N$ in a frequency interval from ν to $\nu + \mathrm{d}\nu$ per unit volume is equal to the *mean occupation number* multiplied by the number of quantum states (degeneracy):

$$\mathrm{d}N = f_{\mathrm{BE}}(\nu)\mathrm{d}g = \frac{\mathrm{d}g}{\mathrm{e}^{h\nu/k_{\mathrm{B}}T} - 1} \tag{8.38}$$

The quantum states in the phase space, consisting of a volume V and a spherical shell in the momentum space (from p to $p + \mathrm{d}p$), are given by $\mathrm{d}g = 2V(4\pi p^2 \mathrm{d}p)/h^3$, where the factor 2 accounts for the two polarization states of electromagnetic waves. Thus, we can write the density of states (DOS), which is the number of quantum states per unit volume per unit frequency interval, as

$$D(\nu) = \frac{1}{V}\frac{\mathrm{d}g}{\mathrm{d}\nu} = \frac{8\pi\nu^2}{c^3} \tag{8.39}$$

Notice that c is the speed of light in the medium and it is assumed that the refractive index of the medium is real and independent of the frequency. Because the Bose–Einstein distribution function gives the mean occupation number of each quantum state, the number of photons per unit volume per unit frequency interval is

$$f(\nu) = \frac{1}{V}\frac{\mathrm{d}N}{\mathrm{d}\nu} = f_{\mathrm{BE}}(\nu)D(\nu) = \frac{8\pi\nu^2}{c^3(\mathrm{e}^{h\nu/k_{\mathrm{B}}T} - 1)} \tag{8.40}$$

Integrating the above equation over all frequencies yields the total number of photons at a given temperature per unit volume. Clearly, the number of photons is not conserved in a blackbody cavity with a fixed volume V.

Since the energy of a photon is $h\nu$, the spectral energy density (energy per unit volume per unit frequency interval) at a fixed temperature T can be written as

$$u_\nu(\nu) = h\nu f(\nu) = \frac{8\pi h\nu^3}{c^3(\mathrm{e}^{h\nu/k_\mathrm{B}T} - 1)} \tag{8.41}$$

For an area element inside the enclosure, the radiant energy flux is related to the energy density and the speed of light by

$$q''_{rad,\nu} = \frac{u_\nu c}{4} \tag{8.42}$$

If a blackbody is placed inside the enclosure, it will absorb all incoming radiant energy that reaches its surface; at thermal equilibrium, it must emit the same amount of energy. After substituting Eq. (8.41) into Eq. (8.42), we obtain the *spectral emissive power* of a blackbody as a function of frequency and temperature as

$$e_{\mathrm{b},\nu}(\nu, T) = \frac{2\pi h\nu^3}{c^2(\mathrm{e}^{h\nu/k_\mathrm{B}T} - 1)} \tag{8.43}$$

Note that the spectral emissive power is the power emitted per unit area per frequency or wavelength interval. To express the spectral emissive power in terms of the wavelength (in vacuum), we can substitute $c = c_0/n$, $\nu = c_0/\lambda$, $\mathrm{d}\nu = -c_0\mathrm{d}\lambda/\lambda^2$, and $e_{\mathrm{b},\nu}\mathrm{d}\nu = -e_{\mathrm{b},\lambda}\mathrm{d}\lambda$ into Eq. (8.43). Therefore,

$$e_{\mathrm{b},\lambda}(\lambda, T) = \frac{2\pi hc_0^2n^2}{\lambda^5(\mathrm{e}^{hc/k_\mathrm{B}\lambda T} - 1)} = \frac{C_1n^2}{\lambda^5(\mathrm{e}^{C_2/\lambda T} - 1)} \tag{8.44}$$

where $C_1 = 3.742 \times 10^8\ \mathrm{W\,m^{-2}\ \mu m^4}$ and $C_2 = 1.439 \times 10^4\ \mu\mathrm{m\ K}$ [10] are called the first and second radiation constants. Equations (8.43) and (8.44) are called *Planck's law* or *Planck's distribution* (of blackbody radiation) in terms of the frequency and wavelength, respectively. It should be noted that the blackbody intensity is $I_{\mathrm{b},\lambda}(\lambda, T) = e_{\mathrm{b},\lambda}(\lambda, T)/\pi$, as in Eq. (2.48), and isotropic inside the whole cavity regardless of the radiative properties of the wall. Furthermore, when there is a small opening, the emitted radiation is diffuse and obeys the blackbody distribution, as shown in Fig. 8.4b. The requirement is that the opening should be sufficiently small compared with the size of the enclosure, but large enough compared to the wavelengths of interest. The concept of blackbody cavity was made clear by Wien in his 1911 Nobel lecture, as seen from the excerpt below:

> … there must exist, in a cavity surrounded by bodies of equal temperature, a radiation energy that is independent of the nature of the bodies. If in the walls surrounding this cavity a small aperture is made through which radiation issues, we obtain a radiation which is independent

of the nature of the emitting body, and is wholly determined by the temperature. The same radiation would also be emitted by a body which does not reflect any rays and which is therefore designated as completely black, and this radiation is called the radiation of a black body or blackbody radiation.

It should be noted that if the refractive index is a weak function of wavelength and absorption by the medium is negligible, c^3 in Eqs. (8.40) and (8.41) should be replaced by $c^2 v_g$ where v_g is the group velocity defined in Eq. (8.13). The group velocity v_g should also be used in Eq. (8.42) to replace c. Nevertheless, the expressions of the emissive power given in Eqs. (8.43) and (8.44) remain the same. In the following, $n = 1$ or vacuum is assumed for Planck's distribution unless otherwise indicated.

Equation (8.43) or (8.44) can be integrated over the whole spectrum to obtain the Stefan–Boltzmann law: $e_b = \sigma_{SB} T^4$. In Fig. 8.5, $e_{b,\lambda}/\sigma_{SB} T^5$ is plotted as a function of λT so that the area under Planck's distribution (solid curve) is $\int_0^\infty \frac{e_{b,\lambda}(\lambda,T)}{\sigma_{SB} T^5} d(\lambda T) = \frac{1}{\sigma_{SB} T^4} \int_0^\infty e_{b,\lambda}(\lambda, T) d\lambda = 1$. The Planck's distribution has a peak and approaches zero at extremely short and long wavelengths. If $C_2/\lambda T \gg 1$, the right-hand side of Eq. (8.44) can be approximated by $C_1 \lambda^{-5} e^{-C_2/\lambda T}$. This is called Wien's formula, which gives good approximation, even beyond the maximum emissive power, as can be seen from Fig. 8.5. At very long wavelengths, Wien's formula underpredicts the emissive power and asymptotically approaches to $C_1 \lambda^{-5}$, suggesting that the emissive power is independent of temperature. Note that the right-hand side of Eq. (8.44) approaches $C_1 T/(C_2 \lambda^4)$ if $C_2/\lambda T \ll 1$, since $e^x - 1 \approx x$ for $x \ll 1$. This is called the Rayleigh–Jeans formula, which is applicable at very long wavelengths, as shown in Fig. 8.5. The significance of the Rayleigh–Jeans formula is that it correctly predicts the temperature dependence of the blackbody spectrum, at very long wavelengths, where Wien's formula fails. The failure of the Rayleigh–Jeans formula at short wavelengths is called the *ultraviolet catastrophe*. The significance of Planck's formula is more than a unified mathematical formulation. It was derived based on the hypothesis of energy quanta that do not exist in classical Newtonian mechanics or Maxwell's electrodynamics. It should be

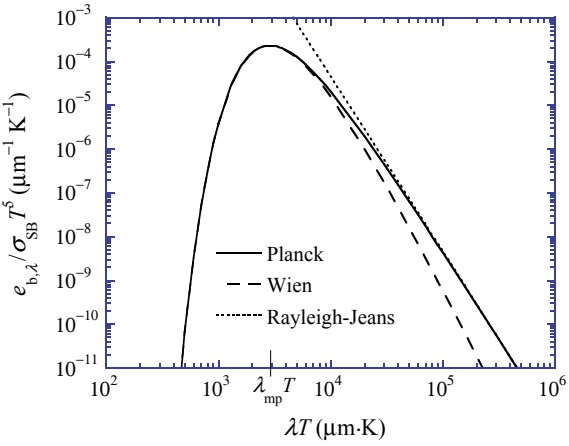

Fig. 8.5 Planck's law for blackbody emissive power

noted that the preceding derivation is based on statistical thermodynamics, presented in Chap. 3, rather than on Planck's original semi-classical oscillator model [4].

Example 8.2 Find the wavelength λ_{mp} at which Planck's distribution reaches a maximum. What is the ratio of the energy emitted at $\lambda < \lambda_{mp}$ to that at $\lambda > \lambda_{mp}$?

Solution By setting the derivative of Eq. (8.44) equal to zero, i.e., $de_{b,\lambda}/d\lambda = 0$, we have

$$\frac{hc}{k_B \lambda T} + 5\exp\left(-\frac{hc}{k_B \lambda T}\right) - 5 = 0$$

This equation can be solved by iteration or numerically to yield

$$\lambda_{mp}\ [\mu m] = \frac{2898\ \mu m\,K}{T\ [K]} \tag{8.45}$$

This is *Wein's displacement law*. The location of λ_{mp} is also marked on Fig. 8.5. To find out the ratio of the energy emitted at $\lambda < \lambda_{mp}$ to that at $\lambda > \lambda_{mp}$, we can numerically evaluate $\int_0^{\lambda_{mp}} e_{b,\lambda}(\lambda, T)d\lambda / \int_{\lambda_{mp}}^{\infty} e_{b,\lambda}(\lambda, T)d\lambda$. The numerical result is approximately 1:3 and independent of temperature.

Example 8.3 Assuming the sun to be a blackbody at 5800 K, calculate the emissive power at the following wavelength intervals: $\lambda < 0.3\ \mu m$, $0.3\ \mu m < \lambda < 0.4\ \mu m$, $0.4\ \mu m < \lambda < 0.7\ \mu m$, $0.7\ \mu m < \lambda < 3\ \mu m$, and $\lambda > 3\ \mu m$. Neglect the absorption by the atmosphere. What is the radiant power arriving at the earth's surface from the sun?

Solution The total emissive power is $\sigma_{SB} T_{sun}^4 = 5.67 \times 10^{-8} \times 5800^4 \approx 64\ MW/m^2$. We can obtain the emissive power in each spectral region by integrating Eq. (8.44), as listed in the following table:

λ (μm)	<0.3	0.3–0.4	0.4–0.7	0.7–3	>3	Total
$\lambda_2 T$ (μmK)	1740	2320	4060	17400	∞	_
$F_{0\to\lambda_2}$	0.03	0.12	0.49	0.98	1	_
$F_{\lambda_1\to\lambda_2}$	0.03	0.09	0.37	0.49	0.02	1
$\Delta E_b\ (MW/m^2)$	1.9	5.8	23.7	31.4	1.3	64.1

Note that $F_{\lambda_1\to\lambda_2}$ represents the fraction of radiation falling between λ_1 and λ_2. The total power emitted by the sun equals the emissive power multiplied by the surface area of the sun. The fraction of the power that reaches the earth equals the solid angle of the earth divided by 4π. Note that the radius of the sun $r_{sun} = 6.955 \times 10^8$ m, the radius of the earth $r_{earth} = 6.378 \times 10^6$ m, and the earth–sun distance $R_{earth-sun} = 1.496 \times 10^{11}$ m. Therefore, the total power that will reach the earth's surface, if the

Table 8.1 Spectral regions expressed in different units

	Wavelength λ (μm)	Wavenumber $\bar{\nu}$ (cm^{-1})	Frequency ν (THz)	Angular frequency ω (10^{14} rad/s)	Photon energy E (eV)
UV	0.01–0.38	$(10 - 0.26) \times 10^5$	30,000–790	1900–50	120–3.3
VIS	0.38–0.76	$(2.6 - 1.3) \times 10^4$	790–400	50–25	3.3–1.6
NIR	0.76–2.5	$(1.3 - 0.4) \times 10^4$	400–120	25–7.5	1.6–0.5
MIR	2.5–25	4000–400	120–12	7.5–0.75	0.5–0.05
FIR	25–1000	400–10	12–0.3	0.75–0.019	0.05–0.0012
MW	10^3–10^5	10–0.1	0.3–0.003	$(19–0.19) \times 10^{-3}$	$(12–0.12) \times 10^{-4}$

absorption by the atmosphere is neglected, is

$$\dot{Q} = 4\pi r_{\text{sun}}^2 \cdot \sigma_{\text{SB}} T_{\text{sun}}^4 \cdot \frac{\pi r_{\text{earth}}^2}{4\pi R_{\text{earth-sun}}^2} \approx 1.8 \times 10^{17} \text{ W}$$

The average irradiation on the earth is: $G = \dot{Q}/\pi r_{\text{earth}}^2 \approx 1377$ W/m^2. This value is very close to the total solar irradiance (TSI), measured outside the earth's atmosphere.

Because of the broad spectral region of electromagnetic waves, alternative units are often used, such as wavelength λ (in vacuum), wavenumber $\bar{\nu} = 1/\lambda$, frequency $\nu = c_0/\lambda$, angular frequency $\omega = 2\pi\nu$, and photon energy $E = h\nu$. Generally speaking, optical radiation covers the spectral region including ultraviolet (UV), visible (VIS), near infrared (NIR), mid infrared (MIR), and far infrared (FIR). Table 8.1 outlines the subdivisions of the spectral region in different units from ultraviolet (UV) to microwave (MW). Note that $\bar{\nu}$ [cm^{-1}] $= 10,000/\lambda$ [μm] and E [eV] $= 1.24/\lambda$ [μm]. Thermal radiation covers part of the UV from $\lambda = 0.1$ μm through some of the MW region.

8.2.2 Radiation Thermometry

The developments of the absolute temperature scale and radiation thermometry are among the most important applications of blackbody radiation [16]. The Stefan–Boltzmann law $e_{\text{b}} = \sigma_{\text{SB}} T^4$ defines an absolute thermodynamic temperature, which is consistent with the one defined by the ideal gas law and the Carnot cycle. While radiation thermometry can serve as a primary standard, most practical radiation thermometers are not absolute instruments because of other considerations such as fast response, easy operation, and low cost. High-temperature furnaces are commonly used as calibration standards. The cavity is a hollow cylinder, made of graphite for

example, with a conical ending and a small aperture. The most accurate calibration source is the fixed-point heat pipe blackbody, for which a pure metal is melted outside the graphite cylinder to maintain a constant temperature in a two-phase state. The freezing temperatures are then used to define the temperature scales (1234.93 K for Ag, 1337.33 for Au, and 1357.77 K for Cu).

To measure the absolute temperature of a thermally radiative body, two blackbody cavities at different temperatures would be needed: one serves as the emitter (blackbody source) and the other as the receiver (radiometer). Quinn and Martin [17] used a blackbody source and a cryogenic radiometer to directly determine the thermodynamic temperatures and measure the Stefan–Boltzmann constant. The experimentally obtained Stefan–Boltzmann constant was $(5.66967 \pm 0.00076) \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$. The difference is 0.13% of the theoretical value $(5.67040 \pm 0.00004) \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$, based on Planck's constant, Boltzmann's constant, and the speed of light. Since the early 1990s, NIST has developed a high-accuracy cryogenic radiometer (HACR) facility to serve as the primary standard for optical radiation measurements. A schematic of the original HACR receiver is shown in Fig. 8.6. The receiver is mounted at the bottom of a liquid helium cryostat in an evacuated chamber, and the optical access is through a Brewster window below the cavity. The HACR facility has gone through some major upgrades in recent years. The receiver cavity is made of copper with a high thermal conductivity and low specific heat at cryogenic temperatures. The inner wall of the cavity is coated with a specular black paint to absorb the incident radiation with an effective absorptance greater than 99.998%. The electrical-substitution technique links the radiant power to the electric power to achieve an overall uncertainty within 0.02% for optical power measurements. Detailed descriptions can be found from Pearson and Zhang [18] and references therein. The cosmic radiation background (in the far-infrared and microwave region), measured with cryogenic bolometers, can be fitted to the blackbody distribution at 2.7 K; this is the temperature of the universe at the present time. The discovery of cosmic radiation background in 1964 and the subsequent measurements and theoretical studies have been recognized by the Nobel Prizes in Physics to Arno Penzias and Robert Wilson in 1978 and to John Mather and George Smoot in 2006.

Most radiation thermometers are based on spectral measurements rather than on the measurement of the total irradiance from the target. When a radiation thermometer is used to measure the temperature of a real surface, the unknown emittance of the surface and the influence of the surrounding radiation are the major issues that affect the measurement. Various methods have been developed to deal with these problems, including the creation of a blackbody cavity on the surface, the two-color method, and the use of a controlled reference source. The development of optical fibers has allowed radiometric temperature measurements for surface locations that are otherwise inaccessible by imaging radiometers. The detailed theory and practice of radiation thermometry can be found from the two book volumes compiled by Zhang et al. [19]. A brief discussion of the basic operational principles of spectral radiation thermometry is given in the following.

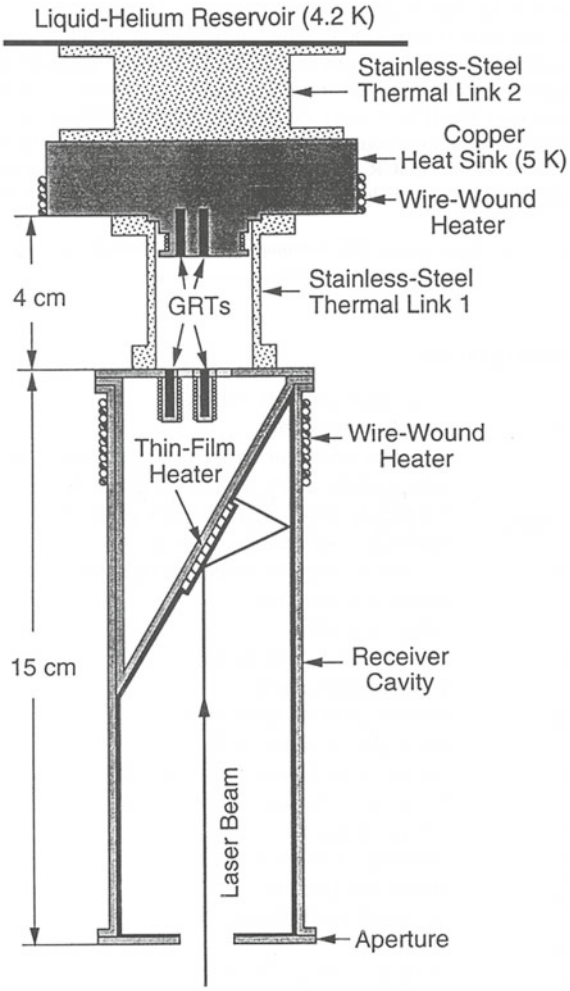

Fig. 8.6 Schematic of the receiver cavity of an absolute cryogenic radiometer, where GRT stands for germanium resistance thermometer, from Pearson and Zhang [18]

The *measurement equation* of a spectral radiation thermometer can be approximated as follows:

$$V_{\mathrm{d}} = C_{\mathrm{I}} I_{\mathrm{ex},\lambda}(\lambda) \tag{8.46}$$

where V_{d} is the detector output signal and C_{I} is an instrument constant that is independent of the target material and temperature. The term $I_{\mathrm{ex},\lambda}(\lambda)$ is called the *exitent spectral radiance*, which includes the radiation emitted by the target and the surroundings, as well as that reflected by the target. The *radiance temperature* T_λ (also called the *brightness temperature*) is defined according to

$$I_{\mathrm{b},\lambda}(\lambda, T_\lambda) = I_{\mathrm{ex},\lambda}(\lambda) \tag{8.47}$$

where $I_{b,\lambda}(\lambda, T_\lambda)$ is the blackbody intensity at the wavelength λ and temperature T_λ. If the surrounding emission and absorption can be neglected, the exitent spectral radiance is due only to the emission; therefore,

$$I_{\mathrm{ex},\lambda}(\lambda) = I_{\mathrm{e},\lambda}(\lambda, T) = \varepsilon'_\lambda I_{\mathrm{b},\lambda}(\lambda, T) \tag{8.48}$$

where ε'_λ is the directional-spectral emittance, and $I_{\mathrm{e},\lambda}(\lambda, T)$ is the intensity emitted by the target. By combining Eqs. (8.47) and (8.48) and applying Wien's formula, the surface temperature is related to the radiance temperature by

$$\frac{1}{T} = \frac{1}{T_\lambda} + \frac{\lambda}{C_2} \ln \varepsilon'_\lambda \tag{8.49}$$

The uncertainty in the measured temperature due to an uncertainty in the emittance is

$$\frac{\delta T}{T} = -\frac{\lambda T}{C_2} \frac{\delta \varepsilon'_\lambda}{\varepsilon'_\lambda} \tag{8.50}$$

The impact of emittance on the temperature measurement decreases as λ decreases. Equation (8.50) suggests that it may be advantageous to choose a wavelength that is somewhat shorter than the wavelength at which $I_{\mathrm{b},\lambda}(\lambda, T)$ is a maximum as given by Wien's displacement law. If the surrounding radiation is not negligible, $I_{\mathrm{ex},\lambda}(\lambda)$ is the sum of the emitted and reflected spectral radiances. In practice, when choosing the operating wavelength, one should also consider the material's properties and the effect of surrounding radiation, as well as the detector availability and sensitivity. Hence, the choice of a radiation thermometer requires a detailed analysis of different effects in the actual measurements.

Example 8.4 Rapid thermal processing is a semiconductor single-wafer manufacturing technique. A lightpipe radiation thermometer, operated at $\lambda = 0.95$ μm, is used to measure the temperature of the wafer. The emittance or emissivity of a plain silicon wafer is approximately 0.7 at this wavelength. Neglect the reflected radiation from the wafer. If the wafer is at a temperature of 1200 K, what is the radiance temperature? If the temperature needs to be determined within an uncertainty of 1 K, how much tolerance in the emittance error is acceptable?

Solution From Eq. (8.49), $T_\lambda \approx 1167$ K, which differs from the actual temperature by approximately 33 K. One can also solve Eqs. (8.47) and (8.48), using Planck's law, and the result is essentially the same. Based on Eq. (8.50), to obtain a temperature within 1 K, the emittance must be determined within an uncertainty of $\delta\varepsilon'_\lambda = 0.0074$. Zhou et al. [20] developed a model to predict the effective emittance of silicon wafers in rapid thermal processing furnaces and showed that, by using a reflective cavity, the temperature measurement uncertainty can be significantly reduced.

8.2.3 Radiation Pressure and Photon Entropy

Like other particles, photon gas also has the property of entropy and can be related to other properties in equilibrium states. Express the energy density in an enclosure of volume V, at thermodynamic equilibrium, with a temperature T as $u = \frac{U}{V} = \frac{4}{c}\sigma_{SB}T^4$. It can be seen that the specific heat at constant volume is $c_v = \left(\frac{\partial u}{\partial T}\right)_V = \frac{16}{c}\sigma_{SB}T^3$. The radiation entropy or photon entropy can therefore be obtained as

$$S = \int_0^T V c_v \frac{\mathrm{d}T}{T} = \frac{16}{3c} V \sigma_{SB} T^3 \tag{8.51a}$$

or

$$s = \frac{16}{3c}\sigma_{SB}T^3 \tag{8.51b}$$

Note that $T = \left(\frac{\partial U}{\partial S}\right)_V$ is satisfied. The Helmholtz free energy $A = U - TS = -\frac{4}{3c}V\sigma_{SB}T^4$. Thus, the radiation pressure is

$$P = -\left(\frac{\partial A}{\partial V}\right)_T = \frac{4}{3c}\sigma_{SB}T^4 \tag{8.52}$$

The force from radiation pressure, albeit small, has some important applications in trapping and manipulating atomic to molecular particles. This technique is called optical traps or optical tweezers; see Lang and Block [21] for a bibliographical review. Arthur Ashkin shared the Nobel Prize in Physics in 2018 "for the optical tweezers and their application to biological systems." Another way to view radiation pressure is that photons or electromagnetic waves carry both energy and momentum. The interaction of electromagnetic waves or photons with matter therefore involves a change of momentum, resulting in a pressure on the object. Radiation pressure was first predicted by J. C. Maxwell in 1873 and experimentally demonstrated by Russian physicist P. Lebedev in 1900 and American physicists E. F. Nichols and G. F. Hull in 1901. Solar radiation pressure plays a role in the formation of a comet's dust tail. Radiation pressure is also important in cosmology concerning the formation and evaluation of the stars and galaxies.

If each photon mode (frequency) is individually considered, the *spectral entropy density* for unpolarized radiation can be expressed as follows [4]:

$$s_\nu(\nu, T) = \frac{8\pi k_B \nu^2}{c^3}\left[\frac{x}{e^x - 1} + \ln\left(\frac{e^x}{e^x - 1}\right)\right] \tag{8.53}$$

where $x = \frac{h\nu}{k_B T}$. Note that $\frac{1}{T} = \left(\frac{\partial s_\nu}{\partial u_\nu}\right)_\nu = \frac{k_B}{h\nu}\ln\left(1 + \frac{8\pi h\nu^3}{c^3 u_\nu}\right)$, which is consistent with Eq. (8.41). Similar to the energy flux (emissive power) and intensity, the *radiation entropy flux* can be obtained by multiplying a factor $c/4$ to Eqs. (8.51b) and (8.53), and the *radiation entropy intensity* can be obtained by dividing the flux by π, because of

the isotropic nature of blackbody radiation. Clearly, electromagnetic radiation carries both energy and entropy.

Example 8.5 Consider the radiation heat transfer between two parallel plates at T_1 and T_2, respectively. Assume each plate has an area of A and both plates are blackbodies. The separation distance is much smaller than $\sqrt{A}$ but much greater than the wavelength of thermal radiation.

(a) How much entropy is generated at each plate? Evaluate the ratio of entropy generation assuming that $T_1 = 2T_2$.
(b) If a thermophotovoltaic receiver is mounted on the lower temperature side to convert thermal radiative energy to electricity (work), what is its maximum achievable efficiency?

Solution

(a) The net energy flow from plate 1 to 2 is $\dot{Q}_{12} = A\sigma_{SB}(T_1^4 - T_2^4)$. The entropy of plate 1 will decrease at the rate of $\mathrm{d}S_1/\mathrm{d}t = -\dot{Q}_{12}/T_1$, and the entropy of plate 2 will increase at the rate of $\mathrm{d}S_2/\mathrm{d}t = \dot{Q}_{12}/T_2$. On the other hand, the net entropy flow from plate 1 to 2 can be calculated as $\dot{S}_{12} = \frac{4}{3}A\sigma_{SB}\left(T_1^3 - T_2^3\right)$. Therefore, $\dot{S}_{\mathrm{gen},1} = -\frac{\dot{Q}_{12}}{T_1} + \dot{S}_{12} = A\sigma_{SB}\left(\frac{1}{3}T_1^3 - \frac{4}{3}T_2^3 + \frac{T_2^4}{T_1}\right)$, $\dot{S}_{\mathrm{gen},2} = A\sigma_{SB}\left(\frac{1}{3}T_2^3 - \frac{4}{3}T_1^3 + \frac{T_1^4}{T_2}\right)$. The combined total entropy generation is equal to $\dot{Q}_{12}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)$, as expected. It can be shown that the entropy generation at each plate is always greater than zero if $T_1 \neq T_2$, or equal to zero if $T_1 = T_2$. When $T_1 = 2T_2$, the entropy generation by plate 1 is about one-quarter and that by plate 2 is about three-quarters of the total entropy generated.

(b) The available energy or exergy of thermal radiation is defined as the maximum work that can be produced by a system with respect to a large reservoir. In the present example, we may assume that the reservoir is at the same temperature as T_2. Suppose an amount of heat is taken from the high-temperature plate. We would like to find out the maximum work that can possibly be produced. Let us consider a reversible heat engine at T_2. The radiative energy leaving surface 1 can still be described by $\dot{Q}_1 = A\sigma_{SB}(T_1^4 - T_2^4)$, and the entropy leaving surface 1 is $\dot{S}_1 = \frac{4}{3}A\sigma_{SB}\left(T_1^3 - T_2^3\right)$. Therefore, the entropy generation in plate 1 cannot be eliminated. In other words, it is impossible to achieve the Carnot efficiency of $\eta_{\mathrm{Carnot}} = 1 - T_2/T_1$. The maximum work can be obtained when the irreversibility at the lower temperature plate is negligible and the heat engine is also reversible. It can easily be shown that the maximum work $\dot{W}_{\max} = \dot{Q}_1 - T_2\dot{S}_1$, and the optimal efficiency is given by

$$\eta_{\mathrm{opt}} = \frac{\dot{W}_{\max}}{\dot{Q}_1} = 1 - \frac{4(1 + y + y^2)}{3(1 + y)(1 + y^2)} \tag{8.54a}$$

where $y = T_1/T_2 \geq 0$. When $y = 2$, we obtain an optimal efficiency $\eta_{\mathrm{opt}} = 37.8\%$, which is less than the Carnot efficiency of 50%, because of the unrecoverable irreversibility at plate 1.

Consider a black receiver on the earth's surface that converts solar radiation to electricity. Since the incoming radiation is from a narrow solid angle, the emitted radiation can be assumed to be in equilibrium with the surroundings at T_0. Therefore, the received radiant power and entropy flux is $\dot{Q}_1 = \phi A \sigma_{SB} T_s^4$ and $\dot{S}_1 = \frac{4}{3}\phi A \sigma_{SB} T_s^3$, where T_s is the temperature of the sun and ϕ is a fraction accounting for the view angle and atmospheric transmittance (neglecting the scattering effect). Assuming a power $\dot{W}$ is developed, the heat transferred to the surroundings is $\dot{Q}_0 = \dot{Q}_1 - \dot{W}$ and the entropy transferred to the surroundings is $\dot{Q}_0/T_0$. In a reversible energy conversion device, the entropy generation must be zero and the maximum efficiency is obtained as

$$\eta_{opt} = \frac{\dot{W}_{max}}{\dot{Q}_1} = 1 - \frac{4}{3}\frac{T_0}{T_s} \tag{8.54b}$$

Different formulas on the optimal efficiency exist for solar energy conversion devices due to the different model assumptions used. For example, we can set $T_1 = T_s$ and $T_2 = T_0$ in Eq. (8.54a) and change the denominator $\dot{Q}_1$ to the absorbed solar radiation $\dot{Q}_{1,s} = A\sigma_{SB} T_1^4$. Then one would obtain Petela's formula [22]:

$$\eta_{opt} = \frac{\dot{W}_{max}}{\dot{Q}_{1,s}} = 1 - \frac{4}{3}\frac{T_0}{T_s} + \frac{1}{3}\left(\frac{T_0}{T_s}\right)^4 \tag{8.54c}$$

Since T_0/T_s is about one-twentieth, the difference between Eqs. (8.54b) and (8.54c) is practically negligible. A comprehensive discussion on energy conversion efficiency can be found from the review of Landsberg and Tonge [23]. This topic is of contemporary interest especially when dealing with near-field radiative energy conversion devices [24].

The next question one may ask is whether temperature can be defined for laser radiation. The answer is *yes,* and the temperature for high-intensity lasers can be very high. An intuitive guess is to define the temperature, based on the intensity I_ν of the laser or the monochromatic radiation, by setting $I_\nu = I_{b,\nu}(\nu, T_\nu)$. The definitions of entropy and thermodynamic temperature for optical radiation are very important for analyzing optical energy conversion systems, such as solar cells, thermophotovoltaic generators, luminescence devices, and laser cooling apparatus [25, 26]. Assume that the monochromatic radiation is from a thermodynamic equilibrium state, such as a resonance cavity that allows only a single mode to exist. The spectral entropy intensity of unpolarized radiation can be written as follows [4, 25]:

$$L_\nu = \frac{2k_B\nu^2}{c^2}\left[\left(1 + \frac{c^2 I_\nu}{2h\nu^3}\right)\ln\left(1 + \frac{c^2 I_\nu}{2h\nu^3}\right) - \frac{c^2 I_\nu}{2h\nu^3}\ln\left(\frac{c^2 I_\nu}{2h\nu^3}\right)\right] \tag{8.55}$$

Thermodynamically, the *monochromatic radiation temperature* can be defined as $\frac{1}{T_\nu(\nu)} = \left(\frac{\partial L_\nu}{\partial I_\nu}\right)_\nu$ and given as

$$\frac{1}{T_\nu(\nu)} = \left(\frac{\partial L_\nu}{\partial I_\nu}\right)_\nu = \frac{k_\mathrm{B}}{h\nu}\ln\left(1 + \frac{2h\nu^3}{c^2 I_\nu}\right) \tag{8.56}$$

This is indeed Planck's distribution of intensity at the same temperature. The expressions can be modified for polarized radiation. When the energy intensity is very high, Eq. (8.56) approaches $T_\nu(\nu) = \frac{c^2 I_\nu}{2k_\mathrm{B}\nu^2}$, which is in the Rayleigh–Jeans limit. The radiation temperature will be proportional to the intensity of the monochromatic radiation and can exceed 10^{10} K, with a 1–mW He-Ne laser at 632.8 nm wavelength [26]. Therefore, for lasers with a moderate intensity, T_ν tends to be so high that the entropy is nearly zero; hence, the laser power can be considered as "work." If a collimated beam is randomly scattered by a rough surface, the scattered radiation will have a much lower intensity because of the increase in the solid angle. The process is accompanied with an entropy increase and is thus irreversible. It is not possible to increase the intensity of the scattered light, back to its original intensity, without leaving any net effect on the environment. On the other hand, if a nearly collimated light is split into two beams with a beamsplitter, the transmitted and reflected beams can interfere with each other to reconstruct the original beam. This process is reversible because the two beams are *correlated*. The correlated beams have lower entropy than those with the same intensity at thermodynamic equilibrium. The concept of temperature is applicable only if the maximum entropy state has been reached [25]. While the definition of the monochromatic radiation temperature is similar to that of the radiance temperature, the physical significance is somewhat different. In the definition of radiance temperature, the concepts of entropy and thermodynamic equilibrium do not enter into consideration.

Consider a gray-diffuse body, for which the emissive power is proportional to the blackbody emissive power, at any frequency and angle of emission. The monochromatic temperature calculated from Eq. (8.56), however, is frequency dependent. This is because the emitted radiation, as a whole, cannot be considered as a blackbody at any temperature. Thermal radiation of this type has been called *dilute blackbody radiation* [23]. This simple example shows that photons at any given frequency can be considered as in a thermodynamic equilibrium but not necessarily in equilibrium with photons at other frequencies. When radiation has two linear polarizations with different intensities, the monochromatic temperatures will be different, even for the two polarizations. In general, it is a function of frequency, direction, and polarization. The requirement is that each subsystem be in a thermodynamic equilibrium, even though it is not in equilibrium with other subsystems at the same spatial location. Photons at different frequencies, with different polarization states, or propagating in different directions, can coexist in their own equilibrium states without any interaction with each other. The concept may be called partial equilibrium, as in the case when the two parts of a cylinder were separated by a moveable adiabatic wall. The mechanical equilibrium would be established to maintain the same pressure on each side, but the temperatures may be different from each other because thermal equilibrium is reached only inside each portion but not between them. Another example is in ultrafast laser heating of metals, as discussed in Chap. 7, where the electron

and phonon systems can be treated as being in separate equilibrium states but not in equilibrium with each other.

The concept of entropy intensity has recently been applied by Caldas and Semiao [27] to study the entropy generation in an absorbing, emitting, and scattering medium, based on the equation of radiative transfer (ERT) introduced in Sect. 2.4.3. The key is that the change in entropy in an elemental path length equals the change in intensity divided by the radiance temperature. The entropy change at steady state can be obtained from Eq. (2.53) in Chap. 2 as follows:

$$\frac{\mathrm{d}L_\lambda}{\mathrm{d}\xi} = \frac{a_\lambda I_{\mathrm{b},\lambda}}{T_\lambda(I_\lambda)} - \frac{(a_\lambda + \sigma_\lambda) I_\lambda}{T_\lambda(I_\lambda)} + \frac{\sigma_\lambda}{4\pi} \int_{4\pi} \frac{I_\lambda(\Omega')}{T_\lambda(I_\lambda)} \Phi(\Omega', \Omega) \mathrm{d}\Omega' \tag{8.57}$$

Like I_λ, the entropy intensity L_λ is a function of wavelength, location, and direction. Note that $I_{\mathrm{b},\lambda} = I_{\mathrm{b},\lambda}(\lambda, T_\mathrm{g})$, where T_g is the local temperature. Usually, $T_\lambda(I_\lambda)$ depends not only on the wavelength but also on the direction for a given location. The term $I_\lambda / T_\lambda(I_\lambda)$, however, is not the same as L_λ. Integrating Eq. (8.57) over the solid angle of 4π and over all wavelengths yields the entropy change in the volume element due to the intensity field variation. Furthermore, the entropy change in the control volume is equal to the total energy absorbed divided by T_g. The energy rate received per unit volume can be expressed as

$$\dot{q} = \int_0^\infty \int_{4\pi} a_\lambda \left(I_\lambda - I_{\mathrm{b},\lambda}\right) \mathrm{d}\Omega \mathrm{d}\lambda \tag{8.58}$$

The above equation works even with scattering since the integration of in-scattering and out-scattering cancels out. The rate of entropy change of the medium due to the net absorption by the matter is simply $\dot{q} / T_g$, which may be either positive or negative. The sum of the entropy change due to the field and that due to the matter is the total entropy change that is attributed to entropy generation by irreversibility. Therefore, we can express the volumetric entropy generation rate in terms of the absorption, emission, and scattering as follows [27]:

$$\dot{s}_\mathrm{gen} = \dot{s}_{\mathrm{abs + emi}} + \dot{s}_\mathrm{sca} \tag{8.59}$$

The entropy generation due to combined absorption and emission is

$$\dot{s}_{\mathrm{abs\text{ - }emi}} = \int_0^\infty \int_{4\pi} a_\lambda \left(I_\lambda - I_{b,\lambda}\right) \left[\frac{1}{T_\mathrm{g}} - \frac{1}{T_\lambda(I_\lambda)}\right] \mathrm{d}\Omega \mathrm{d}\lambda \tag{8.60a}$$

The entropy generation due to scattering is

$$\dot{s}_{\text{sca}} = \int_0^{\infty} \int_{4\pi} \left[\int_{4\pi} \frac{I_\lambda(\Omega')}{4\pi T_\lambda(I_\lambda)} \Phi(\Omega', \Omega) \mathrm{d}\Omega' - I_\lambda(\Omega) \right] \sigma_\lambda \mathrm{d}\Omega \mathrm{d}\lambda \tag{8.60b}$$

Note that Eq. (8.60a) should be treated as the combined absorption and emission effect since entropy generation due to absorption and emission processes cannot be separated. Furthermore, the entropy generation due to either absorption–emission or scattering is always greater than or equal to zero. When a surface is involved in radiative heat transfer, the entropy generation rate per unit area can be expressed as

$$s''_{\text{gen}} = \int_0^{\infty} \int_0^{2\pi} \int_0^{\pi/2} \left[\frac{I_{\text{in},\lambda} - I_{\text{out},\lambda}}{T_{\text{w}}} - (L_{\text{in},\lambda} - L_{\text{out},\lambda}) \right] \cos\theta \sin\theta \mathrm{d}\theta \mathrm{d}\phi \mathrm{d}\lambda \tag{8.61}$$

where T_{w} is the wall temperature, and subscripts "in" and "out" signify the energy or entropy intensity to and from the surface, respectively. If the surface is not a blackbody, the outgoing intensity includes both the emitted and reflected intensities. An alternative approach is to integrate the intensity over the whole sphere with a solid angle of 4π. In Eq. (8.61), the entropy intensity is related to the energy intensity by Eq. (8.55), which is recast in terms of wavelength as follows:

$$L_\lambda(\lambda, I_\lambda) = \frac{2k_{\text{B}}c}{\lambda^4} \left[\left(1 + \frac{\lambda^5 I_\lambda}{2hc^2} \right) \ln\left(1 + \frac{\lambda^5 I_\lambda}{2hc^2} \right) - \frac{\lambda^5 I_\lambda}{2hc^2} \ln\left(\frac{\lambda^5 I_\lambda}{2hc^2} \right) \right] \tag{8.62}$$

The use of Eq. (8.62) may be disputed when multiple reflections occur. The intensity of the emitted radiation is less than that of the blackbody and is reduced by each reflection. The question still remains as to whether the blackbody intensity should be used to calculate the entropy or the actual intensity after each reflection or the combined intensity at any given location. An example is a system of two large parallel plates, separated by vacuum. One of the plates is at a temperature T_1 and is diffuse-gray with an emittance of 0.5. The other plate is insulated and is a perfect reflector (i.e., no emission). It is clear that a thermal equilibrium will be established in the cavity after a long time. Again, the separation distance is much larger than the thermal radiation wavelengths. The radiation leaving surface 1 includes the emitted rays, as well as the first-order and higher order reflected rays. An attempt to define the entropy of the emitted ray and each reflected ray will result in a total entropy intensity greater than the entropy intensity calculated based on the blackbody intensity $I_{\text{b},\lambda}(\lambda, T_1)$. Therefore, to apply the previous analysis in a consistent way and to obtain meaningful results, we must make the following hypotheses:

- The intensity at any given location is additive regardless of where it originates from, as long as it falls within the same solid angle and wavelength intervals. While this sounds obvious, it is untrue when interference effects become important. The resulting intensity is called the combined intensity.

- The monochromatic radiation temperature T_λ, defined in Eq. (8.56), is a function of the combined intensity and is, in general, dependent on the direction and wavelength. The effect of polarization is neglected to simplify the problem. Equation (8.56) must not be applied to each of the reflected or scattered rays. The physical significance is that all the photons, with the same wavevector and frequency, can be considered as a subsystem that is at thermodynamic equilibrium with the temperature $T_\lambda[I_\lambda(\lambda, \theta, \phi)]$.
- The entropy intensity is defined based on the combined intensity, according to Eq. (8.62). While entropy must be additive, the entropy of all individual rays must be calculated based on the monochromatic temperature of the combined intensity. Because the number of photons, intensity, and entropy are additive, the fraction of the entropy of each ray is the same as the ratio of the intensity of that ray to the combined intensity.

With the theories presented in this section, one should be able to perform a second law thermodynamic analysis for a given system, involving radiative transfer of energy. Zhang and Basu [28] investigated entropy flow and generation considering incoherent multiple reflections. Different approximations exist in analyzing the entropy of radiation. For example, the method of dilute blackbody radiation uses a dilution factor and defines an effective temperature for each wavelength [23]. When the process is very complicated, such an effective temperature cannot be easily defined and this definition cannot be applied to multiple reflections. Entropy generation is usually accompanied by the generation of heat, such as heating by friction, electrical resistor, chemical reaction, or the absorption of solar radiation. On the other hand, it appears that entropy generation can occur in radiation without the generation of heat, such as by scattering. The definition of inelastic scattering is based on the conservation of energy (wavelength) and momentum, which does not impose any constraints on the reversibility. Further research is much needed in order to better understand the nature of entropy of radiation and determine the ultimate efficiency of photovoltaic cells and other radiative processes, including laser cooling and trapping. Another area of possible application of radiation entropy is in nanoscale heat conduction using EPRT, as discussed in Chap. 7. The entropy concept may be extended to the phonon system by defining radiation entropy and entropy intensity of phonons. Bright and Zhang [29] extended the concept of radiation entropy in a participating medium to phonon radiation, providing a method to evaluate local entropy generation. The conventional formula for entropy generation in heat diffusion can be derived under the local-equilibrium assumption. Furthermore, the entropy generation mechanism during phonon transport is elucidated as due to the "absorption" of high-frequency phonons and "emission" of lower frequency phonons, arising from the actual phonon scattering processes [29]. There is a need to further develop photon entropy analysis for near-field thermal radiation considering both interferences and photon tunneling [24].

8.2.4 Limitations of Planck's Law

The concept that *a blackbody absorbs all radiant energy that is incident upon it* is purely from the geometric-optics point of view, in which light travels in a straight line and cannot interact with an object that does not intercept the light ray. Another example of the geometric-optics viewpoint is that the transmittance of an iris (open aperture) should be 1, meaning that *all radiation incident on the opening will go through but no radiation outside the opening can go through.* However, for an aperture whose diameter is comparable to the wavelength of the incident radiation, diffraction may become important. As a result, the transmittance could exceed 1 in some cases. Due to the diffraction effect, a particle that is sufficiently small compared to the wavelength will interact with the radiation field, according to the scattering and absorption cross sections, which can be greater than the projected surface area. In some cases, it is possible for the object to absorb more energy than the product of the radiant flux and the projection area. The absorptance can be greater than 1 and thus exceeds the limit set by a blackbody. When such an object is placed in an isothermal enclosure, the emitted energy will be greater than that from a blackbody having the same dimensions. This anomaly has been discussed in detail by Bohren and Huffman [9].

The energy density near the surface within a distance less than the wavelength can be much greater than that given by Eq. (8.41) and increases as the distance is further reduced. When two objects are placed at a distance much smaller than the characteristic wavelength of thermal radiation, i.e., in the near field, photon tunneling can occur and cause significant enhancement of the energy transfer. In recent years, there have been numerous studies of light transmission through small apertures, radiation heat transfer at nanometer distances, and light emission from nanostructures [30–33]. Recent studies have also demonstrated that radiation heat transfer can be greatly enhanced for micro/nanostructures even when they are separated by distances longer than the characteristic wavelength [34]. This is still an open field with many new developments and applications. We will study these phenomena and the underlying physics in the following two chapters.

8.3 Radiative Properties of Semi-infinite Media

The reflection and refraction (transmission) of a semi-infinite isotropic medium are studied based on Maxwell's equations using suitable boundary conditions at the interface between the incident and transmitting media. Only plane waves with different polarizations are considered. Total internal reflection and the associated Goos–Hänchen phase shift are also introduced. For real materials or interfaces, the bidirectional reflectance distribution function (BRDF) is often needed to fully describe the radiative properties. The emittance can be calculated based on the reflectance by introducing Kirchhoff's law.

8.3.1 Reflection and Refraction of a Plane Wave

Consider radiation incident from one medium to another at the interface or the boundary. The boundary that separates the media is assumed to be a smooth plane and extends to infinity. Each medium is homogeneous and isotropic such that there is no scattering within the medium. Therefore, the electric response can be characterized by the relative permittivity or dielectric function ε, and the magnetic response can be characterized by the relative permeability μ. For nonmagnetic materials, the refractive index is related to the dielectric function by $n = \sqrt{\varepsilon}$. Keep in mind that these quantities are, in general, complex and frequency dependent. The real and imaginary parts of the refractive index are often called the optical constants. In this section, we present the general formulation for both magnetic and nonmagnetic materials. For certain crystalline and amorphous solids, like quartz and glass, the refractive index is real in a wide spectral region and is the only parameter needed to fully characterize the optical response of the material. In such a case, the expression can be largely simplified and the results can be easily comprehended. The reduced results will also be presented because of their importance to numerous engineering problems.

The incident radiation is a monochromatic plane wave with an angular frequency ω. As shown in Fig. 8.7, the wavevector of the incident wave is $\mathbf{k}_1^+ = (k_{1x}, 0, k_{1z})$, and the surface normal defines the *plane of incidence*, which is the x-z-plane. The wavevectors of the reflected and transmitted waves must lie in the same plane. The angle of incidence θ_1 is the angle between the incident wavevector and the z-direction, i.e., $\sin\theta_1 = k_{1x}/k_1$ and $\cos\theta_1 = k_{1z}/k_1$, where $k_1^2 = k_{1x}^2 + k_{1z}^2 = \mu_1\varepsilon_1\omega^2/c_0^2$. It is common to study the reflection and the refraction for linearly polarized waves, with either the electric or magnetic field being parallel to the y-axis, because other polarizations can be decomposed into the two polarization components.

When the electric field is in the y-direction, as shown in Fig. 8.7a, the wave is called a transverse-electric (TE) wave or is said to be perpendicularly (s) polarized. The incident electric field can be expressed as follows by omitting the time-harmonic term of $e^{-i\omega t}$ hereafter:

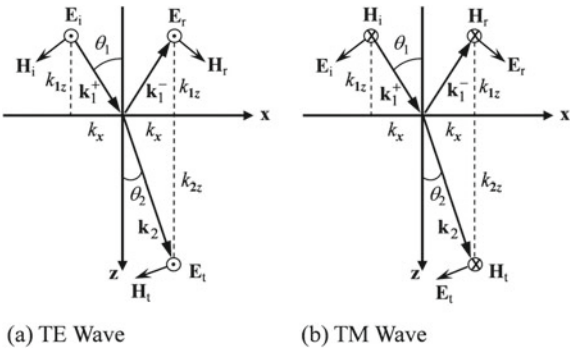

Fig. 8.7 Illustration of reflection and transmission at an interface: **a** TE wave or s-polarization. **b** TM wave or p-polarization

$$\mathbf{E}_\mathrm{i} = \hat{\mathbf{y}} E_\mathrm{i} \mathrm{e}^{\mathrm{i}k_{1z}z + \mathrm{i}k_{1x}x} \tag{8.63}$$

The boundary conditions state that the tangential components of both **E** and **H** must be continuous at the interface. This implies that the x-component of the wavevector must be the same for the incident, reflected, and transmitted waves, i.e., $k_{1x} = k_{2x} = k_x$. Because the angle of reflection must be the same as the angle of incidence (specular reflection), we have $\mathbf{k}_1^- = (k_x, 0, -k_{1z})$. For the transmitted or refracted wave, we have $\mathbf{k}_2 = (k_x, 0, k_{2z})$ and

$$\sin\theta_2 = \frac{k_x}{k_2} = \frac{n_1 \sin\theta_1}{n_2} \tag{8.64}$$

which is called Snell's law. It can be easily visualized by observing the bended image of a chopstick in a bowl of water. Note that $k_{2z}^2 = k_2^2 - k_x^2 = \mu_2\varepsilon_2\omega^2/c_0^2 - k_x^2 = k_2^2\cos^2\theta_2$. Generally speaking, the wavevector components and the refractive indices may be complex. Complex angles can be defined so that Eq. (8.64) is always valid. Near the interface, the nonzero components of the electric and magnetic fields are

$$E_y = \begin{cases} (E_\mathrm{i}\mathrm{e}^{\mathrm{i}k_{1z}z} + E_\mathrm{r}\mathrm{e}^{-\mathrm{i}k_{1z}z})\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z < 0 \\ E_\mathrm{t}\mathrm{e}^{\mathrm{i}k_{2z}z}\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z > 0 \end{cases} \tag{8.65}$$

$$H_x = \begin{cases} -\frac{k_{1z}}{\omega\mu_0\mu_1}(E_\mathrm{i}\mathrm{e}^{\mathrm{i}k_{1z}z} - E_\mathrm{r}\mathrm{e}^{-\mathrm{i}k_{1z}z})\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z < 0 \\ -\frac{k_{2z}}{\omega\mu_0\mu_2}E_\mathrm{t}\mathrm{e}^{\mathrm{i}k_{2z}z}\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z > 0 \end{cases} \tag{8.66}$$

and

$$H_z = \begin{cases} \frac{k_x}{\omega\mu_0\mu_1}(E_\mathrm{i}\mathrm{e}^{\mathrm{i}k_{1z}z} + E_\mathrm{r}\mathrm{e}^{-\mathrm{i}k_{1z}z})\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z < 0 \\ \frac{k_x}{\omega\mu_0\mu_2}E_\mathrm{t}\mathrm{e}^{\mathrm{i}k_{2z}z}\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z > 0 \end{cases} \tag{8.67}$$

where E_i, E_r, and E_t are, respectively, the amplitudes of the incident, reflected, and transmitted electric fields at the interface. It is further assumed that k_x is real so that the amplitude of the field is independent of x. The Fresnel reflection and transmission coefficients for a TE wave are defined as $r_{12,s} = E_\mathrm{r}/E_\mathrm{i}$ and $t_{12,s} = E_\mathrm{t}/E_\mathrm{i}$, respectively. Boundary conditions require that E_y and H_x be continuous at $z = 0$. From Eqs. (8.65) and (8.66), we obtain $1 + r_{12,s} = t_{12,s}$ and $(k_{1z}/\mu_1)(1 - r_{12,s}) = (k_{2z}/\mu_2)t_{12,s}$; thus,

$$r_{12,s} = \frac{E_\mathrm{r}}{E_\mathrm{i}} = \frac{k_{1z}/\mu_1 - k_{2z}/\mu_2}{k_{1z}/\mu_1 + k_{2z}/\mu_2} \tag{8.68a}$$

and

$$t_{12,s} = \frac{E_\mathrm{t}}{E_\mathrm{i}} = \frac{2k_{1z}/\mu_1}{k_{1z}/\mu_1 + k_{2z}/\mu_2} \tag{8.68b}$$

which are generally applicable, as long as each medium is homogeneous and isotropic [8]. For nonmagnetic materials, the previous equations can be written as follows:

$$r_{12,s} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2} \tag{8.69a}$$

and

$$t_{12,s} = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2} \tag{8.69b}$$

The spectral reflectivity, or simply reflectivity, ρ'_λ is given by the ratio of the reflected energy flux to the incident energy flux, and the absorptivity α'_λ is the ratio of the transmitted energy flux to the incident energy flux, since all the photons transmitted through the interface will be absorbed inside the second medium. Terms ending with "-ivity" are typically used for a perfect interface and those with "-tance" are for general surfaces including smooth and rough surfaces, thin films, as well as layered structures. The energy flux is related to the time-averaged Poynting vector, defined in Eq. (8.18b). From Eqs. (8.65) to (8.67), the x- and z-components of the Poynting vector at the interface ($z \to 0$) in medium 1 are

$$\langle S_{1x} \rangle = \frac{1}{2}\mathrm{Re}\left[\frac{k_x^*}{\omega\mu_0\mu_1^*}(E_\mathrm{i} + E_\mathrm{r})(E_\mathrm{i}^* + E_\mathrm{r}^*)\right] \tag{8.70a}$$

and

$$\langle S_{1z} \rangle = \frac{1}{2}\mathrm{Re}\left[\frac{k_{1z}^*}{\omega\mu_0\mu_1^*}(E_\mathrm{i} + E_\mathrm{r})(E_\mathrm{i}^* - E_\mathrm{r}^*)\right] \tag{8.70b}$$

It can be seen that, in general, the reflected wave and the incident wave are coupled and the energy flow cannot be separated into a reflected flux and an incident flux. Under the assumption that medium 1 is lossless (nonabsorbing or nondissipative) and $k_x^2 < k_1^2$, we can write

$$\langle S_{1z} \rangle = \langle S_{\mathrm{i}z} \rangle - \langle S_{\mathrm{r}z} \rangle \tag{8.71}$$

where

$$\langle S_{\mathrm{i}z} \rangle = \frac{k_{1z}}{2\omega\mu_0\mu_1}|E_\mathrm{i}|^2 \text{ and } \langle S_{\mathrm{r}z} \rangle = \frac{k_{1z}}{2\omega\mu_0\mu_1}|E_\mathrm{r}|^2 \tag{8.72}$$

If medium 1 is lossy, there will be additional terms associated with $E_\mathrm{i}E_\mathrm{r}^*$ and $E_\mathrm{i}^*E_\mathrm{r}$. In this case, the power flow normal to the interface cannot be separated as forward and backward terms, because of the cross-coupling terms. Therefore, the lossless condition in medium 1 is required in order to properly define the energy or power reflectivity [35]. This is usually not a problem when radiation is incident from air or a dielectric prism onto a medium. The spectral reflectivity can then be obtained based on the z-components of the reflected and incident Poynting vectors as

$$\rho'_{\lambda,s}(\theta_1) = |E_\mathrm{r}|^2 \big/ |E_\mathrm{i}|^2 = \left|r_{12,s}\right|^2 \tag{8.73}$$

The Poynting vector at the interface in medium 2 can be written as

$$\langle \mathbf{S}_\mathrm{t} \rangle = \frac{1}{2\omega\mu_0} \mathrm{Re}\left(\frac{k_x^* \hat{\mathbf{x}} + k_{2z}^* \hat{\mathbf{z}}}{\mu_2^*} \right) |E_\mathrm{t}|^2 \tag{8.74}$$

which is not parallel to $\mathrm{Re}(\mathbf{k}_2)$ unless $\mathrm{Im}(\mu_2) = 0$. Recall that the plane of constant phase is perpendicular to $\mathrm{Re}(\mathbf{k}_2)$. If medium 2 is dissipative, $\mathrm{Im}(\mathbf{k}_2)$ is parallel to the z-axis and the amplitude will vary along the z-direction. The wave becomes inhomogeneous in medium 2, except when $k_x = 0$ (normal incidence). The definition of the transmitted energy flux at the interface is based on the projected Poynting vector in the z-direction. Hence, the absorptivity is the ratio of the z-components of the transmitted and incident Poynting vectors, viz.,

$$\alpha'_{\lambda,s}(\theta_1) = \frac{\mathrm{Re}(k_{2z}/\mu_2)}{\mathrm{Re}(k_{1z}/\mu_1)} |t_{12,s}|^2 \tag{8.75}$$

Note that $\mathrm{Re}(k_{2z}/\mu_2) = \mathrm{Re}(k_{2z}^*/\mu_2^*)$, and $\mathrm{Re}(k_{1z}/\mu_1) = k_{1z}/\mu_1$ since medium 1 is lossless. It can be shown that $\rho'_{\lambda,s} + \alpha'_{\lambda,s} = 1$, as required by energy conservation: $\langle S_{1z} \rangle = \langle S_{2z} \rangle$ at $z = 0$. For nonmagnetic and nondissipative materials, we have

$$\alpha'_{\lambda,s}(\theta_1) = \frac{n_2 \cos\theta_2}{n_1 \cos\theta_1} |t_{12,s}|^2 \tag{8.76}$$

The reflection and transmission coefficients for the transverse magnetic (TM) wave or parallel (p) polarization are defined as the ratios of the magnetic fields: $r_{12,p} = H_\mathrm{r} / H_\mathrm{i}$ and $t_{12,p} = H_\mathrm{t} / H_\mathrm{i}$, respectively [7]. Hence,

$$r_{12,p} = \frac{H_\mathrm{r}}{H_\mathrm{i}} = \frac{k_{1z}/\varepsilon_1 - k_{2z}/\varepsilon_2}{k_{1z}/\varepsilon_1 + k_{2z}/\varepsilon_2} \tag{8.77a}$$

$$t_{12,p} = \frac{H_\mathrm{t}}{H_\mathrm{i}} = \frac{2k_{1z}/\varepsilon_1}{k_{1z}/\varepsilon_1 + k_{2z}/\varepsilon_2} \tag{8.77b}$$

In the case of nonmagnetic materials, we obtain

$$r_{12,p} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2} \tag{8.78a}$$

and

$$t_{12,p} = \frac{2n_2 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2} \tag{8.78b}$$

At normal incidence, the reflection coefficients calculated based on Eqs. (8.69a) and (8.78a) are related by

$$r_{12,s} = \frac{n_1 - n_2}{n_1 + n_2} = -r_{12,p} \tag{8.79}$$

When both n_1 and n_2 are real, for $n_1 < n_2$, the electric field will experience a phase reversal (phase shift of π) upon reflection but the magnetic field will not. On the other hand, for $n_1 > n_2$, it is the magnetic field that will experience a phase reversal. In fact, based on Maxwell's equations, the electric and magnetic quantities obey a duality when $\rho_e = 0$. They can be interchanged with the following substitutions: $\mathbf{E} \to \mathbf{H}$ and $\mathbf{H} \to -\mathbf{E}$. Note that ε and μ, as well as the polarization states s and p, should also be interchanged. The Poynting vector for a TM wave is $\langle \mathbf{S} \rangle = \frac{1}{2\omega\varepsilon_0} \mathrm{Re}\left(\frac{\mathbf{k}}{\varepsilon}\right) \left| H_y \right|^2$, which is not parallel to $\mathrm{Re}(\mathbf{k})$ when $\mathrm{Im}(\varepsilon_2) \neq 0$. Upon refraction into an absorbing medium, the waves become inhomogeneous and the Poynting vectors for different polarizations may split into different directions [36]. Nevertheless, the constant-amplitude plane is always perpendicular to the z-direction because the amplitude cannot change along the x-y-plane. The reflectivity for p-polarization is

$$\rho'_{\lambda,p}(\theta_1) = \left| r_{12,p} \right|^2 \tag{8.80}$$

Hence, the absorptivity becomes

$$\alpha'_{\lambda,p}(\theta_1) = \frac{\mathrm{Re}(k_{2z}/\varepsilon_2)}{\mathrm{Re}(k_{1z}/\varepsilon_1)} \left| t_{12,p} \right|^2 \tag{8.81}$$

For nonmagnetic and nonabsorbing materials, we have

$$\alpha'_{\lambda,p}(\theta_1) = \frac{n_1 \cos\theta_2}{n_2 \cos\theta_1} \left| t_{12,p} \right|^2 \tag{8.82}$$

If the incident wave is unpolarized or circularly polarized, the reflectivity can be obtained by averaging the values for p- and s-polarized waves, i.e.,

$$\rho'_{\lambda} = \frac{\rho'_{\lambda,p} + \rho'_{\lambda,s}}{2} \tag{8.83}$$

The reflectivity for radiation incident from air ($n_1 \approx 1$) to a dielectric medium ($n_2 = 2$) and that from the dielectric to air are shown in Fig. 8.8 for each polarization as well as for the unpolarized incident radiation. When $n_1 > n_2$, the reflectivity will reach 1 at $\theta_1 = \theta_c = \sin^{-1}(n_2/n_1)$. This angle is called the *critical angle*, and *total internal reflection* occurs at angles of incidence greater than the critical angle. This is the principle commonly used in optical fibers and waveguides, since light is trapped inside the high-index material and propagates along the medium. It can be seen that in total internal reflection, $k_x > k_2$ and k_{2z} becomes purely imaginary. The amplitude of the wave exponentially attenuates in the positive z-direction. This is similar to Eq. (8.37) and makes it an evanescent wave, as shown in Fig. 8.3. The time-averaged Poynting vector is zero in the z-direction. Hence, no energy is transmitted through the boundary.

For the TE wave, the reflectivity increases monotonically with the angle of incidence and reaches 1 at the grazing angle (90°) or at the critical angle when $n_1 > n_2$.

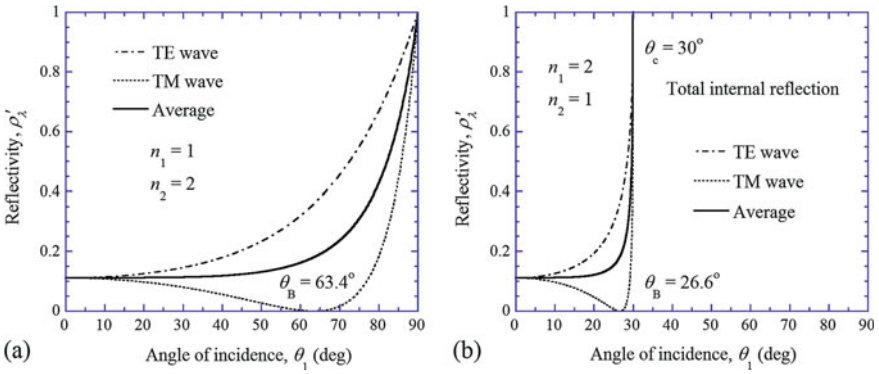

Fig. 8.8 Reflectivity versus the angle of incidence between air and a dielectric: **a** Incident from air to a medium; **b** Incident from a medium to air

The reflectivity for the TM wave, on the other hand, goes through a minimum that is equal to zero. The angle at which $\rho'_{\lambda,p} = 0$ is called the Brewster angle, given by $\theta_B = \tan^{-1}(n_2/n_1)$ for nonmagnetic materials. For p-polarization, all the incident energy will be transmitted into medium 2, without reflection at the Brewster angle. This phenomenon has been used to build polarizers and transmission windows in absolute cryogenic radiometers. The physical mechanism of reflection can also be understood as re-emission by the *induced electric dipoles* in the medium, based on the *Ewald–Oseen extinction theorem*. At the Brewster angle, the electric dipoles induced in the material align in the direction of the reflected wave, and the refracted wave is perpendicular to the reflected wave (i.e., $\theta_1 + \theta_2 = 90°$). The reflective power goes to zero because an electric dipole cannot radiate along its own axis. The situation is changed when magnetic materials are involved, such as a negative index material. The fields radiated by both the induced electric dipoles and *magnetic dipoles* are responsible for the reflection. The Brewster angle can occur for either polarization when the radiated fields cancel each other. A detailed discussion can be found from the publication of Fu et al. [37]. In an absorbing medium, there is a drop in reflectivity for p-polarization, but the minimum is not zero. Furthermore, there exists a *principal angle* at which the phase difference between the two reflection coefficients equals 90° and the ratio of the reflectivity for the TM and TE waves is near the reflectivity minimum [8]; see Problem 8.24.

The reflectivity for radiation incident from air ($n_1 \approx 1$) or a vacuum, at normal incidence, becomes

$$\rho'_{\lambda,\mathrm{n}} = \frac{(n_2 - 1)^2 + \kappa_2^2}{(n_2 + 1)^2 + \kappa_2^2} \tag{8.84}$$

for any polarization. It can be seen that the normal reflectivity will be close to 1, when either $n_2 \ll 1$ or $n_2 \gg 1$. The reflectivity is large for most metals in the infrared, because both n_2 and κ_2 are large. The reflectivity of a conventional superconductor

approaches 1 when the frequency is lower than that of the superconducting energy gap, since $n_2 \to 0$ in this case. On the other hand, $\rho'_{\lambda,\mathrm{n}} \to 0$ when $n_2 \approx 1$ and $\kappa_2 \ll 1$. This can occur in a dielectric material at certain mid-infrared wavelengths and also for most metals in the x-ray region.

8.3.2 *Total Internal Reflection and the Goos–Hänchen Shift*

Total internal reflection (TIR) occurs when light comes from an optically denser material to another material at incidence angles greater than the critical angle determined by Snell's law. As discussed in the preceding section, the amplitude of the reflection coefficient becomes unity at incidence angles greater than the critical angle. Although no energy is transferred from medium 1 to medium 2, there exists an electromagnetic field in the second medium near the surface. This electromagnetic field can store as well as exchange energy with medium 1 at any instant of time.

While evanescent waves do not carry energy into the second medium, there is a shift in the phase of the reflected wave upon TIR. Consider a plane wave of angular frequency ω incident from a semi-infinite medium 1 to medium 2, as shown in Fig. 8.9a. The wavevector $\mathbf{k}_1^+ = k_x\hat{\mathbf{x}} + k_{1z}\hat{\mathbf{z}}$, $\mathbf{k}_1^- = k_x\hat{\mathbf{x}} - k_{1z}\hat{\mathbf{z}}$, and $\mathbf{k}_2 = k_x\hat{\mathbf{x}} + k_{2z}\hat{\mathbf{z}}$, since the parallel wavevector component k_x must be the same as required by the phase-matching boundary condition. The magnitudes of the wavevectors are

$$k_1^2 = k_x^2 + k_{1z}^2 = \varepsilon_1\mu_1\omega^2/c^2 \tag{8.85a}$$

and

$$k_2^2 = k_x^2 + k_{2z}^2 = \varepsilon_2\mu_2\omega^2/c^2 \tag{8.85b}$$

where ε and μ are the relative (ratio to those of a vacuum) permittivity and permeability, respectively, and c is the speed of light in a vacuum (omitting the subscript 0). Assume that the incident wave is p polarized or a TM wave, so that the only nonzero component of the magnetic field is in the y-direction. The magnetic field of the incident wave may be expressed as $\mathbf{H}_\mathrm{i} = (0, H_y, 0)$, where

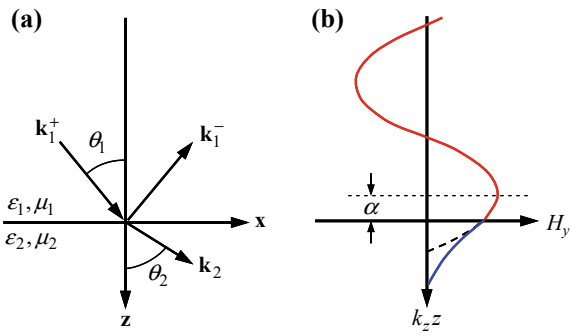

Fig. 8.9 Illustration of total internal reflection. **a** Schematic of the incident, reflected, and transmitted waves at the interface between two semi-infinite media. **b** The magnetic field distribution for a TM wave when total internal reflection occurs

$H_y(x, y, z, t) = H_\mathrm{i} \mathrm{e}^{\mathrm{i}k_{1z}z + \mathrm{i}k_{1x}x - \mathrm{i}\omega t}$. For simplicity, let us omit $\exp(-\mathrm{i}\omega t)$ from now on. Recall that the Fresnel coefficients for a TM wave are defined as the ratios of the reflected or transmitted magnetic field to the incident magnetic field. For example, the Fresnel reflection coefficient is

$$r_\mathrm{p} = \frac{H_\mathrm{r}}{H_\mathrm{i}} = \frac{k_{1z}/\varepsilon_1 - k_{2z}/\varepsilon_2}{k_{1z}/\varepsilon_1 + k_{2z}/\varepsilon_2} \tag{8.86}$$

The field in medium 1 is composed of the incident and reflected fields, and that in medium 2 is the transmitted field. Therefore,

$$\frac{H_y}{H_\mathrm{i}} = \begin{cases} (\mathrm{e}^{\mathrm{i}k_{1z}z} + r_p \mathrm{e}^{-\mathrm{i}k_{1z}z})\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z \le 0 \\ (1 + r_p)\mathrm{e}^{\mathrm{i}k_{2z}z}\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z > 0 \end{cases} \tag{8.87}$$

The electric fields can be obtained by applying the Maxwell equations. Hence, we can write the electric and magnetic fields in both media as follows:

$$\frac{E_x}{H_\mathrm{i}} = \begin{cases} \frac{k_{1z}}{\omega\varepsilon_1\varepsilon_0}(\mathrm{e}^{\mathrm{i}k_{1z}z} - r_p \mathrm{e}^{-\mathrm{i}k_{1z}z})\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z \le 0 \\ \frac{k_{2z}}{\omega\varepsilon_2\varepsilon_0}(1 + r_p)\mathrm{e}^{\mathrm{i}k_{2z}z}\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z > 0 \end{cases} \tag{8.88}$$

and

$$\frac{E_z}{H_\mathrm{i}} = \begin{cases} -\frac{k_x}{\omega\varepsilon_1\varepsilon_0}(\mathrm{e}^{\mathrm{i}k_{1z}z} + r_p \mathrm{e}^{-\mathrm{i}k_{1z}z})\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z \le 0 \\ -\frac{k_x}{\omega\varepsilon_2\varepsilon_0}(1 + r_p)\mathrm{e}^{\mathrm{i}k_{2z}z}\mathrm{e}^{\mathrm{i}k_x x}, & \text{for } z > 0 \end{cases} \tag{8.89}$$

Assume that ε's and μ's are real and furthermore, $\varepsilon_1\mu_1 > \varepsilon_2\mu_2 > 0$. From Eq. (8.85b), we have $k_{2z}^2 = \varepsilon_2\mu_2\omega^2/c^2 - k_x^2$. When $\sqrt{\varepsilon_2\mu_2} < k_x c/\omega < \sqrt{\varepsilon_1\mu_1}$, the incidence angle θ_1 is defined but the refraction angle is not, because k_{2z} becomes imaginary. One can write $k_{2z} = \mathrm{i}\eta_2$, where $\eta_2 = \sqrt{k_x^2 - \varepsilon_2\mu_2\omega^2/c^2}$ is a real positive number. In this case, $|r_p| = 1$ and

$$r_p = \mathrm{e}^{\mathrm{i}\delta} = \mathrm{e}^{-\mathrm{i}2\alpha} \tag{8.90}$$

where $\tan\alpha = (\eta_2/\varepsilon_2)/(k_{1z}/\varepsilon_1)$. Following Haus [38], the magnetic field at $x = 0$ in medium 1 can be written as

$$H_y = 2H_\mathrm{i}\mathrm{e}^{-\mathrm{i}\alpha}\cos(k_{1z}z + \alpha), \quad z \le 0 \tag{8.91a}$$

Similarly, H_y in medium 2 becomes

$$H_y = 2H_\mathrm{i}\mathrm{e}^{-\mathrm{i}\alpha}\cos(\alpha)\mathrm{e}^{-\eta_2 z}, \quad z > 0 \tag{8.91b}$$

The magnetic field at $x = 0$ is plotted in Fig. 8.9b with respect to $k_z z$, at the instant of time when the phase of $H_\mathrm{i}\mathrm{e}^{-\mathrm{i}\alpha - \mathrm{i}\omega t}$ becomes zero. From this figure, one can see that the field decays exponentially in medium 2. As a result, there is a phase shift in medium 1 upon TIR so that the maximum amplitude is shifted from the interface

to $k_z z = -\alpha$. The phase angle of the reflection coefficient $\delta = -2\alpha$ is called the *Goos–Hänchen phase shift,* which depends on the incidence angle θ_1 or k_x. The difference in δ for TE and TM waves in a dielectric prism was used to construct a polarizer called *Fresnel's rhomb*, which can change a linearly polarized wave into a circularly polarized wave, or vice versa [7].

Example 8.6 Calculate the time-averaged Poynting vector near the interface in the case of total internal reflection.

Solution Based on Example 8.1, it can be seen that the Poynting vector $\mathbf{S} = \mathrm{Re}(\mathbf{E}) \times \mathrm{Re}(\mathbf{H})$ is, in general, a function of time. The time-dependent terms that oscillate with 2ω, however, become zero after integration. The time-averaged Poynting vector is $\langle \mathbf{S} \rangle = \frac{1}{2}\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*)$. For z > 0, $\langle S_z \rangle = \frac{1}{2}\mathrm{Re}(E_x H_y^*) = 0$ because k_{2z} is purely imaginary. It can also be shown that $\langle S_z \rangle = 0$ for $z \leq 0$ (see Problem 8.26). Furthermore,

$$\langle S_x \rangle = -\tfrac{1}{2}\mathrm{Re}(E_z H_y^*) = \begin{cases} \frac{k_x}{\omega \varepsilon_1 \varepsilon_0} |H_i|^2 \big[1 + \cos(2k_{1z} z + 2\alpha)\big], & z \leq 0 \\ \frac{k_x}{\omega \varepsilon_2 \varepsilon_0} |H_i|^2 [1 + \cos(2\alpha)] \mathrm{e}^{-2\eta_2 z}, & z > 0 \end{cases} \tag{8.92}$$

Note that $\langle S_x \rangle$ does not have to be continuous across the interface. Depending on whether ε is positive or negative, the sign of $\langle S_x \rangle$ may be parallel or antiparallel to k_x. It should also be noted that $\langle S_x \rangle$ is a sinusoidal function of z in medium 1 and decays exponentially in medium 2 as z approaches infinity.

Newton conjectured that, when a light beam is reflected at the boundary upon TIR, the light corpuscles would penetrate some distance into the optically rarer medium and then reenter the optically denser medium. In addition, he suspected that the path of the beam would be a parabola with its vertex in the rarer medium and, consequently, the actual reflected beam would be shifted laterally with respect to the geometric-optics prediction. From the Poynting vector formulation given in Eq. (8.92), the energy must penetrate into the second medium to maintain the energy flow parallel to the interface and reenter the first medium so that no net energy is transferred across the interface. The actual beams have a finite extension so that the reflected beam in the far field can be separated from the incident beam since the Poynting vector is parallel to the wavevector. The effect of the parallel energy flow indeed causes the reflected beam to shift forward from that expected by the geometric-optics analysis. F. Goos and H. Hänchen were the first to observe the lateral beam shift through a cleverly devised experiment in 1947. A schematic of this experiment is shown in Fig. 8.10, in which a glass plate was used so that the incident light was multiply reflected by the top and bottom surfaces. In the middle of one or both of the surfaces, a silver strip was deposited. This way, the beam reflected by the silver film (solid line) would essentially follow geometric optics and that by total internal reflection would experience a lateral shift. Although the lateral shift is on the order of the wavelength, a large number of reflections (over 100 times) allowed the shift to be observed by a photographic plate. Lotsch [39] published a series of papers on the comprehensive study of the Goos–Hänchen effect. Puri and Birman [40] provided an elegant review of earlier works, including several methods for analyzing the Goos–Hänchen effect. A quantitative study of the Goos–Hänchen effect is presented next.

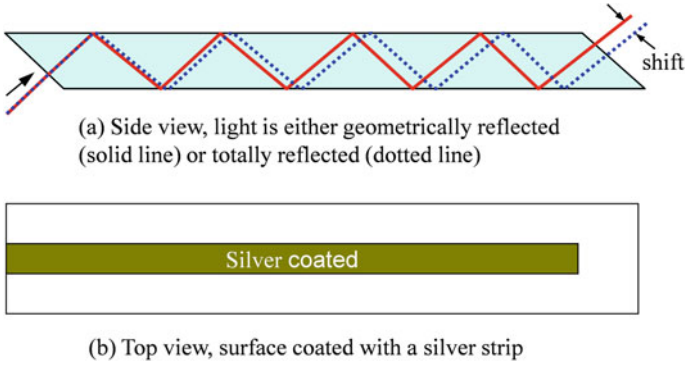

Fig. 8.10 Illustration of the Goos–Hänchen experiment: **a** side view, **b** top view

One way to model the lateral shift is to use a beam of finite width rather than an unbounded plane wave. Another method that is mathematically simpler considers the phase change of an incoming wave packet, which is composed of two plane waves with a slightly different k_x. Upon TIR, the phase shift $\delta = -2\alpha$ for a given polarization is a function of k_x. The difference in the phase shift will cause the reflected beam to exhibit a lateral shift along the interface (x-direction) given as

$$D = -\frac{\mathrm{d}\delta}{\mathrm{d}k_x} = \frac{\varepsilon_1}{\varepsilon_2}\frac{2k_x}{\eta_2 k_{1z}}\frac{k_{1z}^2 + \eta_2^2}{k_{1z}^2 + (\eta_2\varepsilon_1/\varepsilon_2)^2}, \text{ for } p \text{ polarization} \tag{8.93}$$

where we have used $\alpha = \tan^{-1}(\eta_2\varepsilon_1/k_{1z}\varepsilon_2)$. In formulating the above equation, k_x is always taken as positive. Equation (8.92) clearly suggests that $\langle S_x\rangle$ and k_x have the same sign when the permittivity is positive and different signs when the permittivity is negative [41]. When ε_1 and ε_2 have different signs, the lateral shift D will be negative, which implies that the lateral shift is opposite to $\langle \mathbf{S}\rangle_x$ of the incident beam. For a TE wave, one can simply replace ε's by μ's in Eqs. (8.93). For two dielectrics, we have $\mu_1 = \mu_2 = 1$, $\varepsilon_1 = n_1^2$, and $\varepsilon_2 = n_2^2$, where n_1 and n_2 are the refractive indices of medium 1 and 2, respectively. Consequently, Eq. (8.93) reduces to the following:

$$D_s = \frac{2\tan\theta_1}{\eta_2} \text{ for a TE wave} \tag{8.94a}$$

and

$$D_p = \frac{2\tan\theta_1}{\eta_2\left(n_1^2\sin^2\theta_1/n_2^2 - \cos^2\theta_1\right)} \text{ for a TM wave} \tag{8.94b}$$

At grazing incidence, $k_{1z} \to 0$, however, the shift in the direction parallel to the beam is $D\cos\theta_1 = (2/\eta_2)(\varepsilon_2/\varepsilon_1)\sin\theta_1$, which approaches a finite value and does not diverge. At the critical angle, $\theta_1 = \theta_c = \sin^{-1}(n_1/n_1)$, $\eta_2 = \delta = 0$,

and D approaches infinity. This difficulty can be removed by using the Gaussian beam incidence [42]. Quantum mechanics has also been applied to predict the lateral beam shift [39]. The Goos–Hänchen effect also has its analogy in acoustics and is of contemporary interest in dealing with negative index materials, waveguides, and photon tunneling [41, 43, 44].

8.3.3 *Bidirectional Reflectance Distribution Function*

Real surfaces contain roughness or texture that depends on the processing method. A surface appears to be smooth if the wavelength is much greater than the surface roughness height. A highly polished surface can have a roughness height on the order of nanometers. Some surfaces that appear "rough" to human eyes may appear to be quite "smooth" for far-infrared radiation. The root-mean-square (rms) roughness is a commonly used parameter to describe surface roughness. The power spectral density provides more general information on the vertical and spatial extent of surface irregularities. Zhang et al. [31] gave a detailed discussion on the roughness parameters as well as the instruments used for surface characterization.

The reflection of radiation by rough surfaces is more complicated. For randomly rough surfaces, there often exist a peak around the direction of specular reflection, an off-specular lobe, and a diffuse component. When the surface contains periodic structures, such as patterned or microfabricated surfaces, diffraction effects may become important and several peaks may appear. The bidirectional reflectance distribution function (BRDF), which is a function of the angles of incidence and reflection, fully describes the reflection characteristics from a rough surface at a given wavelength. As illustrated in Fig. 8.11, the BRDF is defined as the reflected radiance (intensity) divided by the incident irradiance (flux) at the surface [45]

$$f_{\mathrm{r}}(\lambda, \theta_i, \phi_{\mathrm{i}}, \theta_r, \phi_{\mathrm{r}}) = \frac{\mathrm{d}I_{\mathrm{r}}}{I_{\mathrm{i}} \cos\theta_i \mathrm{d}\Omega_i} \ [\mathrm{sr}^{-1}] \tag{8.95}$$

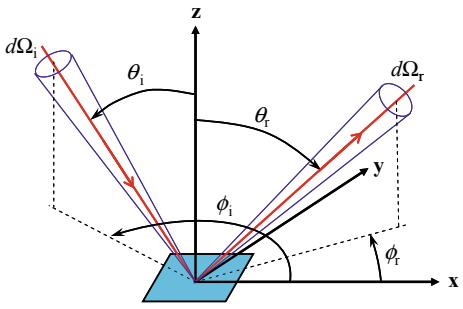

Fig. 8.11 Geometry of the incident and reflected beams in defining the BRDF

where (θ_i, ϕ_i) and (θ_r, ϕ_r) denote the directions of incident and reflected beams, respectively, I_i is the incident irradiance (radiant power per unit area), and dI_r is the reflected radiance (intensity).

The directional-hemispherical reflectance can be obtained by integrating the BRDF over the hemisphere:

$$\rho'_\lambda = \int_{2\pi} f_r \cos\theta_r \, d\Omega_r \tag{8.96}$$

An important principle of the BRDF is reciprocity, which specifies symmetry of the BRDF, with regard to reflection and incidence angles. In other words, the reflectance for energy incident from (θ_i, ϕ_i) and reflected to (θ_r, ϕ_r) is equal to that for energy incident from (θ_r, ϕ_r) and reflected to (θ_i, ϕ_i). Therefore,

$$f_r(\lambda, \theta_i, \phi_i, \theta_r, \phi_r) = f_r(\lambda, \theta_r, \phi_r, \theta_i, \phi_i) \tag{8.97}$$

The BRDF reciprocity is an extension of the Helmholtz reciprocity principle [46]. While the reciprocity principle holds for most passive medium and surfaces, it does not hold in some nonlinear or magnetic media.

For a *diffuse* or *Lambertian* surface, the BRDF is independent of (θ_r, ϕ_r) and is related to the directional-hemispherical reflectance as $f_{r,dif} = \rho'_\lambda/\pi$. On the other hand, the BRDF for an ideal *specular,* or mirror-like, reflector can be represented as

$$f_{r,spe} = \frac{\rho'_\lambda}{\cos\theta_i}\delta_\theta(\theta_r - \theta_i)\delta_\phi(\phi_r - \phi_i - \pi) \tag{8.98}$$

where the Dirac delta function $\delta(x)$ is zero everywhere, except at $x = 0$. Furthermore, the delta functions are normalized such that

$$\int_{2\pi} \delta_\theta(\theta_r - \theta_i)\delta_\phi(\phi_r - \phi_i - \pi) d\Omega_r = 1 \tag{8.99}$$

In general, the BRDF of a real surface should fall between the two extreme cases. It should be noted that for a perfectly smooth surface, the reflectivity calculated from the Fresnel coefficient, discussed in Sect. 8.3.1, can also be understood as the directional-hemispherical reflectance. Further discussions on BRDF models based on geometric optics and physical optics, as well as rigorous solutions of the Maxwell equations, will be given in Chap. 9, where we will also study the effect of surface microstructures on the BRDF and how to characterize a rough surface.

8.3.4 Emittance (Emissivity) and Kirchhoff's Law

Real materials have finite thicknesses. The assumption of semi-infinity or opaqueness requires that the thickness be much greater than the radiation penetration depth. This is usually not a problem for a metal in the visible or infrared spectral regions. When this is not the case, we are dealing with a transparent or semitransparent material, like a glass window. The radiative properties of semitransparent layers and thin films will be studied in the next chapter. Laser beams or light from a spectrophotometer do not extend to infinity and are not perfectly collimated. Nevertheless, as long as the diameter of the beam spot is much greater than the wavelength and the beam divergence is not very large, the directional, spectral reflectivity and absorptivity, calculated from the previous section, are applicable to most situations and can be integrated to obtain the properties for finite conic angles or hemispherical properties. For real materials, we use reflectance and absorptance that depend on the nature of surfaces and coatings.

For real surfaces, the ratio of the *emissive power* of the surface to that of a blackbody at the same temperature defines the hemispherical emittance (or emissivity) $\varepsilon_\lambda^{\mathrm{h}}$. The directional emittance (or emissivity) ε'_λ is defined based on the intensity ratio. The total emittance can be evaluated by integrating the spectral emittance over all wavelengths weighted by the blackbody distribution function. A concise discussion of radiative properties can be found from a popular heat transfer textbook [47] and more complete definitions and relations can be found from Howell et al. [1] and Modest [2].

To establish the relationship between the radiative properties, consider an opaque surface at a temperature T_s inside a vacuum enclosure whose walls are at a temperature T_w, as shown in Fig. 8.12. Whether the surface is inside the enclosure or not, we must have

$$\alpha'_\lambda(\theta,\phi) + \rho'_\lambda(\theta,\phi) = 1 \tag{8.100a}$$

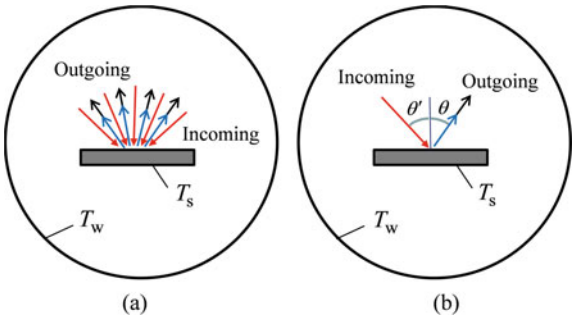

Fig. 8.12 Schematic of a blackbody enclosure for consideration of **a** hemispherical properties and **b** directional properties. Note that the outgoing radiation has two arrows, one represents emission and the other reflection

Because the incoming radiation will either be absorbed or reflected as long as the surface is opaque, the sum of the directional absorptance and the directional-hemispherical reflectance must be unity. Furthermore, the sum of the hemispherical absorptance and the hemispherical–hemispherical reflectance must also be unity:

$$\alpha_{\lambda}^{\mathrm{h}} + \rho_{\lambda}^{\mathrm{h}} = 1 \text{ and } \alpha_{\mathrm{tot}}^{\mathrm{h}} + \rho_{\mathrm{tot}}^{\mathrm{h}} = 1 \tag{8.100b}$$

where the subscript "tot" signifies a total property.

Consider the enclosure at thermal equilibrium where no internal sources or sinks exist in the surface element. The temperatures of the surface element and the wall must be the same. Furthermore, as shown earlier with the blackbody cavity concept, the intensity and emissive power inside the enclosure are independent on the location and direction. Based on Fig. 8.12a for the hemispherical properties, we see that the combination of the emitted and reflected power per unit area must be the same as that of the incident. Energy balance requires that

$$e_{\mathrm{b},\lambda}(\lambda, T_{\mathrm{w}}) = \varepsilon_{\lambda}^{\mathrm{h}} e_{\mathrm{b},\lambda}(\lambda, T_{\mathrm{s}}) + \rho_{\lambda}^{\mathrm{h}} e_{\mathrm{b},\lambda}(\lambda, T_{\mathrm{w}}) \tag{8.101a}$$

and

$$\sigma_{\mathrm{SB}} T_{\mathrm{w}}^4 = \varepsilon_{\mathrm{tot}}^{\mathrm{h}} \sigma_{\mathrm{SB}} T_{\mathrm{s}}^4 + \rho_{\mathrm{tot}}^{\mathrm{h}} \sigma_{\mathrm{SB}} T_{\mathrm{w}}^4 \tag{8.101b}$$

By setting $T_{\mathrm{s}} = T_{\mathrm{w}}$ and combining with Eq. (8.100b), we have

$$\varepsilon_{\lambda}^{\mathrm{h}} = \alpha_{\lambda}^{\mathrm{h}} \text{ and } \varepsilon_{\mathrm{tot}}^{\mathrm{h}} = \alpha_{\mathrm{tot}}^{\mathrm{h}} \tag{8.102}$$

The equality between emittance and absorptance is called Kirchhoff's law. Note that in the literature, both emissivity and absorptivity are commonly used regardless of the nature of the surfaces. The hemispherical properties depend on the directional and spectral behavior of the surface. Furthermore, the hemispherical absorptance depends on the condition of the incident intensity distribution. Hence, the two equalities given in Eq. (8.102) do not hold in general. Special situations exist under ideal assumptions, for example, if a surface is diffuse, $\varepsilon_{\lambda}^{\mathrm{h}} = \alpha_{\lambda}^{\mathrm{h}}$ always holds. Furthermore, if a surface is diffuse-gray, both the equalities in Eq. (8.102) hold without requiring thermal equilibrium [47]. Real surfaces rarely meet these requirements, however.

Equation (8.102) can also be understood by considering the energy balance of the surface, that is, the absorbed radiant power must equal the emitted so that the emittance must be equal to the absorptance. This argument is justifiable for the hemispherical properties. When dealing with directional properties, as can be seen from Fig. 8.12b, the emitted and the reflected components toward the same direction θ should add up to give the blackbody intensity. Under thermal equilibrium, we can obtain the following expression [46]:

$$\varepsilon_{\lambda}'(\lambda, \theta, \phi) + \int_{2\pi} f_{\mathrm{r}}(\lambda, \theta', \phi', \theta, \phi) \cos\theta' \mathrm{d}\Omega' = 1 \tag{8.103a}$$

Here, the integration is over the incident hemisphere. The second term in Eq. (8.103a) is the hemispherical-directional reflectance [1, 2]. When BRDF reciprocity holds, it is the same as the directional-hemispherical reflectance ρ'_λ. For a specular surface, Eq. (8.103a) can be written as

$$\varepsilon'_\lambda(\lambda, \theta, \phi) + \rho'_\lambda(\lambda, \theta', \phi') = 1 \tag{8.103b}$$

where (θ', ϕ') and (θ, ϕ) are a pair of specular incidence and reflection angles, i.e., $\theta' = \theta$ and $\phi' = \phi + 180°$. Zhu and Fan [48] showed that Eq. (8.103b) holds even in nonreciprocal systems such as magneto-optical materials. When BRDF reciprocity holds, the conventional Kirchhoff's law $\varepsilon'_\lambda = \alpha'_\lambda$ for the spectral directional properties can be derived. Therefore, the spectral, directional emittance can be expressed in terms of the spectral, directional-hemispherical reflectance as follows:

$$\varepsilon'_\lambda = 1 - \rho'_\lambda \tag{8.104}$$

When a material is not at thermal equilibrium with its surroundings, its emittance is defined based solely on spontaneous emission and is an intrinsic property of the material that does not depend on the surroundings. On the other hand, the absorptance is defined based on the net absorbed energy by treating stimulated or induced emission as negative absorption. Under appropriate conditions, Kirchhoff's law according to the equality given in Eq. (8.104) is valid for individual polarization with or without thermal equilibrium. The assumptions are: (a) the material under consideration is reciprocal and at local thermal equilibrium, though not necessarily at equilibrium with the surroundings; (b) the external field is not strong enough to alter the material's intrinsic properties or cause a nonlinear effect. We can then compute the directional emittance for an opaque surface or semi-infinite media, from the directional-hemispherical reflectance for incidence from air or a vacuum.

The emittance is typically calculated by averaging over the two polarizations. The hemispherical emittance can then be obtained by integrating the directional emittance so that

$$\varepsilon^{\mathrm{h}}_\lambda = \frac{1}{\pi} \int_0^{2\pi} \int_0^{\pi/2} \varepsilon'_\lambda \cos\theta \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi \tag{8.105}$$

It can be seen from Fig. 8.8a that, when averaged over the two polarizations, the reflectivity changes little until the Brewster angle and then increases to 1 when the incidence angle approaches 90°. The hemispherical emittance for a nonmetallic surface is about 10% smaller than the normal emittance. On the other hand, the hemispherical emittance for metallic surfaces is about 20% greater than the normal emittance. Diffuse emission is a good first-order approximation, even though the surface is smooth and the reflection is specular. Thus, the hemispherical emittance may be approximated by the normal emittance. In most studies, the emittance is calculated from the indirect method, based on the reflectivity and Kirchhoff's law

given in Eq. (8.104). Direct calculations can be accomplished by considering the emission, along with the absorption and transmission, inside the material. According to the fluctuation-dissipation theorem (FDT), thermal emission arises from the induced field that originated from the random charge fluctuation. Wang et al. [49] used the FDT to directly calculate the emittance of a layered structure and demonstrated the equivalence between the direct method and the indirect method based on Eq. (8.104). The fluctuational electrodynamics is essential for the study of near-field radiation and will be carefully discussed in Chap. 10.

The total-hemispherical emittance can be evaluated using Planck's distribution. Therefore,

$$\varepsilon_{\mathrm{tot}}^{\mathrm{h}} = \frac{\int_0^\infty \varepsilon_\lambda^{\mathrm{h}}(\lambda) e_{\mathrm{b},\lambda}(\lambda, T)\mathrm{d}\lambda}{\int_0^\infty e_{\mathrm{b},\lambda}(\lambda, T)\mathrm{d}\lambda} = \frac{\int_0^\infty \varepsilon_\lambda^{\mathrm{h}}(\lambda) e_{\mathrm{b},\lambda}(\lambda, T)\mathrm{d}\lambda}{\sigma_{\mathrm{SB}} T^4} \tag{8.106}$$

The total emittance depends on the surface temperature and the spectral variation of the optical constants. Pure metals usually have a very low emittance, and the emittance can increase due to surface oxidation. Spectrally selective materials that appear to be reflective to visible light may exhibit a large total emittance, greater than 0.9 near ambient temperature; examples are snow and white paint. An earlier compilation of the radiative properties of many engineering materials can be found in Touloukian and DeWitt [50]. The use of surface microstructures to modify emission characteristics will be discussed in the next chapter.

8.4 Dielectric Function Models

Unlike in dilute gases where the molecules are far apart, in solids, the closely packed atoms form band structures. Absorption in solids usually happens in a much broader frequency region or band. Free electrons in metals can interact with the incoming electromagnetic waves or photons, and cause broadband absorption from the visible (or even ultraviolet) all the way to the microwave and longer wavelengths. For semiconductors especially with high impurity (doping) concentrations or at elevated temperatures, both the free electrons and holes contribute to the absorption process. The absorption of a photon makes the electron or the hole transit to a higher energy state within the same band. Therefore, free-carrier absorption is caused by *intraband transitions.* In order to conserve momentum, the carriers must also collide with ionized impurities, phonons, other carriers, grain boundaries, interfaces, and so forth. The collisions act as a *damping force* on the motion of carriers. The Drude model describes the oscillatory movement of an electron, driven by a harmonic field, which is subjected to a damping force. The model is simple in form and predicts the dielectric function of some metals fairly well in a broad spectral region, especially in the mid and far infrared.

Absorption by lattice vibrations or bound electrons, which is important for insulators and lightly doped semiconductors, is due to the existence of electric dipoles formed by the lattice. The strongest absorption is achieved when the frequency equals the vibrational mode of the dipole, i.e., the resonance frequency, which is usually in the mid- to far-infrared region of the spectrum. The contribution of bound electrons is often modeled by the Lorentz model.

Interband transition is the *fundamental absorption process* in semiconductors. An electron can be excited from the valence band to the conduction band by absorbing a photon, whose energy is greater than the energy gap E_g. Because the absorption by electrons is usually weak in semiconductors, a strong absorption edge is formed near the bandgap. In this transition process, both the energy and the momentum must be conserved.

This section discusses the formulation for different contributions to the dielectric function. It should be noted that the real and imaginary parts of the dielectric function are interrelated according to the causality, which is discussed first. Because all naturally occurring and most of the synthesized materials are nonmagnetic at high frequencies, only nonmagnetic materials are considered so that $\mu = 1$ and $n = \sqrt{\varepsilon}$ in the following, except in Sect. 8.4.6.

8.4.1 Kramers–Kronig Dispersion Relations

The real and imaginary parts of an analytic function are related by the Hilbert transform relations. Hendrik Kramers and Ralph Kronig were the first to show that the real and imaginary parts of the dielectric function are interrelated. These relations are called the Kramers–Kronig dispersion relations or K-K relations. The K-K relations can be interpreted as the causality in the frequency domain and are very useful in obtaining optical constants from limited measurements. The principle of causality states that *the effect cannot precede the cause,* or *no output before input.* Some important relations are given here, and a detailed derivation and proofs can be found from Jackson [5], Born and Wolf [8], and Bohren and Huffman [9].

The real part ε' and the imaginary part ε'' of a dielectric function are related by

$$\varepsilon'(\omega) - 1 = \frac{2}{\pi} \wp \int_0^\infty \frac{\zeta \varepsilon''(\zeta)}{\zeta^2 - \omega^2} \mathrm{d}\zeta \tag{8.107a}$$

and

$$\varepsilon''(\omega) - \frac{\sigma_0}{\varepsilon_0 \omega} = -\frac{2\omega}{\pi} \wp \int_0^\infty \frac{\varepsilon'(\zeta) - 1}{\zeta^2 - \omega^2} \mathrm{d}\zeta \tag{8.107b}$$

where σ_0 is the dc conductivity, $\wp$ denotes the Cauchy principal value of the integral, and ζ is a dummy frequency variable. These relations can be written in terms of n and κ as

$$n(\omega) - 1 = \frac{2}{\pi} \wp \int_0^\infty \frac{\zeta \kappa(\zeta)}{\zeta^2 - \omega^2} \mathrm{d}\zeta \tag{8.108a}$$

$$\kappa(\omega) = -\frac{2\omega}{\pi} \wp \int_0^\infty \frac{n(\zeta) - 1}{\zeta^2 - \omega^2} \mathrm{d}\zeta \tag{8.108b}$$

Equations (8.107a), (8.107b) and (8.108a), (8.108b) are the K-K relations, which relate the real part of a causal function to an integral of its imaginary part over all frequencies, and vice versa. A number of sum rules can be derived based on the K-K relations and are useful in obtaining or validating the dielectric function of a given material [15]. The K-K relations can be applied to reflectance spectroscopy to facilitate the determination of optical constants from the measured reflectivity of a material. For radiation incident from a vacuum to a medium with a complex refractive index $(n + \mathrm{i}\kappa)$ at normal incidence, the Fresnel reflection coefficient for TE waves is

$$r(\omega) = |r(\omega)| \mathrm{e}^{\mathrm{i}\phi(\omega)} = \frac{1 - n(\omega) - \mathrm{i}\kappa(\omega)}{1 + n(\omega) + \mathrm{i}\kappa(\omega)} \tag{8.109}$$

where $|r|$ is the amplitude and ϕ the phase shift upon reflection for the electric field. The reflectivity expressed in terms of ω is

$$\rho'_\omega(\omega) = r r^* = |r|^2 \tag{8.110}$$

The amplitude and the phase are related, and it can be shown that

$$\phi(\omega) = -\frac{\omega}{\pi} \wp \int_0^\infty \frac{\ln \rho'_\omega(\zeta)}{\zeta^2 - \omega^2} \mathrm{d}\zeta \tag{8.111}$$

The refractive index and the extinction coefficient can be calculated, respectively, from

$$n(\omega) = \frac{1 - \rho'_\omega}{1 + \rho'_\omega - 2\cos\phi\sqrt{\rho'_\omega}} \tag{8.112a}$$

and

$$\kappa(\omega) = \frac{2\sin\phi\sqrt{\rho'_\omega}}{1 + \rho'_\omega - 2\cos\phi\sqrt{\rho'_\omega}} \tag{8.112b}$$

8.4.2 The Drude Model for Free Carriers

The Drude model describes the frequency-dependent conductivity of metals and can be extended to free carriers in semiconductors. In the absence of an electromagnetic field, free electrons move randomly. When an electromagnetic field is applied, free electrons acquire a nonzero average velocity, giving rise to an electric current that oscillates at the same frequency as the electromagnetic field. The collisions with the stationary atoms result in a damping force on the free electrons, which is proportional to their velocity. The equation of motion for a single free electron is then

$$m_e \ddot{\mathbf{x}} = -m_e \gamma \dot{\mathbf{x}} - e\mathbf{E} \tag{8.113}$$

where e is the absolute charge of an electron, m_e is the electron mass, and γ denotes the strength of the damping due to collision, i.e., the *scattering rate* or the inverse of the relaxation time τ. Assume the electron motion under a harmonic field $\mathbf{E} = \mathbf{E}_0 e^{-i\omega t}$ is of the form $\mathbf{x} = \mathbf{x}_0 e^{-i\omega t}$ so that $\ddot{\mathbf{x}} = -i\omega\dot{\mathbf{x}}$. We can rewrite Eq. (8.113) as

$$\dot{\mathbf{x}} = \frac{e/m_e}{i\omega - \gamma}\mathbf{E} \tag{8.114}$$

The electric current density is $\mathbf{J} = -n_e e\dot{\mathbf{x}} = \tilde{\sigma}(\omega)\mathbf{E}$; therefore, the complex conductivity is

$$\tilde{\sigma}(\omega) = \frac{n_e e^2/m_e}{\gamma - i\omega} = \frac{\sigma_0}{1 - i\omega/\gamma} \tag{8.115}$$

where $\sigma_0 = n_e e^2 \tau / m_e$ is the dc conductivity, as discussed in Chap. 5. Equation (8.115) is called the Drude free-electron model, which describes the frequency-dependent complex conductivity of a free-electron system, in terms of the dc conductivity and the scattering rate, in a rather simple form. The electrical conductivity approaches the dc conductivity at very low frequencies (or very long wavelengths). The dielectric function is related to the conductivity by Eq. (8.28); thus,

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\sigma_0 \gamma}{\varepsilon_0(\omega^2 + i\gamma\omega)} \tag{8.116}$$

where ε_∞, which is on the order of 1, is included to account for contributions, other than the contribution of the free electrons, that are significant at high frequencies. There exist several transitions at the ultraviolet and visible regions for metals, such as *interband transitions*. Note that when $\omega \to \infty$, the real part of the dielectric function of all materials should approach unity, as can be seen from Eq. (8.107a). In the low-frequency limit when $\omega \ll \gamma$, $\tilde{\sigma}(\omega \to 0) \approx \sigma_0$ and $\varepsilon'' \gg \varepsilon'$. Therefore,

$$n \approx \kappa \approx \sqrt{\frac{\sigma_0}{2\varepsilon_0 \omega}} \tag{8.117}$$

This is the Hagen–Ruben equation and is applicable at very long wavelengths [1]. Both the refractive index and the extinction coefficient will increase with the square root of wavelength in vacuum. It is interesting to note that the radiation penetration depth $\delta_\lambda = \lambda/(4\pi\kappa)$ will also increase with the square root of wavelength. As an example, consider gold at $\lambda = 4$ μm with $\kappa = 25$. The penetration depth is 13 nm at this wavelength. If the wavelength is increased to 4 cm, which is well into the microwave region, the penetration depth would increase to 1.3 μm. Generally speaking, metals are highly reflecting in the infrared wavelength region.

The *plasma frequency* is defined according to $\omega_\mathrm{p}^2 = \frac{\sigma_0\gamma}{\varepsilon_0} = \frac{n_\mathrm{e}e^2}{m_\mathrm{e}\varepsilon_0}$. Using the plasma frequency, we can write Eq. (8.116) in a more compact form as follows:

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{\omega(\omega + \mathrm{i}\gamma)} \tag{8.118}$$

If $\omega \gg \gamma$, the dielectric function can be approximated as

$$\varepsilon(\omega) \approx \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{\omega^2}\left(1 - \mathrm{i}\frac{\gamma}{\omega}\right), \text{ when } \omega \gg \gamma \tag{8.119}$$

The plasma frequency falls in the ultraviolet region for most metals. For example, the wavelength corresponding to the plasma frequency is approximately 80 nm for aluminum and 200 nm for tungsten. When $\omega \gg \omega_\mathrm{p}$, as in the x-ray region, $\varepsilon(\omega) \to 1 + \mathrm{i}\gamma\omega_\mathrm{p}^2/\omega^3$. Thus, metals become highly absorptive and not so reflective. Take tungsten as an example. At $\lambda = 1$ nm, the optical constants are $n \approx 1$ and $\kappa = 4 \times 10^{-4}$. The penetration depth is calculated to be $\delta_\lambda = 200$ nm. Because the refractive index is similar to that of air, the reflection is very weak and most of the incident radiation is absorbed within a depth of 1 μm. Some metal foils become semitransparent, for example, the radiation penetration depth in lithium is close to 100 μm at $\lambda = 1$ nm. The Center for X-Ray Optics at Lawrence Berkeley National Laboratory maintains a website on x-ray properties [51]. If $\omega < \omega_p$, the real part of the dielectric function ε' becomes negative, and the extinction coefficient is much greater than the refractive index, i.e., $\kappa \gg n$. According to Eq. (8.84), this corresponds to a high reflectivity. A vanishing real part of the refractive index corresponds to a longitudinal collective oscillation of the electron gas, i.e., a *plasma oscillation*. Plasma oscillations originate from a long-range correlation of electrons caused by Coulomb forces.

Example 8.7 From Table 5.2, calculate the plasma frequency and the electron scattering rate for aluminum. Then calculate its dielectric function and compare the normal reflectivity with data.

Solution For aluminum near room temperature, $n_\mathrm{e} = 18.1 \times 10^{28}$ m^{-3} and $\sigma_0 = 1/r_\mathrm{e} = 3.75 \times 10^7$ m/Ω. From Appendix A, $e = 1.602 \times 10^{-19}$ C, $m_\mathrm{e} = 9.109 \times 10^{-31}$ kg, and $\varepsilon_0 = 8.854 \times 10^{-12}$ C^2/N m^2. Hence, $\gamma = n_\mathrm{e}e^2/m_\mathrm{e}\sigma_0 = 1.4 \times 10^{14}$ rad/s, or the scattering time $\tau = 7.2 \times 10^{-14}$ s, and $\omega_\mathrm{p} = 2.4 \times 10^{16}$ rad/s,

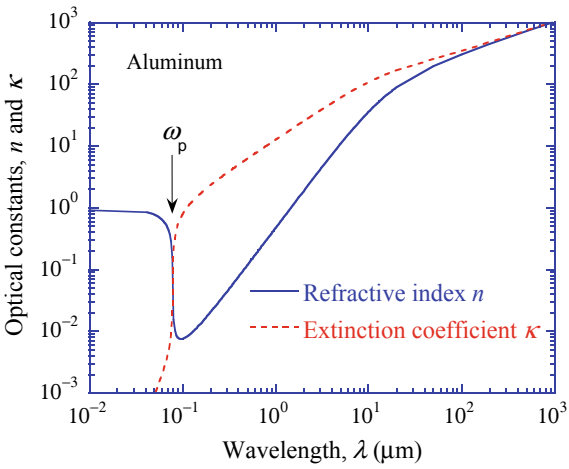

Fig. 8.13 Optical constants of aluminum, calculated from the Drude model

which corresponds to a wavelength of 79 nm. The exact parameters may differ slightly in different references, and sometimes an effective mass is used which is slightly larger than the electron rest mass. The predicted optical constants are plotted in Fig. 8.13, assuming $\varepsilon_\infty = 1$. It can be seen that as the wavelength exceeds 100 μm, the difference between n and κ diminishes. In the region 0.1 μm $< \lambda <$ 200 μm, $n < \kappa$ so that the real part of the dielectric function $\varepsilon' = n^2 - \kappa^2$ becomes negative. A sharp transition occurs at the plasma frequency so that $n \to 1$ and κ decreases rapidly toward higher frequencies.

As shown in Fig. 8.14, the reflectivity calculated from Eq. (8.84) is compared with the measured data for an aluminum film, which was prepared by ultrahigh vacuum deposition and measured in high vacuum to avoid oxidation [15]. The results agree very well at wavelengths greater than 2 μm. For $\lambda < 1$ μm, the contribution from the interband transition causes a reduction in the reflectivity. Note that the simple Drude model did not include these effects and is applicable for long wavelengths only. The established optical constants of metals are based on the measured reflectivity in a broad spectral region by using the K-K relations described in Sect. 8.4.1. The results for a large number of samples are tabulated in *Handbook of the Optical Constants of Solids*, with pertinent references [15].

In some studies, the Drude model is modified by considering the temperature and frequency dependence of the scattering rate and the effective mass. While the Drude model predicts well the radiative properties at room temperature or above, caution should be taken at extremely low temperatures. If the electron mean free path becomes comparable to the distance over which the electric field varies, i.e., the field penetration depth, nonlocal effects become important and the Drude theory breaks down. This can occur at cryogenic temperatures, and a more complex theory called the *anomalous skin effect* theory must then be applied [52].

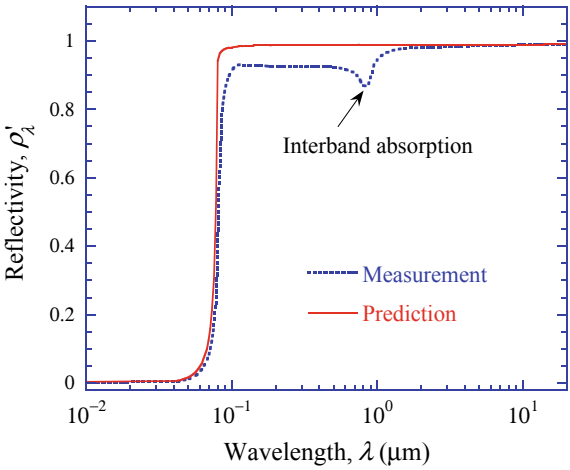

Fig. 8.14 Normal spectral reflectivity of aluminum

8.4.3 The Lorentz Oscillator Model for Phonon Absorption

Vibrations of lattice ions and bound electrons contribute to the dielectric function in a certain frequency region, often in the mid infrared. The refractive index can be calculated using the Lorentz oscillator model, which assumes that a bound charge e is accelerated by the local electric field $\mathbf{E}$, which is assumed to be the same as the applied field here. In contrast to free electrons, a bound charge experiences a restoring force determined by a spring constant K_j. The oscillator is further assumed to have a mass m_j and a damping coefficient γ_j, as shown in Fig. 8.15. The force balance yields the equation of motion for the oscillator:

$$m_j\ddot{\mathbf{x}} + m_j\gamma_j\dot{\mathbf{x}} + K_j\mathbf{x} = e\mathbf{E} \tag{8.120}$$

The solution for a harmonic field $\mathbf{E} = \mathbf{E}_0 \mathrm{e}^{-\mathrm{i}\omega t}$, valid at timescales greater than the relaxation time, is given by

$$\mathbf{x} = \frac{e/m_j}{\omega_j^2 - \mathrm{i}\gamma_j\omega - \omega^2}\mathbf{E} \tag{8.121}$$

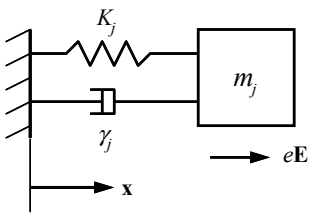

Fig. 8.15 The classical oscillator model

where $\omega_j = (K_j/m_j)^{1/2}$ is the resonance frequency of the jth oscillator. The motion of the single oscillator causes a dipole moment $e\mathbf{x}$. If the number density of the jth oscillator is n_j, the polarization vector, or the dipole moment per unit volume, is $\mathbf{P} = \sum_{j=1}^{N} n_j e\mathbf{x}$, where N is the total number of infrared active phonon modes (oscillators). The constitutive relation gives the polarization as $\mathbf{P} = (\varepsilon - 1)\varepsilon_0\mathbf{E}$. It can be shown that

$$\varepsilon(\omega) = \varepsilon_\infty + \sum_{j=1}^{N} \frac{S_j \omega_j^2}{\omega_j^2 - \mathrm{i}\gamma_j\omega - \omega^2} \tag{8.122}$$

where ε_∞ is a high-frequency constant and $S_j = \omega_{\mathrm{p}j}^2/\omega_j^2 = n_j e^2/(\varepsilon_0 m_j \omega_j^2)$ is called the *oscillator strength.*

At very low frequencies, $\varepsilon(0) = \varepsilon_\infty + \sum_{j=1}^{N} S_j$, which is called the dielectric constant. The real and imaginary parts of the dielectric function (ε', ε'') and optical constants (n, κ) for a simple oscillator are illustrated in Fig. 8.16, near the resonance frequency for $\varepsilon_\infty = 1$. It can be seen from Eq. (8.122) and Fig. 8.16 that, for frequencies much lower or much higher than the resonance frequency ω_j, ε'' and κ are negligible. Only within an interval of γ_j around the resonance frequency is the absorption appreciable. Within the absorption band, the real part of the refractive index decreases with frequency; this phenomenon is called *anomalous dispersion*. It follows that in an interval of width γ_j around the resonance frequency, the Lorentz oscillator is highly reflecting and absorbing, while for higher or lower frequencies, it acts as a transparent material. The real part of the dielectric function becomes negative in a frequency region somewhat higher than ω_j. A more complicated treatment based on quantum mechanics yields a four-parameter model [53]. The previous classical oscillator model can be considered as a good approximation when the relaxation times of the longitudinal and transverse optical phonons are close to each other. In some

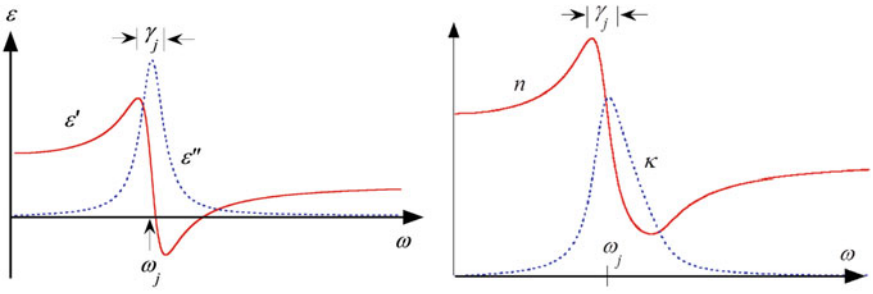

Fig. 8.16 The dielectric behavior predicted by the Lorentz oscillator model. **a** Real part and imaginary part of the dielectric function. **b** Refractive index and extinction coefficient

studies, frequency- and temperature-dependent scattering rate is also considered to model the infrared spectra.

Due to the large number of parameters involved, it is much more difficult to determine the Lorentz oscillator parameters than to determine the Drude parameters. In practice, the oscillator parameters are often treated as adjustable parameters that are obtained by fitting Eq. (8.122) to the measured reflectivity data. The Lorentz model has been applied to a large number of dielectric materials by fitting the reflectance spectra [15]. The author and collaborators have obtained the Lorentz parameters for several perovskite crystals ($LaAlO_3$, $LaGaO_3$, and $NdGaO_3$), thin polyimide films, HfO_2 and Ta_2O_5 films, as well as certain ceramic materials [54].

Example 8.8 The Lorentz model for SiC at room temperature for an ordinary ray is given as follows:

$$\varepsilon(\omega) = \varepsilon_\infty \left[1 + \frac{\omega_{\mathrm{LO}}^2 - \omega_{\mathrm{TO}}^2}{\omega_{\mathrm{TO}}^2 - \mathrm{i}\gamma\omega - \omega^2} \right] \tag{8.123}$$

where $\omega_{\mathrm{LO}} = 969\ \mathrm{cm}^{-1}$ and $\omega_{\mathrm{TO}} = 793\ \mathrm{cm}^{-1}$ are the frequencies corresponding to the longitudinal and transverse optical phonons, respectively, $\gamma = 4.76\ \mathrm{cm}^{-1}$, and $\varepsilon_\infty = 6.7$ [55]. What are the refractive indices at the high- and low-frequency limits? Calculate the normal reflectivity and compare it with the experimental result.

Solution Comparing Eqs. (8.122) and (8.123), we see that the resonance frequency corresponds to the TO phonon frequency, and the oscillation strength is $S_1 = \varepsilon_\infty(\omega_{\mathrm{LO}}^2/\omega_{\mathrm{TO}}^2 - 1) = 3.3$. The high-frequency limit of the refractive index is $n \approx \sqrt{\varepsilon_\infty} = 2.6$, and the low-frequency limit is $n = \sqrt{\varepsilon_\infty + S_1} = 3.16$. Note that transitions that occur in the visible and ultraviolet regions are not included so that the high-frequency limit is approximately 1 μm. On the other hand, because there are no other transitions at long wavelengths, the dielectric constant is approximately the same for zero frequency. The normal reflectivity is calculated using Eq. (8.84) and compared with the data, as shown in Fig. 8.17. The agreement is excellent since the Lorentz parameters were fitted to the experimental data [55]. The phonon band causes a large κ value and hence a high reflectivity (very low emissivity) between $\omega_{\mathrm{TO}} = 793\ \mathrm{cm}^{-1}$ and $\omega_{\mathrm{LO}} = 969\ \mathrm{cm}^{-1}$. This band is called *reststrahlen band*. The German word "reststrahlen" means "residual rays" and the reststrahlen effect indicates the phenomenon of high reflectance in a dielectric material that is otherwise transparent. At $\omega = 1000\ \mathrm{cm}^{-1}$, the reflectivity is nearly 0 such that the emissivity is almost 1. This happens at the edge of the reststrahlen band, where the refractive index increases close to 1 and the extinction coefficient decreases to a very small value. This wavelength is called the *Christiansen wavelength*, and the associated phenomenon is called the *Christiansen effect* [9].

The density-functional perturbation theory (DFPT) can be used to perform first-principles calculations of the lattice dynamics. It can provide phonon dispersions as well as the resonance frequencies of different phonon modes [56]. It should be noted that some optical phonons are symmetric and they cannot be detected by

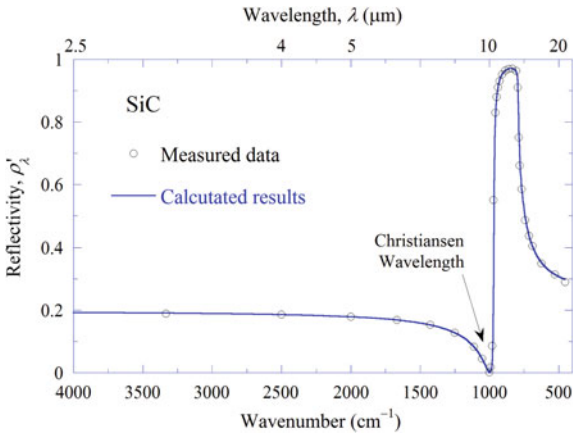

Fig. 8.17 The calculated and measured normal reflectivity of SiC at room temperature

infrared spectroscopy but can show up in Raman spectroscopy like the phonons in Si and Diamond. Note that both the TO and LO vibrational frequencies can be determined with the DFPT including those that are not infrared active but are Raman active [56, 57]. However, the determination of the scattering rate from the first-principles simulation is more challenging. Bao et al. [58] obtained the phonon lifetime and resonance frequencies using an analysis of the ab initio molecular dynamics (MD) trajectories based on the normal modes and spectral density analysis methods, allowing the calculation of the far-infrared dielectric function of GaAs. This method may also be applied to multiple phonon oscillators.

8.4.4 Semiconductors

The absorption coefficient of lightly doped silicon is shown in Fig. 8.18 to illustrate the contribution of different mechanisms [9, 59]. Let us look at the absorption of silicon in the visible and the infrared first, as shown in Fig. 8.18a. At short wavelengths, photon energies are large enough to excite electrons from the valence band to the conduction band. This interband transition causes the absorption coefficient to rise quickly as the photon energy $h\nu$ is increased above the indirect bandgap, which is approximately $E_g = 1.1$ eV at room temperature and decreases somewhat as temperature increases. As the wavelength further increases beyond the absorption edge, the absorption coefficient is affected by the existence of impurities and defects, absorption by free carriers (i.e., intraband or intersubband transitions by electrons and holes), and absorption by lattice vibrations. While the lattice vibration affects certain regions of the spectrum, the free-carrier contribution increases at longer wavelengths. For intrinsic silicon at low temperatures, the free-carrier concentration is very low, and thus silicon is transparent at wavelengths longer than the bandgap wavelength. Lattice absorption occurs in the mid infrared and introduces some absorption for

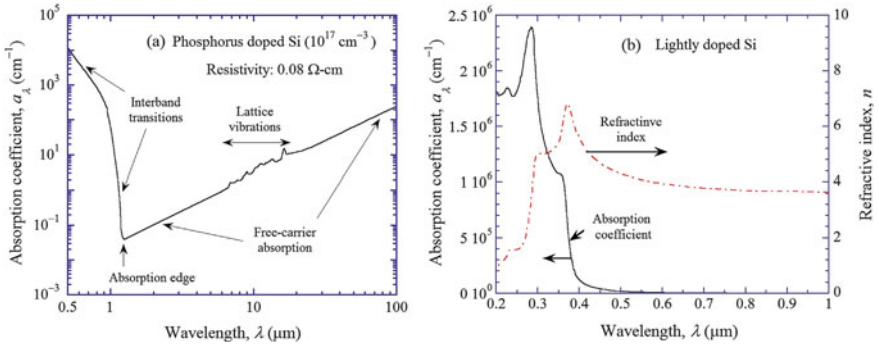

Fig. 8.18 The absorption coefficient and the refractive index of Si at room temperature. **a** Absorption coefficient in the visible and the infrared. **b** Absorption coefficient and refractive index from the ultraviolet to the near infrared

6 μm $< \lambda <$ 25 μm. Free-carrier absorption is important for doped silicon at longer wavelengths. Note that even for intrinsic silicon at high temperatures, thermally excited free carriers dominate the absorption at longer wavelengths; a 0.5-mm-thick silicon wafer is essentially opaque above 1000 K. The free-carrier concentration for intrinsic silicon is about 10^{10} cm^{-3} at 300 K and nearly 10^{18} cm^{-3} at 1000 K. As shown in Fig. 8.18b, the absorption coefficient continues to increase at shorter wavelengths, due to the interband transition associated with the direct bandgap, which dominates the optical characteristics of silicon in the ultraviolet region. This transition also affects the refractive index of silicon at longer wavelengths. Beyond 500 nm, the refractive index of lightly doped Si decreases somewhat as the wavelength increases.

Modeling the *interband transitions* requires quantum theory. First-principles or ab initio calculations have been performed to study the optical absorption spectrum of semiconductors and insulators, considering electron–hole interactions [60, 61]. In a direct-bandgap semiconductor, shown in Fig. 8.19a, the lowest point of the conduction band occurs at the same wavevector as the highest point of the valence band. An electron can be excited from the top of the valence band to the bottom of the conduction band by absorbing a photon of energy that is at least equal to the bandgap energy. When the valence band and the conduction band are parabola-like, the absorption coefficient due to direct bandgap absorption can be expressed as

$$a_{\text{bg}} = A(\hbar\omega - E_{\text{g}})^{1/2} \tag{8.124}$$

where A is a parameter that depends on the effective masses of the electrons and the holes, and the refractive index of the material.

When a transition requires a change in both energy and momentum, as in the case for an indirect bandgap semiconductor shown in Fig. 8.19b, a phonon is either emitted (process 1) or absorbed (process 2) for momentum conservation because the photon itself cannot provide a change in momentum. This kind of transition is called *indirect interband transition.* With the involvement of phonons, the absorption

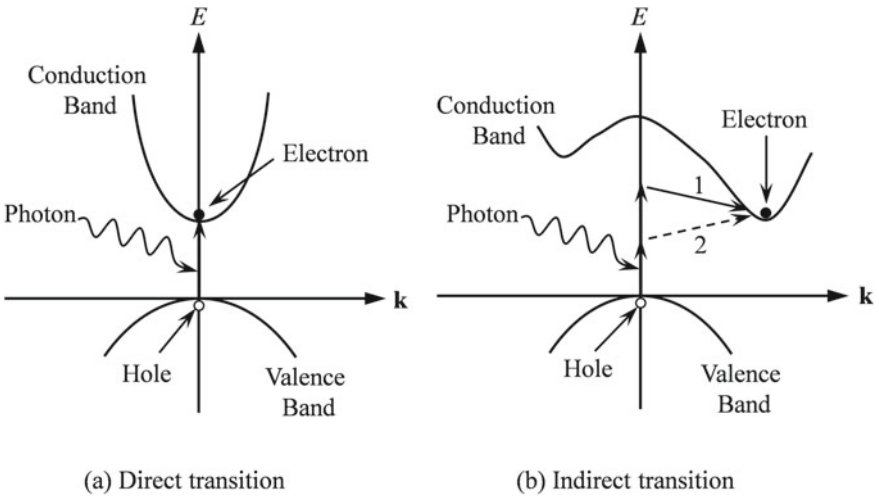

Fig. 8.19 Interband transitions in semiconductors. **a** Direct transition without involving a phonon. **b** Indirect transition involving the emission or absorption of a phonon

coefficient is given as

$$a_{\mathrm{a}}(\omega) = \frac{A(\hbar\omega - E_{\mathrm{g}} + \hbar\omega_{\mathrm{ph}})^2}{\exp(\hbar\omega_{\mathrm{ph}}/k_{\mathrm{B}}T) - 1}, \quad \hbar\omega > E_{\mathrm{g}} - \hbar\omega_{\mathrm{ph}} \tag{8.125}$$

and

$$a_{\mathrm{e}}(\omega) = \frac{A(\hbar\omega - E_{\mathrm{g}} - \hbar\omega_{\mathrm{ph}})^2}{1 - \exp(-\hbar\omega_{\mathrm{ph}}/k_{\mathrm{B}}T)}, \quad \hbar\omega > E_{\mathrm{g}} + \hbar\omega_{\mathrm{ph}} \tag{8.126}$$

where a_{a} and a_{e} correspond to the absorption coefficients for transitions with phonon absorption and emission, respectively, and their values are nonzero only when the photon energy is greater than the bandgap energy subtracted (or added) by the phonon energy. There may be several phonon modes that can cause indirect interband transitions, and their effects on the absorption coefficient can be superimposed [59].

The Drude model can be applied to model the free-carrier contribution for both intrinsic and doped silicon as given in the following [59, 62]:

$$\varepsilon(\omega) = \varepsilon_{\mathrm{bl}} - \frac{N_{\mathrm{e}}e^2/\varepsilon_0 m_{\mathrm{e}}^*}{\omega^2 + i\omega\gamma_{\mathrm{e}}} - \frac{N_{\mathrm{h}}e^2/\varepsilon_0 m_{\mathrm{h}}^*}{\omega^2 + i\omega\gamma_{\mathrm{h}}} \tag{8.127}$$

where the first term on the right $\varepsilon_{\mathrm{bl}}$ accounts for contributions by transitions across the bandgap and lattice vibrations, the second term is the Drude term for transitions in the conduction band (free electrons), and the last term is the Drude term for transitions in the valence band (free holes). Here, N_{e} and N_{h} are the concentrations, m_{e}^* and m_{h}^* the effective masses, and γ_{e} and γ_{h} the scattering rates of free electrons

and holes, respectively. The effective masses of silicon are taken as $m_e^* = 0.27m_0$ and $m_h^* = 0.37m_0$, where m_0 is the electron mass in vacuum.

The value of ε_{bl} is determined using the refractive index and the extinction coefficient of intrinsic silicon. The refractive index of silicon changes from 3.6 at $\lambda = 1$ μm to 3.42 at $\lambda > 10$ μm at room temperature and increases slightly at higher temperatures. Absorption by lattice vibrations occurs in silicon at wavelengths between 6 and 25 μm. To account for the lattice absorption, the extinction coefficients are taken from the tabulated values in *Handbook of the Optical Constants of Solids* [15]. At elevated temperatures or for heavily doped silicon, the effect of absorption by lattice vibrations is negligible compared to the absorption by free carriers. The carrier concentration and the scattering rate depend on the temperature and dopant concentrations. For bulk silicon, the scattering is caused by the collision of electrons or holes with the lattice (phonons) or ionized dopant sites (impurities or defects). The total scattering rates can be calculated by

$$\gamma_e = \gamma_{e-l} + \gamma_{e-d} \text{ and } \gamma_h = \gamma_{h-l} + \gamma_{h-d} \tag{8.128}$$

where the subscripts l and d stand for lattice and defects, respectively. Generally speaking, increasing the defect concentration or temperature gives rise to a larger scattering rate. For intrinsic silicon, the concentration of the thermally excited free electrons and holes is the same and can be found from the relation:

$$N_{th}^2 = N_C N_V \exp(-E_g/k_B T) \tag{8.129}$$

where N_C and N_V are the effective densities of states in the conduction band and the valence band, respectively, and for silicon, $E_g = 1.17 - 0.000473T^2/(T + 636)$ eV. Note that $N_C = 2.86 \times 10^{19}$ cm^{-3} and $N_V = 2.66 \times 10^{19}$ cm^{-3} at 300 K; however, both increase with temperature proportional to $T^{3/2}$. When the dopant concentrations are not very high, the free-carrier concentrations can be obtained from

$$N_e = \frac{1}{2}\left[N_D - N_A + \sqrt{(N_D - N_A)^2 + 4N_{th}^2}\right] \tag{8.130}$$

and $N_h = N_{th}^2/N_e$ when the majority impurities are n-type. When the majority impurities are p-type, the equations become $N_h = \frac{1}{2}\left[N_A - N_D + \sqrt{(N_A - N_D)^2 + 4N_{th}^2}\right]$ and $N_e = N_{th}^2/N_h$. Equation (8.130) has been derived based on complete ionization, which does not hold for heavily doped semiconductors or at very low temperatures. Integration is needed to determine the concentration when complete ionization is not expected, as described by Fu and Zhang [62].

The calculated optical constants n and κ of silicon, for wavelengths in the range between 1 and 100 μm, are shown in Fig. 8.20 at 300 and 1000 K for n-type phosphorus donors. The refractive index changes little for lightly doped silicon, even at high temperatures. The refractive index for heavily doped silicon first decreases and

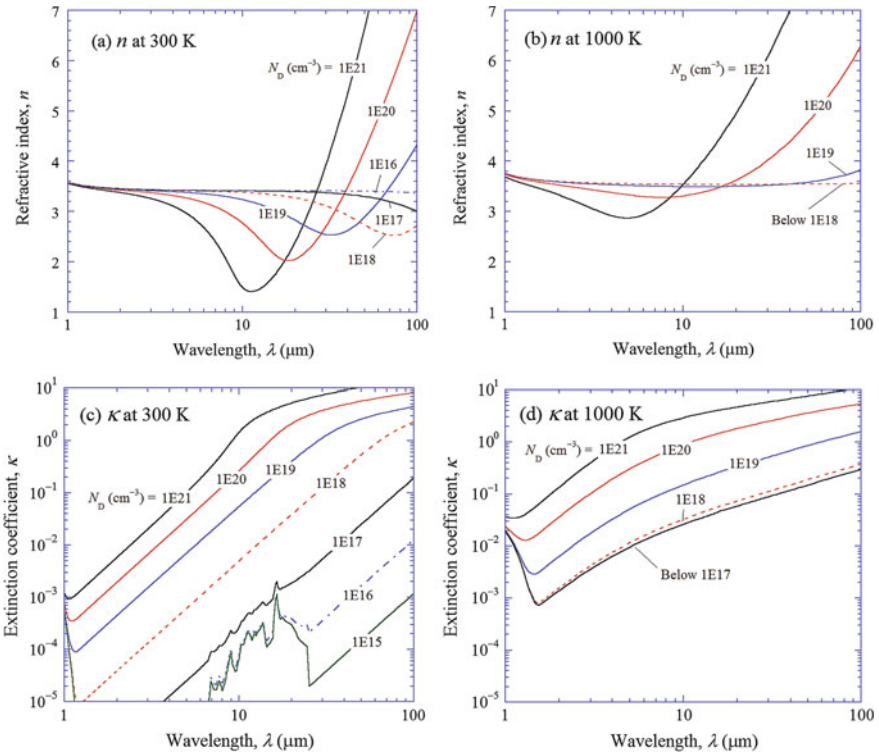

Fig. 8.20 Optical constants of *n*-type phosphorus-doped silicon, at 300 K and 1000 K, for different dopant concentrations

then increases abruptly toward longer wavelengths. The carrier contribution to the extinction coefficient at 300 K is very small for lightly doped silicon, and the lattice contribution can be clearly seen between 6 and 25 μm. As the doping level exceeds 10^{17} cm^{-3}, these phonon features are screened out. This is also true for lightly doped silicon at 1000 K as the thermally excited carriers have a concentration of about 10^{18} cm^{-3}. At 1000 K, κ is essentially the same for $N_D \leq 10^{17}$ cm^{-3} and increases with higher dopant concentrations. At 300 K, the calculated κ at $\lambda > 1.12$ μm decreases with reducing dopant concentration until N_D is less than 10^{10} cm^{-3}, when most carriers are from the thermal excitation rather than the doping. The lattice absorption features become prominent when $N_D \leq 10^{16}$ cm^{-3}. For doping levels under, 10^{18} cm^{-3}, $\kappa \ll n$ unless the wavelength is very long, and silicon behaves as a dielectric. The significance is that the radiation penetration depth can be very large in the mid infrared because of the small κ values. For heavily doped silicon, on the other hand, the Drude model predicts that $n \approx \kappa$ in the long-wavelength limit, just like in a metal. The accuracy of the simple Drude model is subjected to a number of factors, such as the dependence of the effective mass on temperature, dopant concentration, and even frequency. The scattering rate may be frequency dependent as

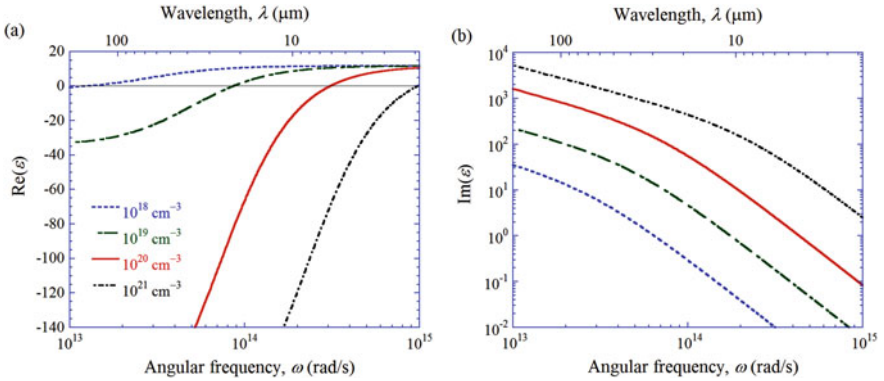

Fig. 8.21 **a** Real and **b** imaginary parts of the dielectric function for *n*-type silicon at 400 K for different dopant concentrations

well. The band structure may be modified for heavily doped silicon. Nevertheless, this model has captured the essential features of the dielectric function of silicon, for wavelengths greater than 0.5 μm, at temperatures from 300 to 1200 K, and with a doping level up to 10^{19} cm^{-3}.

Basu et al. [63] performed a comprehensive study on the ionization models and mobility models of doped silicon. It was pointed out that the ionization model adopted in the previous study of Fu and Zhang [62] would underpredict the carrier concentrations for doping concentrations greater than 10^{17} cm^{-3}. The model recommended by Basu et al. [63] compares well with the mid-infrared transmittance and reflectance spectra of both phosphorus (*p*-type) and boron (*n*-type)-doped silicon films at room temperature. The model can be extended into the temperature range from 250 to 400 K by modeling the temperature-dependent scattering rate using $\gamma(T)/\gamma_0 = (300/T)^{1.5}$, where γ_0 is the impurity scattering rate at room temperature calculated from the mobility model. Figure 8.21 plots the calculated dielectric function of heavily doped silicon at 400 K for angular frequencies between 10^{13} and 10^{15} rad/s with different dopant concentrations. An important feature is that the real part of the dielectric function ε' becomes negative especially for high doping concentrations. Such a metallic behavior due to free carriers can enable surface plasmon resonances and enhance near-field radiative heat transfer as will be discussed in subsequent chapters.

8.4.5 Superconductors

A *superconductor* is a material that exhibits zero resistance and perfect diamagnetism when it is maintained at temperatures below the critical temperature T_c, under a bias current less than the critical current and an applied magnetic field less than the

critical magnetic field. The discovery of high-temperature superconductors in the late 1980s has generated tremendous excitement in the public because the achievement of superconductivity above the boiling temperature of nitrogen (77 K at atmospheric pressure) offers many technological promises. More and more materials have been found to be superconducting at higher and higher temperatures. Extensive studies have been devoted to the infrared properties of superconducting films for applications such as radiation detectors, optical modulators, and other optoelectronic devices [64]. High-temperature superconducting (HTS) materials are made of ceramic structures, such as $YBa_2Cu_3O_{7-\delta}$, where δ is between 0 and 1. The Y-Ba-Cu-O compound behaves as an insulator when $\delta > 0.6$ and as a conductor when $\delta < 0.2$ at room temperature.

In the normal state ($T > T_c$), the dielectric function $\varepsilon(\omega)$ can be modeled as a sum of the free-electron contribution using the Drude model, an intraband absorption that is important for the mid-infrared region by using the Lorentz term, and a high-frequency constant [65]:

$$\varepsilon(\omega) = \varepsilon_\infty + \varepsilon_{\text{Mid-IR}} + \varepsilon_{\text{Drude}} \tag{8.131}$$

The expression of the Drude term is the same as in Eq. (8.116) or (8.118). Although phonon contributions can be neglected compared to the large electronic contributions, a broadband mid-infrared electronic absorption often exists in the HTS materials, which is typically modeled with a Lorentz oscillator that has a large width, or a frequency-dependent scattering rate.

Many properties of superconductors can be explained in terms of a two-fluid model that postulates that *a fluid of normal electrons coexists with a superconducting electron fluid.* These two fluids coexist but do not interact. According to the BCS theory [66], interaction between a pair of free electrons and a phonon (or other thermally generated excitations) leads to the formation of an electron pair, called a *Cooper pair*. The Cooper pairs cannot be scattered by any sources as they move in the lattice structure. In the superconducting state, only a fraction of free electrons f_s is in the condensed phase (or superconducting state) and the remaining electrons are in the normal state. The value of f_s is temperature dependent and goes to zero at T_c. The dielectric function in the superconducting state can be modeled by

$$\varepsilon(\omega) = \varepsilon_\infty + \varepsilon_{\text{Mid-IR}} + (1 - f_s)\varepsilon_{\text{Drude}} + f_s\varepsilon_{\text{Sup}} \tag{8.132}$$

The Drude term remains due to the presence of normal electrons with a number density of $(1 - f_s)n_e$. In Eq. (8.132), the dielectric function of the superconducting electrons can be modeled as

$$\varepsilon_{\text{Sup}} = -\frac{\omega_p^2}{\omega^2} + i\pi\delta(\omega)\frac{\omega_p^2}{\omega} \tag{8.133}$$

where $\delta(\omega)$ is the Dirac delta function. The calculated results are usually fitted with the experimental measurements by adjusting the plasma frequency, the scattering

rate, and the fraction of superconducting electrons. Excellent agreement has been observed between the predicted and experimental values of both the transmittance and the reflectance of superconducting films, at temperatures ranging from 300 down to 10 K [65].

8.4.6 *Metamaterials with a Magnetic Response*

The concept of negative refractive index ($n < 0$) was first postulated by Victor Veselago in 1968 for a hypothetical material that has both negative permittivity and permeability in the same frequency region. In this case, the sign of n should be chosen as negative in $n = \pm\sqrt{\varepsilon\mu}$. Many of the unique features associated with negative index materials (NIMs) were summarized in Veselago's original paper, such as negative phase velocity, reversed Doppler effect, and the prediction of a planar lens. As illustrated in Fig. 8.22a, if n is negative, the phase speed will be negative and light incident from a conventional positive index material (PIM) to a NIM will be refracted to the same side as the incidence. This is called *bending light in the wrong way.* Furthermore, if light can be bent differently, then a planar slab of a NIM can focus light as shown in Fig. 8.22b. The lack of simultaneous occurrence of negative ε and μ in natural materials hindered further study on NIMs for some 30 years. On the basis of the theoretical work by John Pendry and coworkers in the late 1990s, Shelby et al. [67] first demonstrated that a metamaterial exhibits negative refraction at x-band microwave frequencies. In a NIM medium, the phase velocity of an electromagnetic wave is opposite to its energy flux. The electric field, the magnetic field, and the wavevector form a left-handed triplet. For this reason, NIMs are also called left-handed materials (LHMs). Because both ε and μ are simultaneously negative, NIMs are also called double negative (DNG) materials.

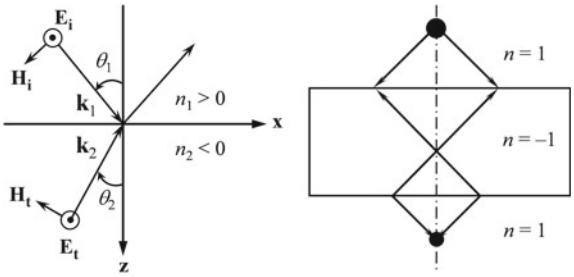

Fig. 8.22 Unique features of a negative index material (NIM). **a** The refracted ray bends toward the same side as the incidence. **b** A slab of NIM can focus light like a lens does. Arrows indicate the wavevector directions. Note that the energy direction is the opposite of the wavevector direction in a NIM

Pendry [68] conceived that a NIM slab with $\varepsilon = \mu = -1$ would perform the dual function of correcting the phase of the propagating components and amplifying the evanescent components, which only exist in the near field of the object. The combined effects could make a perfect lens that eliminates the limitations on image resolution imposed by diffraction for conventional lenses. Despite the doubt cast by some researchers on the concept of a "perfect lens" and even on negative refraction, both the hypotheses of negative refraction and the ability to focus light by a slab of NIM have been verified by analytical, numerical, and experimental methods. Potential applications of NIMs range from nanolithography to novel Bragg reflectors, phase-compensated cavity resonators, waveguides, and enhanced photon tunneling for microscale energy conversion devices [69–71]. Ramakrishna [70] gave an extensive bibliographic review on the theoretical and experimental investigations into NIMs and relevant materials. There has been growing interest in the study of NIMs because of the promising new applications as well as the intriguing new physics. The search for new ways of constructing NIMs also calls for the development of new materials and processing techniques.

The ideal case, where $\varepsilon = \mu = -1$, cannot exist at more than a single frequency because both ε and μ of a NIM must be inherently dependent on the frequency as required by the causality. In addition, real materials possess losses, and hence both ε and μ are complex. The negative index can be realized by considering the complex plane, as illustrated in Fig. 8.23. Note that $\varepsilon = r_\varepsilon e^{i\phi_\varepsilon}$ and $\mu = r_\mu e^{i\phi_\mu}$. Then, we have

$$n = r_n e^{i\phi_n} = \sqrt{r_\varepsilon r_\mu}\, e^{i(\phi_\varepsilon + \phi_\mu)/2} \tag{8.134}$$

Therefore, if both ε' and μ' are negative, n will be negative, but κ will always be positive. Note that a negative n can be obtained as long as $\phi_n > \pi/2$. Generally speaking, one would like to see all the phase angles be close to π so that the loss is minimized. Note that the principal value of the phase is chosen to be from 0 to 2π in the preceding discussion, rather than from $-\pi$ to π. If the latter is chosen, one would obtain a negative κ and a positive n for a NIM. Many metals and polar dielectrics have a negative ε in the visible and the infrared. Furthermore, periodic structures of thin metal wires or strips can dilute the average concentration of electrons and shift the plasma frequency to the far-infrared or longer wavelengths. Negative-μ

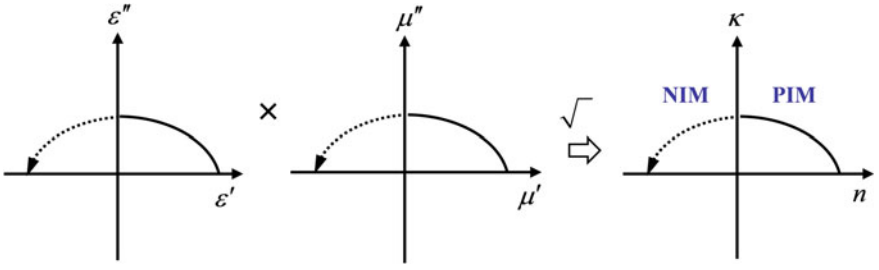

Fig. 8.23 Illustration of a negative refractive index, using the complex planes

materials rarely exist in nature, at the optical frequencies, but can be obtained using metamaterials consisting of split-ring resonator structures at microwave frequencies. These structures can be scaled down to achieve negative μ toward higher frequencies. The combination of repeated unit cells of interlocking copper strips and split-ring resonators makes a metamaterial exhibit a negative ε and μ simultaneously. Based on an effective-medium approach, the relative permittivity and permeability of a NIM can be expressed as functions of the angular frequency ω as follows:

$$\varepsilon(\omega) = 1 - \frac{\omega_{\mathrm{p}}^2}{\omega^2 + i\gamma_{\mathrm{e}}\omega} \tag{8.135}$$

and

$$\mu(\omega) = 1 - \frac{F\omega^2}{\omega^2 - \omega_0^2 + i\gamma_{\mathrm{m}}\omega} \tag{8.136}$$

where ω_{p} is the effective plasma frequency, ω_0 is the effective resonance frequency, γ_{e} and γ_{m} are the damping terms, and F is the fractional area of the unit cell occupied by the split ring. From Eqs. (8.135) and (8.136), both negative ε and μ can be realized in a frequency range between ω_0 and ω_{p} for adequately small γ_{e} and γ_{m}. Here, the values of ω_0, ω_{p}, γ_{e}, γ_{m}, and F depend on the geometry of the unit cell that constructs the metamaterial. These structures can be scaled down to achieve a negative index at higher frequencies.

To illustrate the negative index behavior, Fig. 8.24 shows the calculated refractive index and the extinction coefficient of a hypothetical NIM using the following parameters [72]: $\omega_0 = 0.5\omega_{\mathrm{p}}$, $F = 0.785$, and $\gamma_{\mathrm{e}} = \gamma_{\mathrm{m}} = \gamma = 0.0025\omega_{\mathrm{p}}$. Because of the scaling capability of the metamaterial, the frequency is normalized to ω_{p}. It can be seen that in the frequency range from ω_0 to ω_{p}, where the real parts of ε and μ are negative, n is negative and κ (for small values of γ) is small at frequencies not

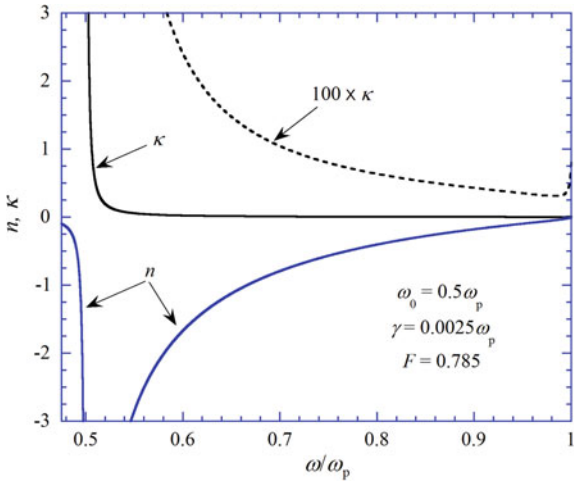

Fig. 8.24 Calculated refractive index of a hypothetical negative index material (NIM)

too close to ω_0. Further discussions of metamaterials and their radiative properties will be given in Chap. 9.

8.5 Experimental Techniques

Measurements of radiative properties (absorptance, emittance, reflectance, and transmittance) of real materials are critical for understanding the physical behavior of materials as well as for thermal analysis and design. The optical constants can be derived from the measured radiative or optical properties. For an opaque object with a smooth surface, measurement of the reflectivity in a broad spectral range can be used through KK relation to determine the complex refractive index if the material is nonmagnetic and isotropic. In the semitransparent region for a smooth slab (thick film), the measured spectral transmittance and reflectance can be used to extract the refractive index and absorption coefficient. For polar materials in the mid infrared (such as SiO_2 and Al_2O_3), the phonon oscillator parameters can be obtained by fitting the measured reflectance spectrum. Bidirectional reflectance and transmittance measurements are often used to study materials with surface roughness or inhomogeneity (such as porous materials or carbon nanotube arrays) due to surface and volume scattering. On the other hand, integrating spheres allow the diffused light to be collected and can be used to measure directional-hemispherical properties. The methods for measuring radiative properties can be grouped generally into two categories: calorimetric measurements and radiometric measurements, discussed in the following.

In a calorimetric technique, the thermal response of the specimen is used to determine the absorptance or emittance of the sample under investigation. Both the steady-state temperature change and transient temperature history can be used to deduce the radiative properties, though the calorimetric methods often use transient thermal responses during the heating or cooling process. The calorimetric method is well suited for measuring the total-hemispherical emittance of opaque materials [50, 73–75]. The sample is suspended in a large vacuum closure whose walls may be cooled with a cryogen such as liquid nitrogen. Due to the large wall area, the rate of net radiative transfer from the surface can be written as $q_{\text{rad}} = \varepsilon_{\text{tot}}^{\text{h}} A_{\text{s}} \sigma_{\text{SB}} (T^4 - T_{\text{w}}^4)$, where $\varepsilon_{\text{tot}}^{\text{h}}$, A_{s}, and T are the emittance, surface area, and temperature of the sample, and T_{w} is the wall temperature. The methods have been used to measure $\varepsilon_{\text{tot}}^{\text{h}}$ of certain metals up to about 1100 K [73, 74] as well as some solids and coatings down to 100 K [76]. In laser calorimetry [15, 77–79], the sample is heated up by a laser beam and the sample temperature depends on the laser power and spectral absorptance. Either the heating curve or cooling curve, after a shuttle is opened or closed, can be used to determine the absorptance based on a suitable thermal model. This method is particularly useful for measuring crystals with very low absorption coefficients [78, 79].

Radiometric techniques are based on the measurement of the radiant power reaching the detector (or receiver) from the source (or emitter). A variety of radiometric

techniques or optical methods exist for measuring the spectral radiative properties. An optical instrument or measurement system typically includes four parts: the source, detector, optical components, and the sample. Sometimes the sample to be studied can be the source (in emission measurements), detector (in absorption measurements), or part of the components. In general, the source can be a lamp, laser, a blackbody cavity, or a thermal emitter. The detector can be a thermopile or a bolometer that measures the radiation based on a temperature rise or a semiconductor detector that is based on photoconductive or photovoltaic principles as described in Chap. 6. Simple optical components include lenses, mirrors, polarizers, filters and windows, beam-splitters, prisms, gratings, and optical fibers. More complicated components, such as an interferometer and a monochromator, to be discussed in subsequent sessions, may combine several simple components.

Besides calorimetry and radiometry, polarimetry and ellipsometry are commonly used to determine the optical constants, from which the radiative properties can be calculated [15, 80]. These methods are largely based on the phase and amplitude of the electric field component, rather than the radiant power or intensity. Cezairliyan et al. [81] used the division-of-amplitude photopolarimeter (DOAP), which can measure the Stokes parameters, to determine the refractive index and extinction coefficient at $\lambda = 633$ nm for cylindrical specimens heated by a pulsed laser. The normal spectral emissivities of molybdenum and tungsten at temperatures between 2000 and 2800 K measured by the polarimetric technique agree well with those measured by the spectral radiometric technique using high-speed pyrometers [81]. Spectroscopic ellipsometers can nowadays perform measurements not only from the ultraviolet to the near-infrared (wavelengths 150–2500 nm) region, but also from the mid- to far-infrared region up to $\lambda = 120\ \mu$m [82].

In the following, we give some general discussions about the sources and detectors, the basics of dispersive instruments and the Fourier-transform spectrometer, along with setups for measuring directional-hemispherical properties with integrating spheres and for measuring spectral, directional emittance of materials at elevated temperatures. Measurements of the bidirectional reflectance and transmittance distribution functions will also be discussed, followed by a section on spectral ellipsometry.

8.5.1 Sources

For thermal radiation, a blackbody is the ideal source since its spectral distribution is well defined, as discussed previously. The radiation from the sun can be approximated as a blackbody at a temperature of about 5800 K. However, sunlight varies with time and atmospheric conditions. Therefore, it cannot be used as a source for quantitative measurements. Because the surface area of the walls must be much greater than the opening, blackbody cavities are bulky and must be carefully designed to maintain a uniform inner wall temperature. This has been done successfully in national metrology laboratories/institutes and used for measuring the Stefan–Boltzmann constant

and the spectral distribution that has resulted in the discovery of Planck's law [3, 16–19]. For high-temperature emittance measurements, blackbody cavities can be used as references. For measuring spectral transmittance and reflectance, since reference methods are commonly used, there is no need to precisely know the radiant power or intensity.

Lasers are quite commonly used in optical measurements since they provide well-collimated and nearly monochromatic radiation at discrete wavelengths. Some gas lasers have a very narrow spectral band due to atomic or molecular transitions, such as He-Ne lasers in the visible (633 nm) and near-infrared (1154 nm) or CO_2 lasers at wavelengths near 10 μm, especially at 10.6 μm. An optical cavity is made of a Fabry–Pérot resonator with two highly reflecting mirrors: one is opaque and the other partially transparent. The beam is reflected back and forth between the mirrors and makes multiple passes through the gain region (lasing medium) before it exits through the partially transparent mirror. This way, the stimulated emission within the gain region can be amplified. At present, semiconductor-based solid-state lasers are very popular and inexpensive. An example is a laser diode with a *p*-*n* junction that is similar to a light-emitting diode (LED). However, LEDs are based on spontaneous emission and produce light in a relatively broad spectral band (30–60 nm width). On the other hand, a diode laser is based on the stimulated emission of a *p*-*i*-*n* junction in which the active region is the intrinsic region (*i*) sandwiched between the *p*- and *n*-type direct bandgap semiconductors, such as GaAs, GaSb, InP, etc. Detailed descriptions of the mechanisms, types, and performances of various laser systems can be found elsewhere. While the laser is a powerful tool for optical measurements, spectrometer systems can quickly produce continuous spectral measurements in a broad wavelength band and thus are the most common instruments for measuring spectral radiative properties.

Incandescent lamps give out light when the filament is heated to an elevated temperature. Tungsten halogen lamps are perhaps the most popular and inexpensive light source for UV, VIS, and NIR measurements. The tungsten filament is heated to about 2800–3200 K in a mixture of inert gas and a halogen gas (such as bromine). The halogen gas reacts with tungsten that is being evaporated from the filament at high temperatures and redeposits the tungsten atoms back onto the filament; this is called the halogen cycle. The bulb is made of fused silica (or quartz) that is an amorphous SiO_2, which has a low coefficient of thermal expansion, high strength, and a high melting temperature. For a regular lamp without a halogen gas, the filament temperature cannot be very high and the tungsten is gradually evaporated and deposited onto the glass wall. At a temperature around 3000 K, the blackbody emission peak is near 1 μm. Since the emissivity of tungsten decreases with increasing wavelength, the emission peak shifts toward a shorter wavelength. Typically, a quartz tungsten-halogen lamp can be used from about 250 to 2500 nm wavelengths. The signal becomes weaker at longer wavelengths due to the reduction of tungsten emissivity. Furthermore, fused silica begins absorbing beyond 3 μm wavelength.

Globar made of SiC with the addition of rare earth oxides has been commonly used for infrared spectroscopy at wavelengths from about 2 to 100 μm. The temperature of the heating element is typically 1300–1650 K. The emissivity of the globar

is from 0.82 to 0.94 at wavelengths from 0.65 to 15 μm [83]. The globar source can be exposed to ambient conditions and lasts for a long time without turning it off, especially with nitrogen purging. Nowadays, most commercial mid-infrared spectrometers use globar emitters with different shapes and somewhat different operating temperatures.

A high voltage can ionize the gas molecules placed in between the cathode and anode. The high-temperature plasma generated by the electric discharge gives out arc light. This is the mechanism of lightning and has been used in arc welding. Common gas discharge lamps include the mercury arc lamp, xenon arc lamp, xenon flash lamp, and deuterium lamp. Although a high-voltage pulse (>20 kV) is needed to initiate the discharging process, gas discharge lamps usually operate with a low DC voltage (around 20 V) and a high current. Deuterium lamps emit unidirectional ultraviolet radiation with high stability. The wavelengths can range from as short as 115 nm (10.8 eV) to about 400 nm. They are commonly used in UV spectroscopic applications. The color temperature of a typical xenon lamp is around 6000 K and the emitted radiation is from 185 to 2000 nm. Hence, the xenon lamp has a closer match to the solar spectrum than other artificial sources. For this reason, xenon lamps are often used in solar simulators, UV/VIS spectrophotometers, and microscopes. Mercury arc lamps emit sharp peaks in the UV region with higher intensities than the xenon lamps with the same power consumption. Since the peaks are centered around 254 nm, the effective color temperature may exceed 10000 K. Another application of the mercury arc lamp is in the far infrared from 30 to 1000 μm, where the intensity of globar decreases more rapidly. In the far-infrared applications, the effective radiance temperature of mercury arc lamps is about 5000 K [84].

Synchrotron radiation or a synchrotron light source uses a circular particle accelerator. When electrons are accelerated under the magnetic field in the storage ring to a relativistic speed, electromagnetic radiation is emitted in the broad spectrum from x-ray to microwave with high brightness, collimation, and stability. Both linear and circular polarizations can be produced. Of course, synchrotron radiation is very expensive and available only in limited facilities. Synchrotron radiation is mostly used in x-ray studies since it is the brightest x-ray source and also in THz radiation (0.1–10 THz) studies where there is a lack of intensive sources [85].

8.5.2 Detectors

Generally speaking, there are two types of radiation detectors: thermal detectors and photon detectors. In a thermal detector, incident radiation causes a temperature variation that can be measured by a transducer that converts it to an electrical signal. In a photon detector or quantum detector, incident photons interact with the materials such as a semiconductor diode and cause electronic transitions to generate electron–hole pairs. Photodiodes can operate in either the photoconductive (PC) mode or the photovoltaic (PV) mode. The former is based on the change in electric conductivity and the latter is based on the voltage or current output, just like a solar cell, due

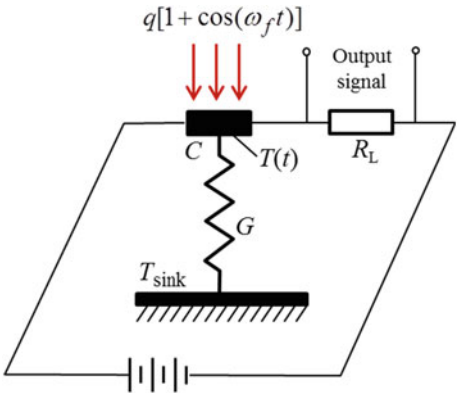

Fig. 8.25 Illustration of a bolometer coupled with a heat sink and a simple measurement system

to absorbed photons. Photon detectors are generally more sensitive and faster than thermal detectors. On the other hand, thermal detectors typically have a broadband response with good linearity.

A thermopile utilizes the thermoelectric effect and combines many thermocouple junctions in a series: the hot junctions are coated black to receive radiation and cold junctions are maintained at the heat sink temperature. Since their invention in the early 1830s, thermopiles combined with a galvanometer had been successfully used for measuring the solar spectrum and for studying blackbody radiation until the bolometer was invented by S. P. Langley in 1880 [3]. Bolometers are based on the temperature dependence of electrical resistance and can be orders of magnitude more sensitive than thermopiles. In the following, we will use a bolometer to illustrate the figures of merit of a radiation detector, specifically a thermal detector.

As shown in Fig. 8.25, a lumped capacitance model is used in which the detector is assumed to be at a uniform temperature with a heat capacitance C and is linked to a heat sink at temperature T_{sink} with a thermal conductance G. For incidence with a radiant power modulated at an angular frequency ω_f, the transient heat conduction equation may be written as [64, 86]

$$C\frac{\mathrm{d}\theta}{\mathrm{d}t} + G\theta = I^2R + \alpha q[1+\cos(\omega_f t)] \tag{8.137}$$

Here, $\theta = T - T_{\text{sink}}$ is the reduced temperature, R is the resistance of the detector element which is a function of temperature, I is the bias current which is assumed to be constant, and α is the absorptance of the detector (often coated with a wavelength-independent absorbing layer). The solution of Eq. (8.137) can be expressed as

$$\theta(t) = \frac{\alpha q\cos(\omega_f t - \phi)}{G_{\text{eff}}\left(1+\omega_f^2\tau^2\right)^{1/2}} + \theta_0 \tag{8.138}$$

where $\phi = \tan^{-1}(\omega_f t)$ is the phase lag, $G_{\text{eff}} = G - I^2(\mathrm{d}R/\mathrm{d}T)$ is the effective thermal conductance, and $\tau = C/G_{\text{eff}}$ is the *time constant*. The last term in Eq. (8.138) can be expressed as $\theta_0 = T_0 - T_{\text{sink}} = (\alpha q + I^2 R_0)/G$, where T_0 is the (average) operating temperature and R_0 is the resistance at T_0. The resistance of the bolometer may be expressed as $R = R_0 + (\mathrm{d}R/\mathrm{d}T)(T - T_0)$. Note that the *temperature coefficient of resistance* (TCR) is defined as $\beta = (\mathrm{d}R/\mathrm{d}T)/R_0$. Usually the bias current is sufficiently small such that $G_{\text{eff}} \approx G$. The time constant determines how high a modulation frequency can be used. Thermal detectors typically have a time constant in the range from milliseconds to seconds.

The responsivity S is the ratio of the output (voltage) signal to the incident power (modulated portion only). For a bolometer based on Eq. (8.138), we see that

$$S = \frac{\alpha\beta I R_0}{G_{\text{eff}}\left(1 + \omega_f^2 \tau^2\right)^{1/2}} \tag{8.139}$$

High TCR is critical for a high sensitivity. For this reason, superconductor bolometers have been developed that use the sharp resistance transition just above the critical temperature [64]. Some solids like vanadium dioxide (VO_2) experience a phase change above room temperature, with a large negative TCR during the insulator (semiconductor) to metal transition (around 340 K). This phenomenon has been used to build uncooled microbolometer arrays for infrared imaging applications.

Another figure of merit is called the noise equivalent power (*NEP*), which is the noise floor that limits the sensitivity since any signal below *NEP* cannot be distinguished from the noise. The *NEP* of a thermal detector depends on the background fluctuation called background noise, phonon noise due to the random exchange of thermal energy through the conductance G, Johnson noises of the detector resistance and load resistance due to random charge fluctuations, and the $1/f$ noise that is inversely proportional to frequency of the electronic signal [64, 86, 87]. By operating at cryogenic temperatures (e.g., using liquid helium), *NEP* can be reduced by orders of magnitude. The detectivity D^* is often used for comparing the sensitivity of different detectors and is defined as

$$D^* = \frac{\sqrt{AB}}{NEP} \tag{8.140}$$

Here, A is the detector area and B is the bandwidth in Hz. The units of D^* are usually expressed in terms of cm $\text{Hz}^{1/2}$ W^{-1}.

Another type of thermal detector that is commonly used in infrared spectrometers is the pyroelectric detector, which is based upon the thermally induced polarization change in pyroelectric materials. Commonly used pyroelectric materials are lithium tantalate ($LiTaO_3$), triglycine sulfate (TGS) or deuterated triglycine sulfate (DTGS), and lead zirconate titanate (PZT), which also has a large piezoelectric effect. When oscillating radiation is absorbed by a pyroelectric material, the temperature variation will change the degree of polarization, resulting in an oscillating voltage signal on

the load resistor. Room-temperature DTGS detectors have been adopted for many Fourier-transform infrared spectrometers. While it has a relatively low D^* compared to photon detectors, as shown in Fig. 8.26, a DTGS detector has a nearly flat spectral response and can be used in a large spectral range, extending to the far-infrared region (wavelengths up to 1000 μm). In the far infrared, liquid-helium-cooled bolometers are often used since the detectivity can be increased by orders of magnitude. Note that photoconductive mercury–cadmium–telluride (HgCdTe or MCT) detectors with various compositions and bandgaps are often used to achieve higher sensitivity in the mid infrared. The detectivity of a narrower band (higher sensitivity) MCT-1 and a wider band (lower sensitivity) MCT-2 is also shown in Fig. 8.26, along with a more sensitive photovoltaic indium antimonite (InSb) detector that is useful up to 5.5 μm. A photon detector based on a semiconductor diode has a cutoff frequency since the incoming photon energy must exceed the bandgap of the semiconductor material. Furthermore, as the frequency increases, a portion of the photon energy that exceeds the bandgap is lost to heat. It should be noted that MCT PC detectors have been reported with poor linearity and sometimes need to be calibrated and corrected [88].

In the visible and near-infrared region, photodiode detectors such as GaP (150–550 nm), Si (190–1100 nm), Ge (800–1800 nm), and InGaAs (900–2600 nm) can be used in their applicable spectral range with high linearity and detectivity. For measurements with very low light signals, such as measuring scattered light or single photon detection, photomultiplier tubes (PMTs) can be used from UV to NIR due to their extremely high sensitivities. A photomultiplier tube is a quartz vacuum tube where electrons are generated at the photocathode by photoemission as described in Chap. 6. The emitted electrons undergo a set of electrodes (called dynodes) where secondary emission occurs to release additional electrons. More and more electrons

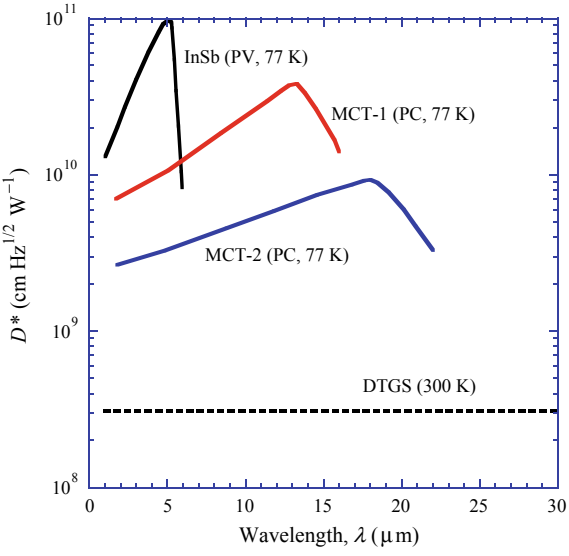

Fig. 8.26 Detectivity (D^*) of several infrared detectors, where the photovoltaic InSb and photoconductive MCT detectors are cooled by liquid nitrogen and the DTGS detector operates at room temperature

are generated and accelerated through multiple dynode stages with higher and higher electrical potential. The electrical current emitted by the photocathode can be amplified by up to eight orders of magnitude by the time the electrons reach the anode where they are collected [86, 87].

Though the original purpose of this invention was to develop better electronic memory for information storage, the charge-coupled device (CCD) has become a major technology in digital images and spectroscopic applications, such as digital cameras and the Hubble space telescope. CCDs are silicon devices that contain an array of metal-oxide-semiconductor capacitors. The charges (electrons or holes) generated by incoming photons through photoelectric effect are first stored in the potential well created by the gate electrodes and then shifted (transferred) to the next capacitor by applying appropriate clock pulses to the gate electrodes. Using a pulse train, the charges stored in each capacitor are eventually transferred to a terminal row to be read out serially and matched up with the location to provide a map of the incoming photon flux. This ingenious conception has led to development of the digital camera technology and certain UV/VIS/NIR spectrometers. In recent years, active pixel sensors based on complementary metal-oxide-semiconductor (CMOS) have been developed as the alternative technology for imaging applications. In a CMOS-based device, each photodetector (pixel) has its own amplifier so that the generated photocurrent is read out simultaneously by an integrated circuit array. At present, most of the commercial infrared focal plane arrays or IR cameras use CMOS technology as the integrated read-out device.

While detectors can be calibrated to measure the actual radiant power as a power meter, radiation detectors are often used for relative measurements according to the ratio of the sample signal to a reference signal. In some applications, such as solar irradiance measurements, radiation thermometry, absolute radiometry, and thermal imaging where accurate radiant power measurements are required, calibrations of the detector responsivity and optical throughput are necessary. Standard instruments such as the electrically self-calibrated radiometers, absolute cryogenic radiometers, and blackbody sources are often employed for these purposes [17, 18].

8.5.3 Dispersive Instruments

Before the laser was invented, most of the light sources were polychromatic. Interference filters, prisms, or gratings are typically used to obtain nearly monochromatic radiation. Multilayered dielectric (and sometimes metallic) films can be coated on a substrate to form interference filters that allow radiation from a narrow spectral band to pass through. An example is the Fabry–Perot interferometry that has sharp transmission peaks as to be further discussed in Chap. 9. A prism can effectively deflect broadband light into different directions, achieving nearly monochromatic radiation in selected directions. However, for spectroscopic applications, most contemporary instruments use surface relief gratings whose surfaces are corrugated periodically.

Polychromatic irradiation on a periodically corrugated surface is diffracted toward different directions (i.e., spatially dispersed according to the wavelength) based on the grating equation [38]:

$$\sin\theta_{\mathrm{d}}^{(m)} = \sin\theta_{\mathrm{i}} + \frac{m\lambda}{\Lambda} \tag{8.141}$$

where $\theta_{\mathrm{d}}^{(m)}$ is the angle of diffraction that depends on the angle of incidence θ_{i}, wavelength λ, grating period Λ, and diffraction order m (which is an integer). The zeroth order ($m = 0$) corresponds to specular reflection that is in the same direction for all wavelengths. When m is not equal to zero (positive or negative), the diffraction angle is wavelength dependent and different monochromatic radiation can be spatially dispersed upon reflection. Further discussion of grating theory will be given in Chap. 9.

Figure 8.27 shows the configuration of a Czerny–Turner monochromator with a rotating grating. The grating period must be greater than the maximum measurable wavelength. For a grating spectrometer, depending on the wavelength range, there may be tens to several thousands of grooves per millimeter. Some modern grating monochromators employ a linear array of sensors (CCD) without moving parts. It should be noted that suitable shortwave cutoff filters are needed to prevent unwanted radiation with higher diffraction orders from reaching the detector. This is because the diffraction angle depends on the product, $m\lambda$. Hence, radiation from shorter

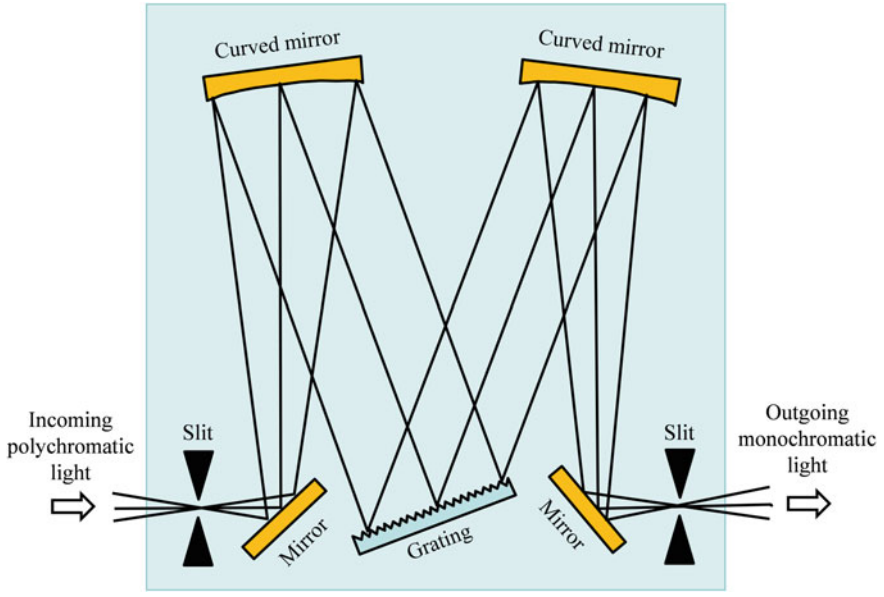

Fig. 8.27 Illustration of a Czerny–Turner grating monochromator. The grating can be rotated to vary the output wavelength

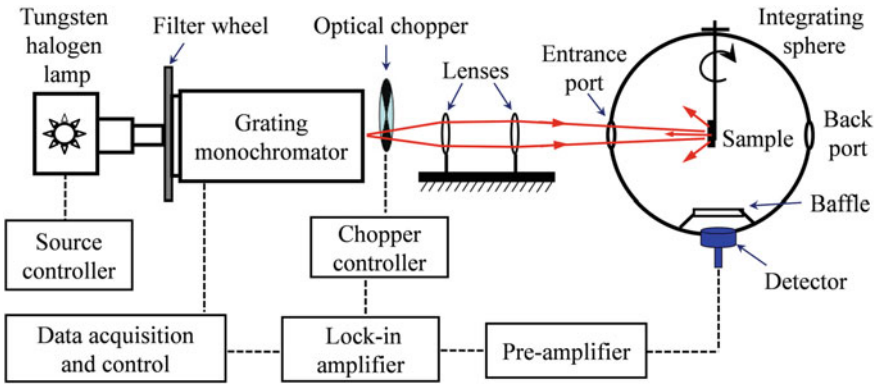

Fig. 8.28 A setup for measuring directional-hemispherical radiative properties using a monochromator and an integrating sphere [89, 90]

wavelengths and higher orders can have the same θ_d as for radiation from longer wavelengths and lower orders. The spectral resolution is often determined by the width of the exit slit. However, the ultimate *resolving power* (i.e., the ratio of the wavelength divided by the spectral resolution $\delta\lambda$) is limited by mN, where N is the total number of illuminated grooves [38]. A variety of commercial monochromators and spectrophotometers are available to meet the specific requirements for spectral range, sensitivity, and resolution.

Typically, grating spectrophotometers are used from the ultraviolet to the near infrared. At wavelengths beyond 2500 nm, the Fourier-transform infrared spectrometer has become the prevailing choice, as discussed in the subsequent section. A custom-built setup for measuring the directional-hemispherical reflectance and transmittance is shown in Fig. 8.28. The system is composed of a halogen lamp, a grating monochromator with a filter wheel, a chopper, and two lenses, which guide the monochromatic radiation to a 200-mm-diameter integrating sphere [89–91]. The inner wall of the integrating sphere is coated with a polytetrafluoroethylene (PTFE) diffuse reflector. Incident radiation can be focused on a spot size of approximately 6 mm $\times$ 6 mm on the center of the sphere, where the sample is mounted on a rotary holder through the top port. By rotating the sample holder, the beam can be directed either onto the back port of the sphere (covered by a PTFE plate) to obtain the reference signal or onto the sample to obtain the sample signal. A baffle placed above the detector located at the bottom port prevents the direct illumination of the detector by the first reflection of the sample or reference. Two photon detectors can be mounted at the bottom port of the sphere interchangeably: a Si photodiode for wavelengths from 300 to 1050 nm and a Ge photodiode for wavelengths from 1000 to 1800 nm. The detector signal is sent to a transimpedance pre-amplifier that has eight decades of dynamic range with a linear response. Afterward, the voltage signal is collected by a lock-in amplifier at the chopping frequency of 400 Hz. For an opaque sample, the directional-hemispherical reflectance can be obtained from the ratio of the sample signal to the reference signal. Furthermore, the back-mount method can also be

used for convenient measurement. In the back-mount configuration, the sample is placed at the back port outside the sphere, interchangeably with the PTFE reference [91]. A number of factors can affect the accuracy of the integrating sphere measurements and, therefore, calibration and corrections are often necessary to reduce the measurement uncertainty [92].

8.5.4 Fourier-Transform Infrared Spectrometer

Developed in the late 1960s, Fourier-transform infrared (FTIR) spectrometers have become a versatile tool for infrared spectral characterization of materials, including spectral transmittance, reflectance, absorptance, and emittance [93–97]. As schematically shown in Fig. 8.29, an FTIR system utilizes Michelson interferometer that consists of a beamsplitter, a fixed mirror, and a moving mirror to produce interference effects. The strength of the output optical signal depends on the relative position of the moving mirror. If the path lengths between the beamsplitter and the two mirrors are the same, the situation is identified as zero path difference (ZPD), and the power reaching the detector will be the largest since constructive interferences occur at all wavelengths. For monochromatic incident light, a periodic signal will reach the detector as the moving mirror travels due to the alternating constructive and destructive interferences. For polychromatic incident radiation, the detector receives a time-varying signal called an *interferogram*, which is a Fourier transform of the incident radiation weighted by the spectral efficiency of the optical system and detector responsivity. In general, the interferogram appears somewhat like a sinc function with a peak at the ZPD. Unlike dispersive spectrometers, the FTIR detector receives a time-varying signal that carries information about the radiative power in

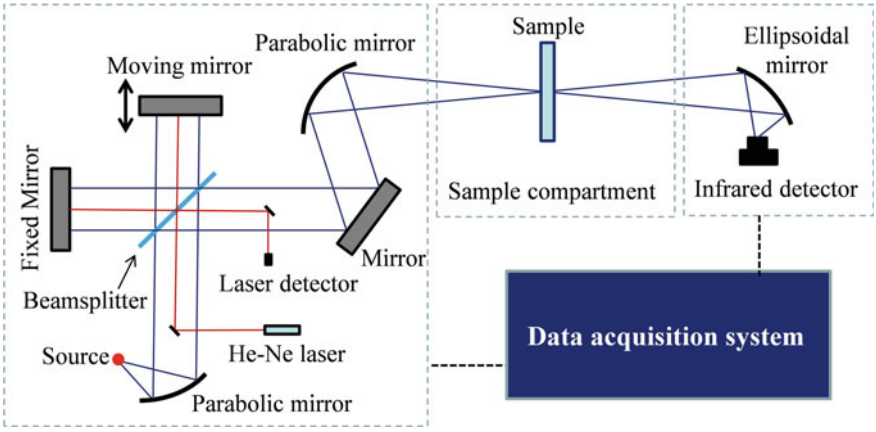

Fig. 8.29 Illustration of the Fourier-transform infrared spectrometer. Accessories can be introduced in the sample compartment for reflectance measurements

a broadband. Suppose the spectrum of the source is $I(\bar{\nu})$. After passing through the Michelson interferometer, the spectral radiant power arriving at the detector is a periodic function with a dc component:

$$\frac{1}{2}[1+\cos(2\pi x\bar{\nu})]\xi(\bar{\nu})I(\bar{\nu})$$

where $\bar{\nu}$ is the wavenumber, x is the path length difference, and $\xi(\bar{\nu})$ is the optical efficiency. The one-half term is due to the interferometer since the other half of the energy is reflected back to the source.

Since only the modulated part contributes to the interferogram, the detector output signal for polychromatic incidence is given as follows:

$$V(x)=\frac{1}{2}\int_0^\infty \xi(\bar{\nu})S(\bar{\nu})I(\bar{\nu})\cos(2\pi\bar{\nu}x)\mathrm{d}\bar{\nu} \tag{8.142}$$

where $S(\bar{\nu})$ is the detector responsivity. The data acquisition system performs an inverse fast Fourier transform with a computer to generate a relative spectral response function, typically called the single beam spectrum [93]:

$$I^*(\bar{\nu})=\int_{-\infty}^{\infty} V(\bar{\nu})\cos(2\pi\bar{\nu}x)\mathrm{d}x \tag{8.143}$$

If the velocity of the moving mirror is u, then $x = 2ut$. Hence, the detector receives a time-varying signal. However, the frequency $f = 2u\bar{\nu}$ typically falls in the range from several hundred to several thousand Hz depending on the wavelength of the incident radiation. This frequency can easily be measured by a thermal detector [94].

As shown in Fig. 8.29, a He-Ne laser with a well-characterized wavelength is used to precisely determine the location of the moving mirror with respect to the fixed mirror (i.e., the path length difference). The laser beam goes through the same interferometer to generate a sinusoidal wave that is detected by a photodiode detector. This enables high wavenumber accuracy for the resulting spectrum.

Spectral transmittance can be measured by dividing the spectrum with the sample by the reference spectrum when the sample is moved out of the optical paths as shown in Fig. 8.29. Reflectance accessories can be used both for specular and diffuse reflectance measurements [63, 91, 96]. Various other accessories can be used with FTIR spectrometers including attenuated total reflectance (ATR) that is based on evanescent waves [15, 93].

FTIR spectrometers have several advantages over dispersive spectrophotometers, such as high throughput, high signal-to-noise ratio, high resolution, and short measurement time. They are particularly suitable for measurements at wavelengths

beyond 2 μm. While FTIR spectrometers have very high wavelength accuracy, caution must be taken with regard to its radiometric accuracy in order to quantitatively measure the radiative properties [88, 94, 95].

Figure 8.30 shows the optical layout of a custom-built spectral emissometer that allows the heated sample to be rotated to measure the directional emittance for each polarization [97]. A blackbody calibration source is used as the reference and a flip mirror allows the emission signal from either the sample surface or the blackbody to be collected by the FTIR through its side port. The ellipsoidal reflector collects the radiation through a half-cone angle of approximately 3° from the sample surface and focuses it again onto the iris, which is used to limit the collecting area on the sample and to adjust the amount of radiation reaching the spectrometer. The opening of the iris can be matched to the FTIR system to achieve a spectral resolution of 1 cm^{-1}. For most measurements, however, a resolution of 4 cm^{-1} is usually sufficient. The parabolic reflector converts the radiation to a nearly collimated beam with a diameter of about 25 mm before sending it to the FTIR. A liquid-nitrogen-cooled InSb detector can also be used, which has a higher detectivity but a narrow spectral range from 2.0 to 5.5 μm. An IR wire-grid polarizer is mounted next to the iris for measuring the emittance with a chosen polarization.

The heater assembly is also shown on the right of Fig. 8.30. The sample was compressed on a copper disk, which was nickel plated to prevent oxidation. The nickel-plated copper disk maintained a uniform temperature underneath the sample. The copper surface was also polished before nickel plating to reduce thermal contact resistance. A coil heater was located at the back of the copper disk with an alumina plate inserted in between for electrical insulation. A K-type thermocouple probe

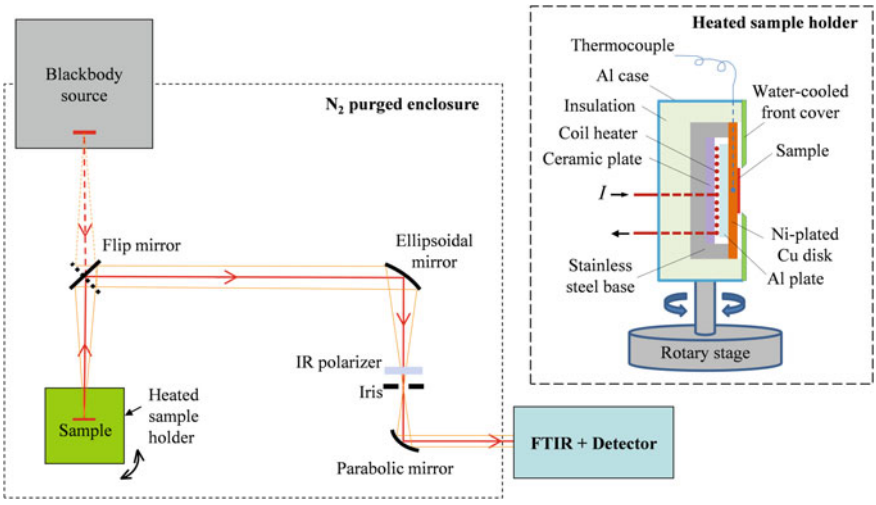

Fig. 8.30 Schematic of optical layout for the spectral emissometer for measuring angular-dependent emittance for each polarization. The heated sample holder is mounted on a rotary stage as shown on the right [97]

with oxidation-resistive sheathing was embedded inside the copper disk for sample temperature measurement. The thermocouple is also used with a PID temperature controller to set and control the sample temperature. The heater assembly was placed in refractory materials and mounted inside a metal box. The sample temperature can reach 1000 K with a power input around 140 W. The front cover of the heater assembly was water-cooled with an aperture of 25 mm in diameter. The heater assembly was mounted on a rotary stage to change the emission angle. The emissometer has been used to measure a SiC substrate for calibration and the coherent emission from an asymmetric Fabry–Perot planar multilayer structure as well as a metamaterial structure by excitation of magnetic polaritons [97]. More discussions of multilayers and magnetic polaritons will be given in the subsequent chapter.

8.5.5 *BRDF and BTDF Measurements*

To measure BRDF or BTDF, both (θ_i, ϕ_i) and (θ_r, ϕ_r) need to be changed while the distance between the sample and the detector should be fixed [45]. For in-plane measurements, the plane of incidence is the same as the plane of reflection so that the azimuthal angles can be fixed [98]. Figure 8.31 shows a diagram of a laser scatterometer. The laser beam is in a fixed position; rotating the detector allows the change of the polar angle of incidence θ_i, while rotating the detector around the sample using the goniometer allows the change of the reflection angle θ_r. In the actual setup, the detector is allowed to move out of the horizontal plane so that

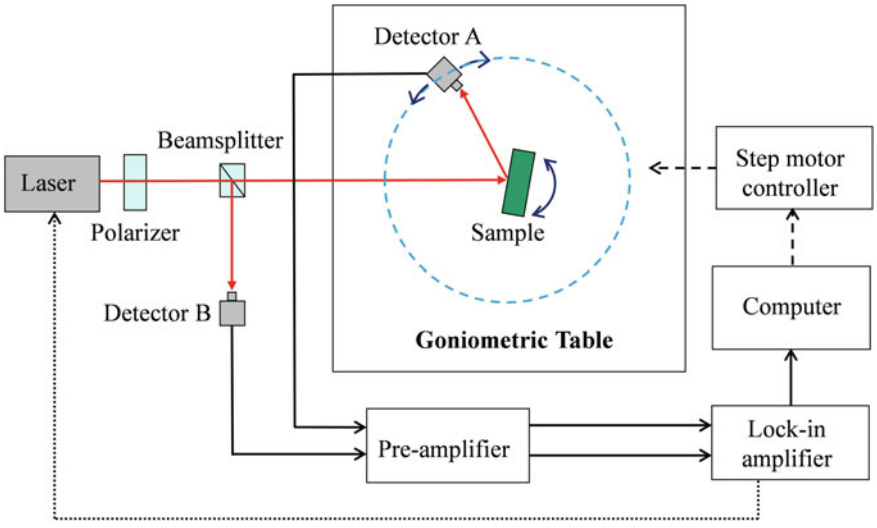

Fig. 8.31 Schematic of a laser scatterometer for BRDF/BTDF measurements [45]

the azimuthal angle of reflection ϕ_r can also be changed. This facility is called the three-axis automated scatterometer (TAAS) and was developed in the author's lab [45]. The three rotary stages are independently controlled by step motors with very high angular precision (better than 0.01°). The incident laser beam is parallel to the optical table and the sample is vertically mounted.

A fiber-coupled diode laser system serves as a collimated light source. A thermoelectric-cooled temperature controller maintains the laser at a constant operation temperature to achieve superior power stability less than 0.2%. The lock-in amplifier provides an alternating current (typically 400 Hz) to the laser controller and measures the detector signal (after the pre-amplifier) at the same modulation frequency to eliminate the effect of background radiation. The wavelength of the laser can be chosen using different laser diodes. A linear polarizer is used to polarize the incident light either parallel or perpendicular to the plane of incidence. Then, the light is split into two paths by a beamsplitter. The majority is transmitted through the beamsplitter to the sample and then reflected/scattered by the sample. The output power is measured with a signal detector A, whose signal is sent to a pre-amplifier. A smaller portion is reflected by the beamsplitter and measured by a reference detector B, whose signal is also sent to the pre-amplifier. Si and Ge photodiode detectors measure the radiant power in the wavelength range from 350 to 1100 nm and from 800 to 1800 nm, respectively. The transimpedance pre-amplifiers convert the current signal from the detectors to a voltage output with resistance values switchable from 10 to 10^9 Ω to achieve an eight-order dynamic range. A typical solid angle of the signal detector (with respect to the center of the laser spot on the sample) is $\Delta\Omega_r = 1.84 \times 10^{-4}$ sr, resulting in a half-cone angle of 0.45° [45]. The lock-in amplifier and step motors are connected to a desktop with the LabView environment for data acquisition and automatic rotary-stage control. It should be noted that the BRDF within ± 2.5° of the retroreflection direction cannot be measured since the sample detector would block the incident beam at this position. Laser diodes at 635, 891, 977, and 1550 nm have been employed in several investigations [45, 54, 99].

In the experiment, the detector output signal is proportional to the solid angle $\Delta\Omega_r$. The denominator of Eq. (8.95) gives the incident radiant power reaching the detector. Hence, the BRDF or BTDF can be obtained from the following measurement equation [45, 98]:

$$f_s = \frac{1}{P_i} \frac{P_s}{\cos\theta_r \Delta\Omega_r} \tag{8.144}$$

where f_s refers to either BRDF or BTDF, P_i is the laser power incident on the sample, P_s is the scattered power reaching the signal detector, and $\Delta\Omega_r$ is the solid angle of the detector area viewed from the beam centered on the sample. During the measurements, the beamsplitter ratio is first calibrated. Detector A is rotated to behind the sample while the sample is removed. The ratio of the signal from detector B to that from detector A gives an instrument constant C_I. In the measurements, $P_s/P_i = C_I V_A/V_B$, where V_A and V_B are the output voltages from the lock-in amplifier for detectors A and B, respectively. As long as the responses are linear, there is

no need to calibrate the detector responses. The use of the reference detector also eliminates the effect of laser power instability during the measurements. The specular reflectance, R_{sp}, can also be measured by positioning the signal detector in the specular direction with $\theta_r = \theta_i$ using the equation $R_{sp} = f_r \cos\theta_r \Delta\Omega_r$. The laser beam diameter (FWHM) is 3–5 mm, which is much smaller than the detector aperture whose diameter is 8 mm. This allows the specularly reflected power to be fully captured by the signal detector. In the measurements, V_A and V_B are averaged over many measurements at a given position to reduce the random error and the resulting uncertainty is typically within 5%.

8.5.6 *Ellipsometry*

When linearly polarized light is incident on an isotropic surface with or without a film, if the incident wave is either *p*- or *s*-polarized, the reflected wave is also linearly polarized as shown in Fig. 8.7. The ratio of the Fresnel reflection coefficients can be expressed as

$$\tilde{R} = \frac{r_p}{r_s} = \tan(\Psi)e^{i\Delta} \tag{8.145}$$

where Ψ and Δ are called the ellipsometric angles or parameters. Note that $\tan(\Psi)$ is the amplitude ratio and Δ is the phase difference between the two Fresnel reflection coefficients. For an opaque and nonmagnetic medium, at given angle of incidence, $\tilde{R}$ is a function of the optical constants (n, κ). If both $\tan(\Psi)$ and Δ can be measured, then the optical constants can be obtained. If a thin dielectric film is coated on an opaque substrate with known optical properties, then the refractive index and the film thickness can be simultaneously determined [80, 96]. It is also possible to use ellipsometry to study anisotropic crystals [82]. This is the principle of ellipsometry and various methods can be used to measure the ellipsometric parameters under oblique incidence [80, 82].

Figure 8.32 shows the rotating analyzer setup where a monochromatic beam (either from a laser or spectrometer) is incident at θ_i, which is usually greater than 60°. The incident wave on the sample is linearly polarized but with both *s*- and

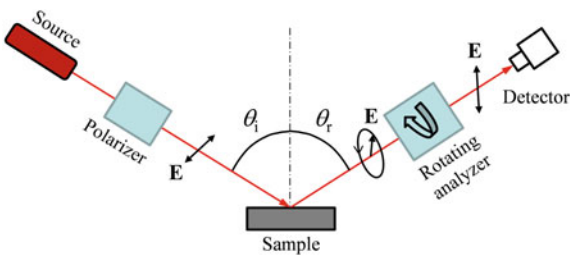

Fig. 8.32 Schematic of a rotary analyzer ellipsometer

p-components. In this case, the reflected wave is, in general, elliptically polarized. The analyzer is another linear polarizer that is rotated to give out a sinusoidal signal to the detector. A quarter-wave compensator can sometimes be used. Furthermore, the incident polarizer can also be rotated to change the incident wave polarization. Through suitable data processing including regression analysis, the two ellipsometric parameters can be determined at each wavelength. With the development of spectral ellipsometers and the extension to the mid- to far-infrared regions, ellipsometry has become a complementary and alternative technique for the study of radiative properties of materials [82, 100].

8.6 Summary

In this chapter, we used the macroscopic Maxwell equations to derive the plane wave equation and subsequently defined the optical properties for isotropic materials. Planck's law was derived based on statistical mechanics. After a brief discussion of radiation thermometry, radiation pressure and photon entropy were then introduced. The reflection and refraction of waves at a smooth interface were derived based on the electromagnetic wave theory. This chapter also presented the dielectric functions for metals, dielectrics, semiconductors, superconductors, as well as materials with a magnetic response or metamaterials. The concept of NIM or DNG materials, as well as their unique features, was also explained. The last section surveyed the experimental techniques typically used for measuring radiative properties from ultraviolet to the far infrared. This chapter serves as the foundation of the subsequent chapters, in which we will provide extensive discussions on the radiative properties of semitransparent materials, windows, multilayers, periodic gratings, rough surfaces, as well as evanescent waves, surface polaritons, photon tunneling, and near-field radiative heat transfer.

Problems

8.1 Write the wave equation in the 1D scalar form as $\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2}$, where c is a positive constant. Prove that any analytical function f can be its solution as long as $\psi(x,t) = f(x \pm ct)$. Plot ψ as a function of x for two fixed times t_1 and t_2. Show that the sign determines the direction (either forward or backward) and c is the speed of propagation. Develop an animated computer program to visualize wave propagation.

8.2 Consider an electromagnetic wave propagating in the positive z-direction, i.e., $\mathbf{k} = k\hat{\mathbf{z}}$. Plot the vibration ellipse, and compare it with Fig. 8.2 for two cases: (1) $\mathbf{a} = 3\hat{\mathbf{x}}$ and $\mathbf{b} = \hat{\mathbf{x}} + 2\hat{\mathbf{y}}$ and (2) $\mathbf{a} = 3\hat{\mathbf{x}}$ and $\mathbf{b} = -2\hat{\mathbf{x}} + \hat{\mathbf{y}}$. Consider the spatial dependence of the electric field at a given time, say $\omega t = 2\pi m$, where m is an integer. Discuss how $\mathbf{E}$ will change with kz for the following two cases:

(3) $\mathrm{Re}(\mathbf{E}_0) = 3\hat{\mathbf{x}}$ and $\mathrm{Im}(\mathrm{E}_0) = 0$ and (4) $\mathrm{Re}(\mathbf{E}_0) = 3\hat{\mathbf{x}}$ and $\mathrm{Im}(\mathbf{E}_0) = -3\hat{\mathbf{y}}$. The polarization is said to be right handed if the end of the electric field vector forms a right-handed coil or screw in space at any given time. Otherwise, it is said to be left handed. Discuss the handedness for all four cases.

8.3 Integrate Eq. (8.17) over a control volume to show that the energy transferred through the boundary into the control volume is equal to the sum of the storage energy change and energy dissipation. Write an integral equation using Gauss's theorem.

8.4 Derive the wave equation in Eq. (8.20) for a conductive medium; show Eq. (8.9) is a solution if k is complex, as given in Eq. (8.21). Many books use $\mathbf{E} = \mathbf{E}_0 \mathrm{e}^{\mathrm{i}(\omega t - \mathbf{k}\cdot\mathbf{r})}$ instead of Eq. (8.9) as the solution; how would you modify Eqs. (8.21) and (8.22)? Show that the complex refractive index must be defined as $\tilde{n} = n - \mathrm{i}\kappa$, where $\kappa \geq 0$.

8.5 Calculate the refractive index, the absorption coefficient, and the radiation penetration depth for the following materials, based on the dielectric function values at room temperature.

(a) Glass (SiO_2): $\varepsilon = 2.1 + \mathrm{i}0$ at 1 μm; $\varepsilon = 1.8 + \mathrm{i}0.004$ at 5 μm.
(b) Germanium: $\varepsilon = 21 + \mathrm{i}0.14$ at 1 μm; $\varepsilon = 16 + \mathrm{i}0.0003$ at 20 μm.
(c) Gold: $\varepsilon = -10 + \mathrm{i}1.0$ at 0.65 μm; $\varepsilon = -160 + \mathrm{i}2.1$ at 2 μm.

8.6 Consider a metamaterial with $\mu = -1 + \mathrm{i}0.01$ and $\varepsilon = -2 + \mathrm{i}0.01$; determine the refractive index and the extinction coefficient. Calculate the radiation penetration depth. Do a quick Internet search on negative index materials, and briefly describe what you have learned.

8.7 Find the magnetic field $\mathbf{H}$ for the wave given in Eq. (8.37). Show that the time-averaged Poynting vector is parallel to the x-axis. That is, the z-component of $\langle \mathbf{S} \rangle$ for such a wave vanishes. Briefly describe the features of an evanescent wave.

8.8 Write Planck's distribution in terms of wavenumber $\bar{\nu} = 1/\lambda$, i.e., the emissive power in terms of the wavenumber: $e_{b,\bar{\nu}}(\bar{\nu}, T)$. What is the most probable wavenumber in cm^{-1}? Compare your answer with the most probable wavelength obtained from Wien's displacement law in Eq. (8.45). Explain why the constants do not agree with each other. Cosmic background radiation can be treated as blackbody radiation at 2.7 K; what is the wavenumber corresponding to the maximum emissive power?

8.9 Based on the geometric parameters provided in Example 8.3 and neglecting the atmospheric effect, calculate the total intensity of the solar radiation arriving at earth's surface. Calculate the spectral intensity for solar radiation at 628 nm wavelength. A child used a lens to focus solar radiation to a small spot on a piece of paper and set fire this way. Does the beam focusing increase the intensity of the radiation? The lens diameter is 5 cm, and the distance between the lens and the paper is 2.5 cm. What are the focus size and the heat flux at the focus? Neglect the loss through the lens.

8.10 For a surface at $T = 1800$ K, with an emissivity of 0.6, what are the radiance temperatures at $\lambda = 0.65$ μm and 1.5 μm? If a conical hole is formed with a half-cone angle of 15°, what is the effective emittance and the radiance temperature at $\lambda = 0.65$ μm?

8.11 Derive Planck's law for a medium with a refractive index $n \neq 1$ in terms of the medium wavelength λ_m, $e_{b,\lambda_m}(\lambda_m, T)$ from Eq. (8.43). Assume that n is not a function of frequency (i.e., the medium is nondispersive) in the spectral region of interest. How does it compare with Eq. (8.44)?

8.12 Express Eq. (8.53) in terms of wavelength, $s_\lambda(\lambda, T)$. Find an expression of the entropy intensity for blackbody radiation, $L_\lambda(\lambda, T)$, and show that $L_\lambda(\lambda, T) = \frac{c}{4\pi} s_\lambda(\lambda, T)$.

8.13 Assume that all the blue light at λ in the range between 420 and 490 nm of solar radiation is scattered by the atmosphere and uniformly distributed over a solid angle of 4π sr. What are the monochromatic temperatures of the scattered radiation at $\lambda = 420$ and 490 nm?

8.14 A diode-pumped solid-state laser emits continuous-wave (cw) green light at a wavelength of 532 nm with a beam diameter of 1.1 mm. If the beam divergence is 2×10^{-7} sr, what would be the spot size at a distance of 100 m from the laser (without scattering)? If the output optical power is 2 mW and the spectral width is $\delta\lambda = 0.1$ nm (assuming a square function), what is the average intensity of the laser beam? Find the monochromatic radiation temperature of the laser when it is linearly polarized. Suppose the laser hits a rough surface and is scattered into the hemisphere isotropically. Find the radiation temperature of the scattered radiation and the entropy generation caused by scattering.

8.15 In Example 8.5, the two plates are blackbodies. Assume that the plates are diffuse-gray surfaces with emissivities ε_1' and ε_2'. Calculate the entropy generation rate in each plate per unit area. How will you determine the optimal efficiency for an energy conversion device installed at plate 2? For $T_1 = 1500$ K, $T_2 = 300$ K, and $\varepsilon_2' = 1$, plot the optimal efficiency versus ε_1'.

8.16 The concept of dilute blackbody radiation can be used as an alternative method to calculate the entropy generation of a two-plate problem as in Problem 8.15. Assume that the multiply reflected rays are at not in equilibrium with each other. Rather, each ray retains its original entropy and can be treated as having an effective temperature of T_1 or T_2 depending on which plate the ray is emitted from. How would you evaluate the entropy transfer from plate 1 to 2 and the entropy generation by each plate then?

8.17 Calculate the entropy generation rate per unit volume for Example 2.7. Further, calculate the entropy generated at each surface, assuming that surface 2 is at 300 K.

8.18 The conversion efficiency of thermophotovoltaic devices is wavelength dependent, and the optical constants are wavelength dependent as well. Perform a literature search to find some recent publications in this area. Use the entropy concept to determine the ultimate efficiency of a specific design. Based on your analysis, propose a few suggestions for further improvement of the particular design you have chosen.

8.19 Derive the Fresnel reflection coefficient for a TM wave, following the derivation given in the text for a TE wave.

8.20 Show that $\rho'_{\lambda,s} + \alpha'_{\lambda,s} = 1$, where $\rho'_{\lambda,s}$ is given in Eq. (8.73) and $\alpha'_{\lambda,s}$ is given in Eq. (8.75). Discuss why the z-component of the time-averaged Poynting vector must be continuous at the boundary but not the x-component.

8.21 For nonmagnetic lossy media with $\varepsilon_1 = \varepsilon'_1 + i\varepsilon''_1$ and $\varepsilon_2 = \varepsilon'_2 + i\varepsilon''_2$, expand Eq. (8.70b) and compare your results with Eq. (8.71).

8.22 For plane waves incident from air to a nonmagnetic material with $\varepsilon = -2 + i0$ (negative real), show that the reflectivity is always 1 regardless of the angle of incidence and the polarization. What can you say about k_{2z} and $\langle S_{2z} \rangle$? Is the wave in the medium a homogeneous wave or an evanescent wave?

8.23 The refractive index of glass is approximately 1.5 in the visible region. What is the Brewster angle for glass when light is incident from air? Calculate the reflectance and plot it against the incidence angle for p-polarization, s-polarization, and random polarization. Redo the calculation for incidence from glass to air, and plot the reflectance against the incidence angle. At what angle does total internal reflection begin and what is this angle called?

8.24 Denote the incidence angle at which the ratio of the reflectance for TM and TE waves is minimized as θ_M. For radiation incident from air to a medium with $n = 2$ and $\kappa = 1$, determine θ_M and compare it with the principle angle θ_P, at which the phase difference between the two reflection coefficients equals to $\pi/2$. [Hint: Use graphs to prove the existence of θ_M and θ_P.]

8.25 For incidence from glass with $n = 1.5$ to air, calculate the Goos–Hänchen phase shift δ for both TE and TM waves. Plot δ as a function of the incidence angle θ_1.

8.26 Show that the normal component of the time-averaged Poynting vector is zero in both the incident and transmitting media when total internal reflection occurs. Furthermore, derive Eq. (8.92).

8.27 Calculate the Goos–Hänchen lateral shift upon total internal reflection from a dielectric with $n = 2$ to air. Plot the lateral shift for both TE and TM waves as a function of θ_1. Discuss the cause and the physical significance of the lateral beam shift.

8.28 A perfect conductor can be understood based on the Drude free-electron model by neglecting the collision term. The dielectric function becomes $\varepsilon(\omega) = 1 - \omega_p^2/\omega^2$, where ω_p is the plasma frequency. For radiation incident from air to a perfect conductor, calculate the phase shift when $\omega = \omega_p/2$ for TE and TM waves as a function of the incidence angle. Use Eq. (8.93) to calculate the lateral beam shift for a TM wave and modify it for a TE wave. Do you expect a sign difference between the TE and TM waves?

8.29 Calculate and plot the emissivity (averaged over the two polarizations) versus the zenith angle for the materials and wavelengths given in Problem 8.5. Calculate and tabulate the normal and hemispherical emissivities for all cases.

8.30 Calculate the optical constants and the radiation penetration depth for either gold or silver at room temperature, using the Drude model, and plot them as functions of wavelength. In addition, calculate the reflectivity and plot it against

wavelength. Compare the results using the Hagen–Ruben equation. How will the scattering rate and the plasma frequency change if the temperature is raised to 600 K?

8.31 Calculate the normal emissivity of MgO from 2000 to 200 cm^{-1} (5 to 50 μm) using the Lorentz model with two oscillators having the following parameters: $\varepsilon_\infty = 3.01$; $\omega_1 = 401$ cm^{-1}, $\gamma_1 = 7.62$ cm^{-1}, and $S_1 = 6.6$; $\omega_2 = 640$ cm^{-1}, $\gamma_2 = 102.4$ cm^{-1}, and $S_2 = 0.045$. Can you develop a program to calculate the hemispherical emissivity and plot it against the normal emissivity for a comparison?

8.32 Find the Brewster angles for light incident from air to a NIM with (a) $\varepsilon_2 = -2$ and $\mu_2 = -2$, (b) $\varepsilon_2 = -1$ and $\mu_2 = -4$, and (c) $\varepsilon_2 = -8$ and $\mu_2 = -0.5$.

8.33 Use the online resources posted on the author's webpage [54] to calculate the absorption coefficient and normal reflectivity of intrinsic doped silicon for $0.5\ \mu m < \lambda < 25\ \mu m$.

8.34 First reproduce Fig. 8.21 for the dielectric function at 400 K and then calculate the dielectric function at 300 K for the same doping concentrations. Furthermore, calculate the real and imaginary parts of the refractive index of *n*-type doped silicon with a dopant concentration of 10^{19} cm^{-1} and plot them versus angular frequency.

8.35 Suppose a NIM can be described by Eqs. (8.135) and (8.136) with the following parameters: $\omega_p = 4.0 \times 10^{14}$ rad/s (i.e., $\lambda_p = 4.71\ \mu m$), $\omega_0 = 2.0 \times 10^{14}$ rad/s (i.e., $\lambda_0 = 9.42\ \mu m$), $\gamma = 0$, and $F = 0.785$. Assume a wave is propagating in such a medium in the region of $n < 0$ with a wavevector $\mathbf{k} = k_x \hat{\mathbf{x}}$, where $k_x = k = |n|\omega/c_0$. Show that the group velocity is in the negative *x*-direction. Also show that the Poynting vector is in the same direction as the group velocity.

8.36 Suppose a NIM can be described by Eqs. (8.135) and (8.136) with the following parameters: $\omega_p = 4.0 \times 10^{14}$ rad/s (i.e., $\lambda_p = 4.71\ \mu m$), $\omega_0 = 2.0 \times 10^{14}$ rad/s (i.e., $\lambda_0 = 9.42\ \mu m$, and $F = 0.5$. Calculate and plot the refractive index and the extinction coefficient in the spectral region from 2 to 15 μm, for $\gamma = 0$, 10^{12}, and 10^{13} rad/s.

8.37 What is a detector? What is a bolometer? What is a radiometer? If you are asked to buy a detector for infrared radiation measurement for the wavelength range between 2 and 16 μm, discuss how you would select a detector and why. [Hint: Do some online search.]

8.38 A bolometer uses a thin YBCO film on a sapphire substrate whose area is 2 mm × 2 mm, operating at 90 K. The thickness of the sapphire plate is 25 μm. The thermal conductance between the detector element and a heat sink is $G = 8.4 \times 10^{-5}$ W/K. The resistance $R_0(90\ K) = 200\ \Omega$ and $\beta = 1.5\ K^{-1}$. Assume the absorptance $\alpha = 0.7$. Calculate the time constant for different bias currents, $I = 0.1, 0.2$ and 0.3 mA. Calculate and plot the detector responsivity as a function of modulation frequency ω_f between 0.1 and 10 Hz for each bias current value given above. Neglect the heat capacity of the YBCO film. The density and specific heat of sapphire at the operating temperature are $\rho = 3970$ kg/m^3 and $c_p = 102$ J/kg K.

References

1. J.R. Howell, M.P. Mengüç, R. Siegel, *Thermal Radiation Heat Transfer*, 6th edn. (CRC Press, New York, 2016)
2. M.F. Modest, *Radiative Heat Transfer*, 3rd edn. (Academic Press, New York, 2013)
3. E.S. Barr, Historical survey of the early development of the infrared spectral region. Am. J. Phys. **28**, 42–54 (1960)
4. M. Planck, *The Theory of Heat Radiation* (Dover Publications, New York, 1959) (originally published in German in 1913 and translated to English by M. Masius in 1914)
5. J.D. Jackson, *Classical Electrodynamics*, 3rd edn. (Wiley, New York, 1998)
6. D.J. Griffiths, Introduction to Electrodynamics, 4th edn. (Cambridge University Press, Cambridge, UK, 2017)
7. A.J. Kong, *Electromagnetic Wave Theory*, 2nd edn. (Wiley, New York, 1990)
8. M. Born, E. Wolf, *Principles of Optics*, 7th edn. (Cambridge University Press, Cambridge, UK, 1999)
9. C.F. Bohren, D.R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983)
10. NIST, http://physics.nist.gov/cuu/index.html. Last Accessed 24 Dec 2018
11. Z.M. Zhang, K. Park, Group front and group velocity in a dispersive medium upon refraction from a nondispersive medium. J. Heat Transfer **126**, 244–249 (2004)
12. P. Yeh, C. Gu, *Optics of Liquid Crystal Displays*, 2nd edn. (Wiley, New York, 2010)
13. T.J. Chui, J.A. Kong, Time-domain electromagnetic energy in a frequency-dispersive left-handed medium. Phys. Rev. B **70**, 205106 (2004)
14. J.M. Zhao, Z.M. Zhang, Electromagnetic energy storage and power dissipation in nanostructures. J. Quant. Spectrosc. Radiat. Transfer **151**, 49–57 (2015)
15. E.D. Palik (ed.), *Handbook of the Optical Constants of Solids*, vol. I, II, and III (Academic Press, San Diego, CA, 1998)
16. Z.M. Zhang, Surface temperature measurement using optical techniques. Annu. Rev. Heat Transfer **11**, 351–411 (2000)
17. T.J. Quinn, J.E. Martin, A radiometric determination of the Stefan-Boltzmann constants and thermodynamic temperatures between –40 °C and +100 °C. Philos. Trans. Royal Soc. London, Ser. A **316**, 85–189 (1985)
18. D.A. Pearson, Z.M. Zhang, Thermal-electrical modeling of absolute cryogenic radiometers. Cryogenics **39**, 299–309 (1999)
19. Z.M. Zhang, B.K. Tsai, G. Machin, *Radiometric Temperature Measurements*, I. Fundamentals & II. Applications (Academic Press, Amsterdam, 2010)
20. Y.H. Zhou, Y.-J. Shen, Z.M. Zhang, B.K. Tsai, D.P. DeWitt, A Monte Carlo model for predicting the effective emissivity of the silicon wafer in rapid thermal processing furnaces. Int. J. Heat Mass Transfer **45**, 1945–1949 (2002)
21. M.J. Lang, S.M. Block, Resource Letter: LBOT-1: Laser-based optical tweezers. Am. J. Phys. **71**, 201–215 (2003)
22. R. Petela, Exergy of heat radiation. J. Heat Transfer **86**, 187–192 (1964)
23. P.T. Landsberg, G. Tonge, Thermodynamic energy conversion efficiencies. J. Appl. Phys. **51**, R1–R20 (1980)
24. E.J. Tervo, E. Bagherisereshki, Z.M. Zhang, Near-field radiative thermoelectric energy converters: a review. Frontiers Energy **12**, 5–21 (2018)
25. C.E. Mungan, Radiation thermodynamics with application to lasing and fluorescent cooling. Am. J. Phys. **73**, 315–322 (2005)
26. C. Essex, D.C. Kennedy, R.S. Berry, How hot is radiation? Am. J. Phys. **71**, 969–978 (2003)
27. M. Caldas, V. Semiao, Entropy generation through radiative transfer in participating media: analysis and numerical computation. J. Quant. Spectrosc. Radiat. Transfer **96**, 423–437 (2005)
28. Z.M. Zhang, S. Basu, Entropy flow and generation in radiative transfer between surfaces. Int. J. Heat Mass Transfer **50**, 702–712 (2007)

29. T.J. Bright, Z.M. Zhang, Entropy generation in thin films evaluated from phonon radiative transport. J. Heat Transfer **132**, 101301 (2010)
30. Z.M. Zhang, C.J. Fu, Q.Z. Zhu, Optical and thermal radiative properties of semiconductors related to micro/nanotechnology. Adv. Heat Transfer **37**, 179–296 (2003)
31. Z.M. Zhang, H. Ye, Measurements of radiative properties of engineered micro-/nanostructures. Annu. Rev. Heat Transfer **16**, 351–411 (2013)
32. X.L. Liu, L.P. Wang, Z.M. Zhang, Near-field thermal radiation: recent progress and outlook. Nanoscale Microscale Thermophys. Eng. **19**, 98–126 (2015)
33. B. Zhao, Z.M. Zhang, Design of optical and radiative properties of solids, in *Handbook of Thermal Science and Engineering: Radiative Heat Transfer,* ed. by F.A. Kulachi (Springer Nature, 2017), pp. 1023–1068
34. D. Thompson, L. Zhu, R. Mittapally, S. Sadat, Z. Xing, P. McArdle, M.M. Qazilbash, P. Reddy, E. Meyhofer, Hundred-fold enhancement in far-field radiative heat transfer over the blackbody limit. Nature **561**, 216–221 (2018)
35. B. Salzberg, A note on the significance of power reflections. Am. J. Phys. **16**, 444–446 (1948); Z.M. Zhang, Reexamination of the transmittance formula of a lamina. J. Heat Transfer **119**, 645–647 (1997)
36. P. Halevi, A. Mendoza-Hernández, Temporal and spatial behavior of the Poynting vector in dissipative media: refraction from vacuum into a medium. J. Opt. Soc. Am. **71**, 1238–1242 (1981)
37. C.J. Fu, Z.M. Zhang, P.N. First, Brewster angle with a negative index material. Appl. Opt. **44**, 3716–3724 (2005)
38. H.A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, Englewood Cliffs, NJ, 1984)
39. H.K.V. Lotsch, Beam displacement at total reflection: The Goos-Hänchen effect, I, II, III, IV. Optik **32**(2), 116–137; (3), 189–204; (4), 299–319; (6), 553–569 (1970–1971)
40. A. Puri, J.L. Birman, Goos-Hänchen beam shift at total internal reflection with application to spatially dispersive media. J. Opt. Soc. Am. A **3**, 543–549 (1986)
41. D.-K. Qing, G. Chen, Goos-Hänchen shifts at the interfaces between left- and right-handed media. Opt. Lett. **29**, 872–874 (2004); X.L. Hu, Y.D. Huang, W. Zhang, D.-K. Qing, J.D. Peng, Opposite Goos-Hänchen shifts for transverse-electric and transverse-magnetic beams at the interface associated with single-negative materials. Opt. Lett., **30**, 899–901 (2005)
42. H.M. Lai, F.C. Cheng, W.K. Tang, Goos-Hänchen effect around and off the critical angle. J. Opt. Soc. Am. A **3**, 550–557 (1986)
43. I.V. Shadrivov, A.A. Zharov, Y.S. Kivshar, Giant Goos-Hänchen effect at the reflection from left-handed metamaterials. Appl. Phys. Lett. **83**, 2713–2715 (2003)
44. X. Chen, C.-F. Li, Lateral shift of the transmitted light beam through a left-handed slab. Phys. Rev. E **69**, 066617 (2004)
45. Y.J. Shen, Q.Z. Zhu, Z.M. Zhang, A scatterometer for measuring the bidirectional reflectance and transmittance of semiconductor wafers with rough surfaces. Rev. Sci. Instrum. **74**, 4885–4892 (2003)
46. W.C. Snyder, Z. Wan, X. Li, Thermodynamic constraints on reflectance reciprocity and Kirchhoff's law. Appl. Opt. **37**, 3464–3470 (1998)
47. T.L. Bergman, A.S. Lavine, F.P. Incropera, D.P. DeWitt, *Fundamentals of Heat and Mass Transfer*, 8th edn. (Wiley, New York, 2017)
48. L. Zhu, S. Fan, Near-complete violation of detailed balance in thermal radiation. Phys. Rev. B **90**, 220301(R) (2014)
49. L.P. Wang, S. Basu, Z.M. Zhang, Direct and indirect methods for calculating thermal emission from layered structures with nonuniform temperatures. J. Heat Transfer **133**, 072701 (2011)
50. Y.S. Touloukian, D.P. DeWitt, Thermal Radiative Properties, in *Thermophysical Properties of Matter,* TPRC Data Series, vol. 7, 8, and 9 ed. by Y.S. Touloukian, C.Y. Ho (IFI Plenum, New York, 1970–1972)
51. LBL-CXRO, http://henke.lbl.gov/optical_constants. Last Accessed 25 Dec 2018
52. W.M. Toscano, E.G. Cravalho, Thermal radiation properties of the noble metals at cryogenic temperatures. J. Heat Transfer **98**, 438–445 (1976)

53. F. Gervais, B. Piriou, Temperature dependence of transverse-and longitudinal-optic modes in TiO_2 (rutile). Phys. Rev. B **10**, 1642–1654 (1974)
54. Z.M. Zhang, http://zhang-nano.gatech.edu. Last Accessed 15 Jan 2019
55. W.G. Spitzer, D. Kleinman, D. Walsh, Infrared properties of hexagonal silicon carbide. Phys. Rev. **113**, 127–132 (1959)
56. S. Baroni, S. de Gironcoli, A. Dal Corso, P. Giannozzi, Phonons and related crystal properties from density-functional perturbation theory. Rev. Mod. Phys. **73**, 515–562 (2001)
57. G.-M. Rignanese, X. Gonze, G. Jun, K. Cho, A. Pasquarello, First-principles investigation of high-k dielectrics: Comparison between the silicates and oxides of hafnium and zirconium. Phys. Rev. B **69**, 184301 (2004)
58. H. Bao, B. Qiu, Y. Zhang, X. Ruan, A first-principles molecular dynamics approach for predicting optical phonon lifetimes and far-infrared reflectance of polar materials. J. Quant. Spectrosc. Radiat. Transfer **113**, 1683–1688 (2012)
59. P.J. Timans, The thermal radiative properties of semiconductors, in *Advances in Rapid Thermal and Integrated Processing*, ed. by F. Roozeboom (Kluwer Academic Publishers, The Netherlands, 1996), pp. 35–101
60. L.X. Benedict, E.L. Shirley, R.B. Bohn, Theory of optical absorption in diamond, Si, Ge, and GaAs. Phys. Rev. B **57**, R9385–R9387 (1998)
61. M. Rohlfing, S.G. Louie, Electron-hole excitations and optical spectra from first principles. Phys. Rev. B **62**, 4927–4944 (2000)
62. C.J. Fu, Z.M. Zhang, Nanoscale radiation heat transfer for silicon at different doping levels. Int. J. Heat Mass Transfer **49**, 1703–1718 (2006)
63. S. Basu, B.J. Lee, Z.M. Zhang, Infrared radiative properties of heavily doped silicon at room temperature", J. Heat Transfer, 132, 023301, 2010; "Near-field radiation calculated with an improved dielectric function model for doped silicon. J. Heat Transfer **132**, 023302 (2010)
64. Z.M. Zhang, A. Frenkel, Thermal and nonequilibrium responses of superconductors for radiation detectors. J. Supercond. **7**, 871–884 (1994)
65. A.R. Kumar, Z.M. Zhang, V.A. Boychev, D.B. Tanner, L.R. Vale, D.A. Rudman, Far-infrared transmittance and reflectance of $YBa_2Cu_3O_{7-\delta}$ films on Si substrates. J. Heat Transfer **121**, 844–851 (1999)
66. J. Bardeen, L.N. Cooper, J.R. Schrieffer, Theory of superconductivity. Phys. Rev. **108**, 1175–1204 (1957)
67. R.A. Shelby, D.A. Smith, S. Schultz, Experimental verification of a negative index of refraction. Science **292**, 77–79 (2001)
68. J.B. Pendry, Negative refraction makes a perfect lens. Phys. Rev. Lett. **85**, 3966–3969 (2000)
69. Z.M. Zhang, C.J. Fu, Unusual photon tunneling in the presence of a layer with a negative refractive index. Appl. Phys. Lett. **80**, 1097–1099 (2002)
70. S.A. Ramakrishna, Physics of negative refractive index materials. Rep. Prog. Phys. **68**, 449–521 (2005)
71. X. Zhang, Z. Liu, Superlenses to overcome the diffraction limit. Nature Mat. **7**, 435–441 (2008)
72. C.J. Fu, Z.M. Zhang, D.B. Tanner, Energy transmission by photon tunneling in multilayer structures including negative index materials. J. Heat Transfer **127**, 1046–1052 (2005)
73. K.G. Ramanathan, S.H. Yen, High-temperature emissivities of copper, aluminum, and silver. J. Opt. Soc. Am. **67**, 32–38 (1977)
74. S.X. Cheng, P. Cebe, L.M. Hanssen, D.M. Riffe, A.J. Sievers, Hemispherical emissivity of V, Nb, Ta, Mo, and W from 300 to 1000 K. J. Opt. Soc. Am. B **4**, 351–356 (1987)
75. J. Hameury, B. Hay, J.R. Filtz, Measurement of total hemispherical emissivity using a calorimetric technique. Int. J. Thermophys. **28**, 1607–1620 (2007)
76. D. Giulietti, A. Gozzini, M. Lucchesi, R. Stampacchia, A calorimetric technique for measuring total emissivity of solid materials and coatings at low temperatures. J. Phys. D Appl. Phys. **12**, 2027–2036 (1979)
77. D.A. Pinnow, T.C. Rich, Development of a calorimetric method for making precision optical absorption measurements. Appl. Opt. **12**, 984–992 (1973)

78. H.B. Rosenstock, D.A. Gregory, J.A. Harrington, Infrared bulk and surface absorption by nearly transparent crystals. Appl. Opt. **15**, 2075–2079 (1976)
79. D. Bunimovich, L. Nagli, A. Katzir, Absorption measurements of mixed silver halide crystals and fibers by laser calorimetry. Appl. Opt. **33**, 117–119 (1994)
80. R.M.A. Azzam, N.M. Bashara, *Ellipsometry and Polarized Light* (North Holland, Amsterdam, 1987)
81. A. Cezairliyan, S. Krishanan, J.L. McClure, Simultaneous measurements of normal spectral emissivity by spectral radiometry and laser polarimetry at high temperatures in millisecond-resolution pulse-heating experiments: Application to molybdenum and tungsten. Int. J. Thermophys. **17**, 1455–1473 (1996)
82. S. Schöche, T. Hofmann, R. Korlacki, T.E. Tiwald, M. Schubert, Infrared dielectric anisotropy and phonon modes of rutile TiO_2. J. Appl. Phys. **113**, 164102 (2013)
83. C.A. Mitchell, Emissivity of globar. J. Opt. Soc. Am. **52**, 341–342 (1962)
84. M.F. Kimmitt, J.E. Walsh, C.L. Platt, K. Miller, M.R.F. Jensen, Infrared output from a compact high pressure arc source. Infrared Phys. Technol. **37**, 471–477 (1996)
85. R.A. Lewis, A review of terahertz sources. J. Phys. D Appl. Phys. **47**, 374001 (2014)
86. R.H. Kingston, *Detection of Optical and Infrared Radiation* (Springer, Berlin, 1978)
87. R.J. Keyes (ed.), *Optical and Infrared Detectors* (Springer, Berlin, 1980)
88. Z.M. Zhang, C.J. Zhu, L.M. Hanssen, Absolute detector calibration applied to nonlinearity error correction in FT-IR measurements. Appl. Spectrosc. **51**, 576–579 (1997)
89. H. Lee, A. Bryson, Z. Zhang, Measurement and modeling of the emittance of silicon wafers with anisotropic roughness. Int. J. Thermophys. **28**, 918–933 (2007)
90. X.J. Wang, J.D. Flicker, B.J. Lee, W.J. Ready, Z.M. Zhang, Visible and near-infrared radiative properties of vertically aligned multi-walled carbon nanotubes. Nanotechnology **20**, 215704 (2009)
91. P. Yang, C.Y. Chen, Z.M. Zhang, A dual-layer structure with record-high solar reflectance for daytime radiative cooling. Sol. Energy **169**, 316–324 (2018)
92. L.M. Hanssen, K.A. Snail, Integrating spheres for mid- and near-infrared reflection spectroscopy, in *Handbook of Vibrational Spectroscopy*, ed. by J.M. Chalmers, P.R. Griffiths (New York, Wiley, 2002), pp. 1175–1192
93. P.R. Griffiths, J.A. de Haseth, *Fourier Transform Infrared Spectrometer* (Wiley, New York, 1986)
94. M.I. Flik, Z.M. Zhang, Influence of nonequivalent detector responsivity on FT-IR photometric accuracy. J. Quant. Spectrosc. Radiat. Transfer **47**, 293–303 (1992)
95. S.G. Kaplan, L.M. Hanssen, R.U. Datla, Testing the radiometric accuracy of Fourier transform infrared transmittance measurements. Appl. Opt. **36**, 8896–8908 (1997)
96. T.J. Bright, J.I. Watjen, Z.M. Zhang, C. Muratore, A.A. Voevodin, Optical properties of HfO_2 thin films deposited by magnetron sputtering: from the visible to the far-infrared. Thin Solid Films **520**, 6793–6802 (2012); T.J. Bright, J.I. Watjen, Z.M. Zhang, C. Muratore, A.A. Voevodin, D.I. Koukis, D.B. Tanner, D.J. Arenas, Infrared optical properties of amorphous and nanocrystalline Ta_2O_5 thin films. J. Appl. Phys. **114**, 083515 (2013)
97. L.P. Wang, S. Basu, Z.M. Zhang, Direct measurement of thermal emission from a Fabry-Perot cavity resonator. J. Heat Transfer **134**, 072701 (2012); L.P. Wang, Z.M. Zhang, Measurement of coherent thermal emission due to magnetic polaritons in subwavelength microstructures. J. Heat Transfer **135**, 091014 (2013)
98. P.Y. Barnes, E.A. Early, A.C. Parr, *Spectral Reflectance* (NIST Special Publication, Washington DC, 1998), pp. 250–248
99. Q.Z. Zhu, Z.M. Zhang, Anisotropic slope distribution and bidirectional reflectance of a rough silicon surface. J. Heat Transfer **126**, 985–993 (2004); H.J. Lee, Z.M. Zhang, Measurement and modeling of the bidirectional reflectance of SiO_2 coated Si surfaces. Int. J. Thermophys. **27**, 820–839 (2006)
100. E. Franke, C.L. Trimble, M.J. DeVries, J.A. Woollam, M. Schubert, F. Frost, Dielectric function of amorphous tantalum oxide from the far infrared to the deep ultraviolet spectral region measured by spectroscopic ellipsometry. J. Appl. Phys. **88**, 5166–5174 (2000)