EXPLORATION OF THE SOLAR SYSTEM BY INFRARED REMOTE SENSING

Second edition

R. A. HANEL
B. J. CONRATH
D. E. JENNINGS
and
R. E. SAMUELSON

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1992

Second edition 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 11/14 pt *System* LaTeX 2_ε [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Exploration of the solar system by infrared remote sensing / R. A. Hanel . . . [et al.].–2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0 521 81897 4
1. Planets–Remote sensing. 2. Infrared astronomy. 3. Outer space–Exploration. I. Hanel, R. A.
QB603.I52 E88 2002

523.2–dc21 2002025665

ISBN 0 521 81897 4 hardback

Contents

1

Foundation of radiation theory

In this chapter we review the physical foundation of remote sensing. Except for possible gravitational effects, information accessible to a distant observer must be sensed as electromagnetic radiation, either in the form of reflected or refracted solar or stellar radiation, or in the form of thermal or nonthermal emission. We restrict the discussion to passive techniques. Active methods, involving the generation of electromagnetic radiation (radar, lidar), are not explicitly treated. However, the physical principles discussed in this text are equally applicable to passive and active methods. In either case a discussion of the measurement and interpretation of remotely sensed data must be based on electromagnetic theory. In Section 1.1 we begin with that theory by reviewing Maxwell's equations. The application of the principle of energy conservation to Maxwell's equations leads to the Poynting theorem with the Poynting vector describing radiative energy transport; this is discussed in Section 1.2. However, the Poynting vector does not characterize more complex phenomena, such as reflection, refraction, polarization, or interference; all of these phenomena play significant roles in many aspects of remote sensing. Their study requires, first, a derivation of the wave equation from Maxwell's formulas, and second, finding appropriate solutions for the electric and magnetic field vectors; this is the subject of Section 1.3. Polarization is briefly reviewed in Section 1.4. Effects of electromagnetic waves striking an interface between two media and the conditions that must be satisfied at the boundary are treated in Section 1.5. The derived conditions are then applied to the boundary to find expressions for reflected and refracted waves. These expressions, the Fresnel equations, are discussed in Section 1.6. The same boundary conditions are used again in Section 5.6 to describe the behavior of thin films employed in many ways in remote sensing instruments. The Planck function is introduced in Section 1.7. In Section 1.8, we return to the Poynting vector in a discussion of quantities used in the theory of radiative transfer, such as spectral intensity and radiative flux.

1.1 Maxwell's equations

Electromagnetic radiation between the red limit of the visible spectrum and the microwave region is called the infrared. In round numbers the infrared covers the spectral range from 1 to 1000 μm. Although only the range from 0.35 to 0.75 μm is truly visible to the human eye, the region between 0.75 and 1 μm is often considered as a part of the 'visible' spectrum because many detectors common to that spectral domain, such as conventional photomultipliers, photographic film, and charge-coupled silicon devices, work well up to about 1 μm. At the far end of the infrared spectrum, tuned circuits, waveguides, and other elements associated with radio and microwave technology become the commonly employed detection tools.

Whatever the wavelength, electromagnetic radiation obeys the laws expressed by Maxwell's equations. These equations describe the interrelationship of electric and magnetic quantities by field action, in contrast to action at a distance, which up to Maxwell's time (1873) was the generally accepted point of view. The field concept goes back to Michael Faraday. In all likelihood, the concept suggested itself to him in experiments with magnets and iron filings in which lines of force become almost an observable reality. However, it was left to James Clerk Maxwell to give the field concept a far-reaching and elegant mathematical formulation. Fifteen years after the publication of Maxwell's treatise (1873), Heinrich Hertz (1888) discovered electromagnetic waves, an experimental verification of Maxwell's theory.

In differential form, using the rationalized system and vector notation, the first pair of Maxwell's equations is (e.g. Sommerfeld, 1952):

$$\dot{\mathbf{D}} + \mathbf{J} = \boldsymbol{\nabla} \times \mathbf{H} \tag{1.1.1}$$

and

$$\dot{\mathbf{B}} = -\boldsymbol{\nabla} \times \mathbf{E}, \tag{1.1.2}$$

where $\mathbf{D}$ and $\mathbf{B}$ are the electric displacement and magnetic induction, and $\mathbf{E}$ and $\mathbf{H}$ the electric and magnetic field strengths, respectively; $\mathbf{J}$ is the current density. The dot symbolizes differentiation with respect to time. Definitions of the curl ($\boldsymbol{\nabla}\times$) and the divergence ($\boldsymbol{\nabla}\,\cdot$) operators are given in Appendix 1. The concept of the electric displacement was introduced by Maxwell. The first equation includes Ampère's law and the second represents Faraday's law of induction.

Besides the main equations (1.1.1) and (1.1.2), two more expressions are traditionally considered part of Maxwell's equations,

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \rho \tag{1.1.3}$$

and

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0. \tag{1.1.4}$$

Equation (1.1.3) defines the electric charge density, ρ, while Eq. (1.1.4) states the nonexistence of magnetic charges or monopoles. Strictly from symmetry considerations of Maxwell's equations one may be led to postulate the existence of magnetic charges, but despite many attempts none has been found.

By applying the divergence operator to Eq. (1.1.1) and substituting ρ for $\boldsymbol{\nabla} \cdot \mathbf{D}$, one arrives at the electric continuity equation,

$$\dot{\rho} + \boldsymbol{\nabla} \cdot \mathbf{J} = 0, \tag{1.1.5}$$

which states the conservation of electric charge: a change in the charge density of a volume element must be associated with a current flow across the boundary of that arbitrarily chosen element. The continuity equation in fluid dynamics is an analogous expression of the conservation of mass.

In order to study the interaction of matter with electric and magnetic fields, three material constants are introduced: the electric conductivity, σ,

$$\mathbf{J} = \sigma \mathbf{E}, \tag{1.1.6}$$

the dielectric constant, ε,

$$\mathbf{D} = \varepsilon \mathbf{E}, \tag{1.1.7}$$

and the magnetic permeability, μ,

$$\mathbf{B} = \mu \mathbf{H}. \tag{1.1.8}$$

Equation (1.1.6) is a form of Ohm's law. Since $\mathbf{J}$ is the current density (A m^{-2}) and $\mathbf{E}$ the electric field strength (V m^{-1}), σ is expressed in Ω^{-1} m^{-1}. The inverse conductivity is the resistivity. In the rationalized system the dielectric constant is conveniently written

$$\varepsilon = \varepsilon_0 \varepsilon_{\mathrm{rel}}, \tag{1.1.9}$$

where ε_0 is the dielectric constant of free space (see Appendix 2 for numerical values) and $\varepsilon_{\mathrm{rel}}$ is a dimensionless quantity, which is unity for free space and which has the same value as the dielectric constant in the Gaussian system of units. The permeability is

$$\mu = \mu_0 \mu_{\mathrm{rel}}, \tag{1.1.10}$$

where μ_0 represents the permeability of free space. The relative permeability is unity for free space, larger than unity for paramagnetic materials, and less than unity for diamagnetic substances.

Maxwell's equations are linear. However, the parameters that describe material properties may become nonlinear in exceptionally strong fields, such as in powerful lasers. In these cases nonlinear terms have to be included. The linear material equations, Eqs. (1.1.6) to (1.1.8), are not applicable to ferroelectric or ferromagnetic substances where the relationship between the electric field strength, **E**, and the electric displacement, **D**, or between the magnetic field strength, **H**, and the magnetic induction, **B**, are not only nonlinear, but show hysteresis effects as well. In any case, Maxwell's equations are the foundation of electromagnetism, which includes optics and infrared physics.

1.2 Conservation of energy and the Poynting vector

The Poynting theorem expresses the conservation of energy in electromagnetism. If one takes the scalar product of Eq. (1.1.1) with **E** and of Eq. (1.1.2) with **H**, and adds the results one finds

$$\mathbf{H} \cdot \dot{\mathbf{B}} + \mathbf{E} \cdot \dot{\mathbf{D}} + \mathbf{E} \cdot \mathbf{J} = \mathbf{E} \cdot (\boldsymbol{\nabla} \times \mathbf{H}) - \mathbf{H} \cdot (\boldsymbol{\nabla} \times \mathbf{E}). \tag{1.2.1}$$

With the vector identity

$$\mathbf{E} \cdot (\boldsymbol{\nabla} \times \mathbf{H}) - \mathbf{H} \cdot (\boldsymbol{\nabla} \times \mathbf{E}) \equiv -\boldsymbol{\nabla} \cdot (\mathbf{E} \times \mathbf{H}) \tag{1.2.2}$$

and the definition

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \tag{1.2.3}$$

one obtains

$$\mathbf{H} \cdot \dot{\mathbf{B}} + \mathbf{E} \cdot \dot{\mathbf{D}} + \mathbf{E} \cdot \mathbf{J} + \boldsymbol{\nabla} \cdot \mathbf{S} = 0. \tag{1.2.4}$$

This is the Poynting theorem; **S** is the Poynting vector. The first two terms in Eq. (1.2.4) represent rate of change of the magnetic and electric energy densities in the field. The third term, $\mathbf{E} \cdot \mathbf{J}$, describes the energy dissipated by the motion of electric charges. Generally, this motion results in Joule heating and, therefore, in losses to the energy stored in the field. The last term, $\boldsymbol{\nabla} \cdot \mathbf{S}$, represents the net flow of electromagnetic energy across the boundaries of the chosen volume. All terms of Eq. (1.2.4) are measured in J m^{-3} s^{-1}, which is energy per unit volume and unit time. Since the divergence operator corresponds to a differentiation with respect to space coordinates, the units of $\mathbf{S}$ are J m^{-2} s^{-1} or W m^{-2}, thus $\mathbf{S}$ is an energy flux through a surface element.

The definition of the Poynting vector, Eq. (1.2.3), requires that $\mathbf{S}$ be orthogonal to both $\mathbf{E}$ and $\mathbf{H}$. In order to better visualize the relative orientation of these three vectors, we align a Cartesian coordinate system so that the x-axis coincides with the direction of the Poynting vector. The components of $\mathbf{S}$ along the y- and z-axes, as well as the components of $\mathbf{E}$ and $\mathbf{H}$ in the direction of the x-axis, must then be zero: $S_y = S_z = E_x = H_x = 0$. The vectors $\mathbf{E}$ and $\mathbf{H}$ do not have components in the direction of energy transport represented by $\mathbf{S}$. Electromagnetic waves are transverse, in contrast to sound waves, which are longitudinal. To investigate the relative orientation between $\mathbf{E}$ and $\mathbf{H}$, we use the second of Maxwell's equations (Eq. 1.1.2) and the explicit expression of the curl operator (see Appendix 1). With the assumption that μ is constant and E_x and H_x equal zero, one obtains one scalar equation for each of the $\hat{\mathbf{j}}$- and $\hat{\mathbf{k}}$-directions ($\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are the unit vectors in the x-, y-, and z-directions):

$$\mu \frac{\partial H_y}{\partial t} = \frac{\partial E_z}{\partial x}; \qquad \mu \frac{\partial H_z}{\partial t} = -\frac{\partial E_y}{\partial x}. \tag{1.2.5}$$

Except for a static field, which is not of interest in this context, Eq. (1.2.5) indicates that H_y must be zero if E_z vanishes and, conversely, H_z must disappear when E_y is zero. These conditions require $\mathbf{E}$ and $\mathbf{H}$ to be at right angles to each other; $\mathbf{E}$, $\mathbf{H}$, and $\mathbf{S}$ form a right-handed, orthogonal system of vectors.

1.3 Wave propagation

In an isotropic, stationary medium, the material constants σ, ε, and μ are uniform and constant scalars. The first pair of Maxwell's equations may then be stated:

$$\varepsilon \dot{\mathbf{E}} + \sigma \mathbf{E} = \boldsymbol{\nabla} \times \mathbf{H} \tag{1.3.1}$$

and

$$\mu \dot{\mathbf{H}} = -\boldsymbol{\nabla} \times \mathbf{E}. \tag{1.3.2}$$

If one differentiates Eq. (1.3.1) with respect to time and multiplies by μ, one obtains

$$\varepsilon\mu\ddot{\mathbf{E}} + \sigma\mu\dot{\mathbf{E}} = \mu\frac{\partial}{\partial t}(\boldsymbol{\nabla}\times\mathbf{H}). \tag{1.3.3}$$

Application of the curl operator to Eq. (1.3.2) yields

$$\mu\boldsymbol{\nabla}\times\dot{\mathbf{H}} = -\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{E}). \tag{1.3.4}$$

For a medium at rest the order of differentiation with respect to space and time may be interchanged. Applying the vector identity

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{E}) \equiv \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E}) - \boldsymbol{\nabla}^2\mathbf{E} \tag{1.3.5}$$

and assuming the medium to be free of electric charges $[(\boldsymbol{\nabla}\cdot\mathbf{E}) = 0]$ leads to

$$\varepsilon\mu\ddot{\mathbf{E}} + \sigma\mu\dot{\mathbf{E}} = \nabla^2\mathbf{E}. \tag{1.3.6}$$

The Laplace operator, ∇^2, is defined in Appendix 1. This partial differential equation characterizes wave and relaxation phenomena. Again, we assume the x-axis to be aligned with the Poynting vector, so that $E_x = 0$. To simplify matters further, we rotate the coordinate system around the x-axis until the y-axis coincides with the direction of the electric field strength, so that $E_z = 0$ also. Only the y-component of $\mathbf{E}$ remains and Eq. (1.3.6) becomes a scalar equation for the unknown $E_y(x, t)$,

$$\varepsilon\mu\ddot{E}_y + \sigma\mu\dot{E}_y = E_y''. \tag{1.3.7}$$

We denote differentiation with respect to time by a dot and with respect to a space coordinate (in this case with respect to x) by a prime. The assumption $E_y = T(t)X(x)$ separates the variables,

$$\varepsilon\mu\frac{\ddot{T}}{T} + \sigma\mu\frac{\dot{T}}{T} = \frac{X''}{X} = -k^2. \tag{1.3.8}$$

Since the left side depends only on the variable t and the middle part only on the variable x, Eq. (1.3.8) can only be satisfied if the left and the middle part equal a constant, $-k^2$. The reason for choosing a negative square and the physical meaning of k will become apparent later. With the introduction of k, Eq. (1.3.8) yields two ordinary differential equations:

$$\varepsilon\mu\ddot{T} + \sigma\mu\dot{T} + k^2 T = 0 \tag{1.3.9}$$

and

$$X'' + k^2 X = 0. \tag{1.3.10}$$

A solution of Eq. (1.3.10) is readily shown to be

$$X = A\,\mathrm{e}^{\pm \mathrm{i}kx}. \tag{1.3.11}$$

The amplitude A is not defined by Eq. (1.3.10); it is determined by boundary conditions. For convenience we use notation with complex arguments in the treatment of wave phenomena. To simplify notation we omit the amplitudes but reintroduce them when needed. To solve Eq. (1.3.9) one may assume a solution of exponential form,

$$T = \mathrm{e}^{pt}, \tag{1.3.12}$$

which yields a characteristic equation for p,

$$\varepsilon\mu p^2 + \sigma\mu p + k^2 = 0. \tag{1.3.13}$$

We make two choices for p. In the first case we find the roots of Eq. (1.3.13) for p, assuming the coefficients ε, μ, σ, and k to be real quantities. Later, we will be interested in periodic solutions of Eq. (1.3.12), which imply $p = \pm\,\mathrm{i}\omega$. In that case, if $\sigma \neq 0$, at least one of the coefficients must be complex. The roots of Eq. (1.3.13) for p are

$$p = -\frac{\sigma}{2\varepsilon} \pm \left(\frac{\sigma^2}{4\varepsilon^2} - \frac{k^2}{\varepsilon\mu}\right)^{\frac{1}{2}}. \tag{1.3.14}$$

The parameter p is complex because the term with σ^2 in the parentheses is generally smaller than the term containing k^2,

$$E_y = \exp\left[-\frac{\sigma t}{2\varepsilon} \pm \mathrm{i}\left(\frac{k^2}{\varepsilon\mu} - \frac{\sigma^2}{4\varepsilon^2}\right)^{\frac{1}{2}} t\right] \exp\left(\pm\,\mathrm{i}kx\right). \tag{1.3.15}$$

E_y is an oscillating function of t and x. Before we discuss the physical content of Eq. (1.3.15) we consider the meaning of some of the quantities involved. It is convenient to introduce new terms pertinent to the description of optical phenomena in the infrared. Consider the inverse product $\varepsilon^{-1}\mu^{-1}$, which has the dimension of the square of a velocity, m^2 s^{-2}. This is the propagation velocity, v, of electromagnetic waves in a medium with dielectric constant ε and permeability μ. For free space this velocity is the velocity of light, c. We have

$$v = (\varepsilon\mu)^{-\frac{1}{2}}; \qquad c = (\varepsilon_0\mu_0)^{-\frac{1}{2}}. \tag{1.3.16}$$

Consequently

$$\frac{c}{v} = \left(\frac{\varepsilon\mu}{\varepsilon_0\mu_0}\right)^{\frac{1}{2}} = (\varepsilon_{\mathrm{rel}}\mu_{\mathrm{rel}})^{\frac{1}{2}} = n. \tag{1.3.17}$$

The ratio of the propagation velocity of free space to that of a medium is the refractive index, n, of the medium. In this case both n and k are real quantities. Since μ_{rel} is nearly unity for most materials of importance in the infrared, the refractive index can often be approximated by $n \sim (\varepsilon_{\mathrm{rel}})^{\frac{1}{2}}$.

The constant k has the dimension of inverse length; it is the number of radians per meter, the angular wavenumber. Therefore,

$$k\lambda = 2\pi, \tag{1.3.18}$$

where λ is the wavelength in meters. The angular frequency, ω, measured in radians per second, is then

$$\omega = kv. \tag{1.3.19}$$

The frequency, f, in hertz (cycles per second), and the wavenumber, ν, in m^{-1}, are

$$f = \frac{\omega}{2\pi}; \qquad \nu = \frac{k}{2\pi}. \tag{1.3.20}$$

Even for a wavelength of 1000 μm the frequency is approximately 3×10^{11} Hz, a very high frequency compared with radio waves. The FM broadcast band is about 100 MHz or 10^8 Hz, for comparison. The term $k^2/\varepsilon\mu$ in Eq. (1.3.15) is simply ω^2

and the solution for E_y becomes:

$$E_y = \exp\left(-\frac{\sigma t}{2\varepsilon}\right) \exp\left\{\pm \mathrm{i}\omega \left[1 - \left(\frac{\sigma}{2\varepsilon\omega}\right)^2\right]^{\frac{1}{2}} t\right\} \exp\left(\pm \mathrm{i}kx\right). \tag{1.3.21}$$

As required for a second order differential equation, Eq. (1.3.21) represents two solutions, indicated by the $\pm$ signs. One solution describes a wave traveling in the direction of x (outgoing wave, opposite signs, $+\,-$ or $-\,+$), and the other, a wave traveling in the opposite direction (incoming wave, equal signs, $+\,+$ or $-\,-$). If the amplitudes of these waves are equal, only a standing wave exists. For a nonconductive medium, where σ is zero, the solution for the outgoing wave simplifies to

$$E_y(\sigma = 0) = \mathrm{e}^{\pm \mathrm{i}(kx - \omega t)}, \tag{1.3.22}$$

which is a plane, unattenuated wave traveling in the x-direction. This case is shown in Fig. 1.3.1 by the periodic curve marked ‘0’.

For a weakly conducting material – dry soil or rocks, for example – two effects may be noted. First, due to the factor $\exp\left(-\sigma t/2\varepsilon\right)$ in Eq. (1.3.21), the amplitudes of the waves diminish exponentially with time. Materials with good optical transmission properties must, therefore, be electrical insulators, but not all insulators are transparent. For many substances the frequency dependence of the refractive index is due to quantum mechanical resonances. Equation (1.3.17) is valid for low frequencies where v and n can be determined from the static values of ε and μ, but not necessarily at infrared or visible wavelengths. The second effect to be noted in Eq. (1.3.21) concerns a frequency shift by the factor $[1 - (\sigma/2\varepsilon\omega)^2]^{\frac{1}{2}}$. As long as σ is small compared with $2\varepsilon\omega$, as in the case marked 0.05 in Fig. 1.3.1, the frequency shift is negligible, but it becomes noticeable for the case $\sigma/2\varepsilon\omega = 0.2$. If σ is equal to or larger than $2\varepsilon\omega$ – that is, if the conduction current is comparable to or larger than the displacement current, as in metals – then the square root in Eq. (1.3.21) becomes zero or imaginary; in either case periodic solutions disappear and only an exponential decay exists, shown by curve 1 of Fig. 1.3.1.

Now we return to the choice of p in Eq. (1.3.12). With the assumption $p = \pm \mathrm{i}\omega$ the solution for T becomes

$$T = \mathrm{e}^{\pm \mathrm{i}\omega t}, \tag{1.3.23}$$

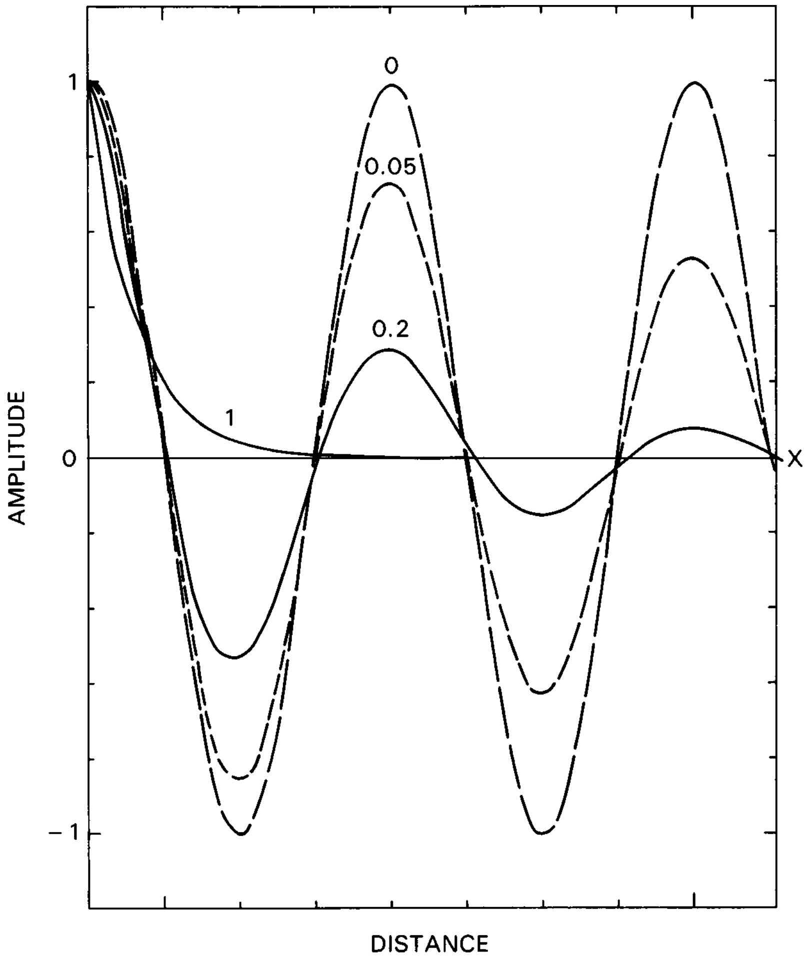

Fig. 1.3.1 Amplitudes of electromagnetic waves propagating in a medium. The parameter refers to the ratio of conduction to displacement current. If this ratio is zero the material is transparent. If this ratio is one or larger, such as in metals, only an exponential decay exists.

but in this case k is complex. We have

$$k = (\varepsilon\mu\omega^2 + \mathrm{i}\sigma\mu\omega)^{\frac{1}{2}} = \frac{\omega}{c}(n_\mathrm{r} + \mathrm{i}n_\mathrm{i}), \tag{1.3.24}$$

where n_r is the real and n_i the imaginary part of the refractive index, n. Squaring Eq. (1.3.24) and setting the real and imaginary parts of both sides equal leads to equations for the real part of k,

$$\frac{\omega n_\mathrm{r}}{c} = \omega\left(\frac{\varepsilon\mu}{2}\left\{\left[1 + \left(\frac{\sigma}{\varepsilon\omega}\right)^2\right]^{\frac{1}{2}} + 1\right\}\right)^{\frac{1}{2}}, \tag{1.3.25}$$

and for the imaginary part,

$$\frac{\omega n_\mathrm{i}}{c} = \omega\left(\frac{\varepsilon\mu}{2}\left\{\left[1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2\right]^{\frac{1}{2}} - 1\right\}\right)^{\frac{1}{2}}. \tag{1.3.26}$$

Therefore, E_y may also be expressed by

$$E_y = \mathrm{e}^{\pm\mathrm{i}\omega(n_\mathrm{r}+\mathrm{i}n_\mathrm{i})x/c}\mathrm{e}^{\pm\mathrm{i}\omega t}. \tag{1.3.27}$$

The term $n = n_\mathrm{r} + \mathrm{i}n_\mathrm{i}$ is the complex refractive index, a concept that is used in the discussion of the interaction of radiation with solid matter (Sections 3.7.b and 3.8).

So far we have concerned ourselves with the electric field strength, **E**. Now we return to the magnetic field strength, **H**. Following a similar procedure for **H** as for **E** leads to analogous equations. After multiplication by ε and differentiation with respect to time of Eq. (1.3.2), one obtains

$$\varepsilon\mu\ddot{\mathbf{H}} = -\varepsilon\frac{\partial}{\partial t}(\boldsymbol{\nabla}\times\mathbf{E}). \tag{1.3.28}$$

If one applies the curl operator to Eq. (1.3.1) one finds

$$\varepsilon\frac{\partial}{\partial t}(\boldsymbol{\nabla}\times\mathbf{E}) + \sigma(\boldsymbol{\nabla}\times\mathbf{E}) = \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{H}). \tag{1.3.29}$$

Multiplication of Eq. (1.3.2) by σ and substitution of this as well as Eq. (1.3.29) into Eq. (1.3.28) yields

$$\varepsilon\mu\ddot{\mathbf{H}} + \sigma\mu\dot{\mathbf{H}} = \boldsymbol{\nabla}^2\mathbf{H}, \tag{1.3.30}$$

which is identical in form with Eq. (1.3.6) for the electrical field strength. The solution for **H** is, therefore, analogous to that for **E**. For $\sigma = 0$, and for the **E** vector in the y-direction only, a component of **H** in the z-direction exists. With the help of Eq. (1.3.22), Eq. (1.3.2) reduces to

$$\mu\frac{\partial H_z}{\partial t} = -\frac{\partial E_y}{\partial x} = -\mathrm{i}k\mathrm{e}^{\pm\mathrm{i}(kx-\omega t)}. \tag{1.3.31}$$

For a periodic function, integration with respect to time is accomplished by dividing by $(-\mathrm{i}\omega)$ and, since $k\mu^{-1}\omega^{-1}$ equals $\varepsilon^{\frac{1}{2}}\mu^{-\frac{1}{2}}$,

$$H_z = \frac{k}{\mu\omega}\mathrm{e}^{\pm\mathrm{i}(kx-\omega t)} = \left(\frac{\varepsilon}{\mu}\right)^{\frac{1}{2}} E_y = mE_y. \tag{1.3.32}$$

The factor m has the dimension of a conductance or, equivalently, of a reciprocal resistance. This resistance is called the wave resistance or, more generally, the optical wave impedance of the medium. For free space the wave resistance is $\sim 377\ \Omega$. For maximum efficiency transmitting and receiving antennas must be matched to that impedance. Similarly, electrical transmission lines must be terminated by their conjugate wave impedances to avoid reflections. In optics an analogous situation exists. No reflection takes place at the interface of two media if their wave impedances are matched, a consideration important for the design of antireflection coatings.

A wave ($\sigma = 0$) traveling in the x-direction, such as described by Eq. (1.3.22), is displayed in Fig. 1.3.2. The electric field strength **E** has a component only in the y-direction and the magnetic vector **H** has one only in the z-direction. The Poynting vector **S** lies along the x-axis. In the time dt the whole pattern moves the distance dx with velocity v in the direction of **S**.

The case shown in Fig. 1.3.2 is for a nonabsorbing medium. To find the relationship between H_z and E_y for an absorbing medium we apply the solution for E_y (Eq. 1.3.27) to the second of Maxwell's equations (Eq. 1.3.2) and find

$$\mu\frac{\partial H_z}{\partial t} = -\frac{\partial E_y}{\partial x} = -\mathrm{i}\frac{\omega}{c}(n_\mathrm{r} + \mathrm{i}n_\mathrm{i})E_y, \tag{1.3.33}$$

which leads after integration (division by $-\mathrm{i}\omega$) to

$$H_z = \frac{(n_\mathrm{r} + \mathrm{i}n_\mathrm{i})}{\mu c}E_y. \tag{1.3.34}$$

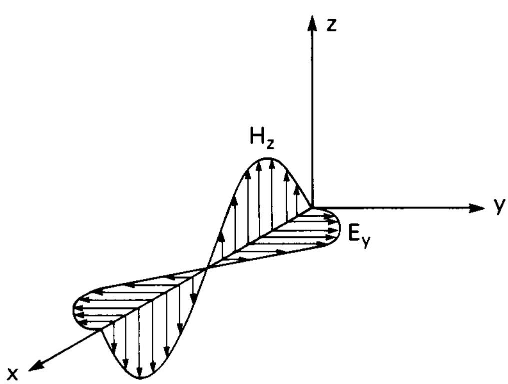

Fig. 1.3.2 Electric (E_y) and magnetic (H_z) vectors in a linearly polarized electromagnetic wave propagating along the x-axis.

Since $(n_\mathrm{r} + \mathrm{i}n_\mathrm{i})$ can be expressed by an amplitude, $(n_\mathrm{r}^2 + n_\mathrm{i}^2)^{\frac{1}{2}}$, and a phase angle, $\gamma = \mathrm{arctg}\, n_\mathrm{i}/n_\mathrm{r}$, we obtain

$$H_z = \pm \frac{1}{\mu c} \left(n_\mathrm{r}^2 + n_\mathrm{i}^2\right)^{\frac{1}{2}} \mathrm{e}^{\mathrm{i}\gamma} E_y. \tag{1.3.35}$$

In a conductive material E_y and H_z are still at right angles to each other and to **S**, but they are phase shifted by an angle γ, and not in phase as shown for a nonabsorbing medium in Fig. 1.3.2.

1.4 Polarization

Now we return to waves in nonabsorbing media. The wave shown in Fig. 1.3.2 is linearly polarized in the y-direction. Traditionally, the direction of the electric field strength, **E**, and the Poynting vector define the plane of polarization. Linearly polarized waves are also possible in the z-direction or at any angle in the y–z plane. The vector **E** may be decomposed into its y- and z-components,

$$\mathbf{E} = \hat{\mathbf{j}} E_y + \hat{\mathbf{k}} E_z. \tag{1.4.1}$$

A linearly polarized wave with an arbitrary plane of polarization may be visualized as the superposition of two waves of the same frequency and phase, one linearly polarized in the y- and the other in the z-direction. But what is the consequence when two waves, E_y and E_z, of the same frequency, both linearly polarized, but with a distinct difference in phase and with different amplitudes, are superimposed? By phase difference we mean differences between the **E** vectors and not between **E** and **H**, which occur only in absorbing media. Since Maxwell's equations are linear, the corresponding vectors, E_y and E_z, of the two waves must be added. The resulting vector sum, **E**, is then the combined field strength. The same applies to the **H** vectors; **E** and **H** are still orthogonal. However, the tip of **E** does not describe a strictly sinusoidal curve in a plane, as shown in Fig. 1.3.2, but rather a curve in space that progresses uniformly along the x-axis; the projection in the y–z plane is not a straight line but an ellipse. We call such a wave elliptically polarized (Fig. 1.4.1). Conversely, an elliptically polarized wave may be decomposed into two linearly polarized waves. If the amplitudes of both superimposed waves are equal, the ellipse becomes a circle and we speak of circular polarization. In that case the end point of the **E** vector travels on a spiral of constant radius around the x-axis. The end point of a circularly or elliptically polarized wave can form a right- or a left-handed spiral. Unfortunately, according to tradition, a right-handed spiral is called a left-handed polarization because in the nineteenth century right- and left-handedness was judged by the observer facing the beam of light. Polarization phenomena play important

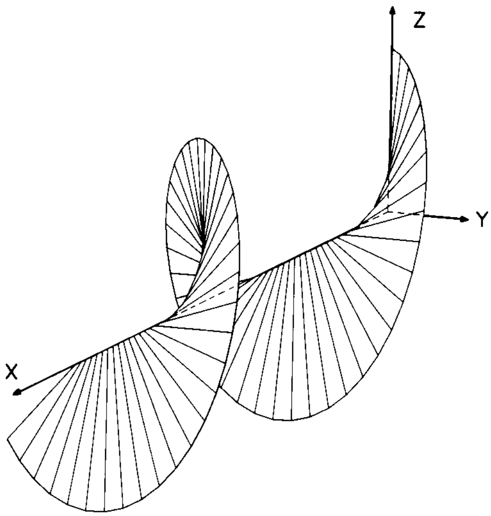

Fig. 1.4.1 Electric vector in an elliptically polarized wave propagating in the x-direction. The magnetic vector is orthogonal to the electric vector.

roles in instrument design, in the theory of reflection and refraction, and in theories of scattering of radiation by particles.

1.5 Boundary conditions

So far we have discussed electromagnetic phenomena in a homogenous medium; now we consider two media and the conditions at their interface. We restrict the discussion to transparent substances. In the media (medium 1 above and medium 0 below the boundary) there exist electric and magnetic fields. In this section the index zero does not refer to free space. At the dividing surface between both domains the fields can be decomposed into two components normal and tangential to the boundary. Consider first the normal component of **B**. To deal with the discontinuity in ε and μ across the dividing surface we consider a small volume that contains a small region of the surface between media 1 and 0 (Fig. 1.5.1). The area of this volume element exposed to medium 1 is δA_1 plus the circumference, s, times $\delta h/2$. The area exposed to medium 0 is δA_0 plus the other half of the circumferential area. Instead of the abrupt change at the boundary we let B_n change gradually from the value $B_\mathrm{n}^{(1)}$ at the surface δA_1 to the value $B_\mathrm{n}^{(0)}$ at the surface δA_0. Applying Gauss' theorem to this volume yields

$$\int_{\text{Volume}} (\boldsymbol{\nabla} \cdot \mathbf{B})\, \mathrm{d}V = \int_{\text{Surface}} \mathbf{B} \cdot \mathrm{d}\mathbf{A}. \tag{1.5.1}$$

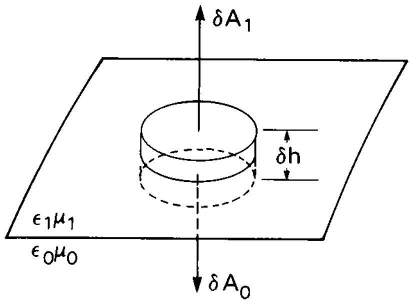

Fig. 1.5.1 Surface element of the boundary of two media of different electromagnetic properties. One half of the volume element is in medium 1, and the other half is in medium 0.

Since $(\boldsymbol{\nabla} \cdot \mathbf{B})$ is zero [Eq. (1.1.4)] the integrals in Eq. (1.5.1) must also be zero. The right side may be expressed by

$$\int \mathbf{B} \cdot \mathrm{d}\mathbf{A} = B_{\mathrm{n}}^{(1)} \cdot \delta A_1 - B_{\mathrm{n}}^{(0)} \cdot \delta A_0 + \left(B_{\mathrm{t}}^{(1)} + B_{\mathrm{t}}^{(0)}\right) s \frac{\delta h}{2} = 0. \tag{1.5.2}$$

Let δh become very small; the contribution from the circumferential area diminishes. Since the areas δA_1 and δA_0 are equal

$$B_{\mathrm{n}}^{(1)} - B_{\mathrm{n}}^{(0)} = 0. \tag{1.5.3}$$

At the interface the normal components of the induction are identical in both media; B_{n} is continuous across the boundary.

The behavior of the component of **D** normal to the boundary may be treated similarly, except that the integrals are not necessarily zero. In this case the charge density ρ must be taken into account. In the transition from the volume element to the surface element, the volume density becomes a surface density, ρ_{surf}, given by

$$D_{\mathrm{n}}^{(1)} - D_{\mathrm{n}}^{(0)} = \rho_{\mathrm{surf}}. \tag{1.5.4}$$

In the presence of a surface charge the normal component of the electric displacement changes abruptly. In the absence of a surface charge, D_{n} is continuous across the boundary.

To investigate the tangential components of **E** and **H** consider a closed loop (Fig. 1.5.2). The loop consists of the elements δs_1, δs_0, and two short connectors, each of length δh. The surface normal of the loop d**A** is in the direction of unit

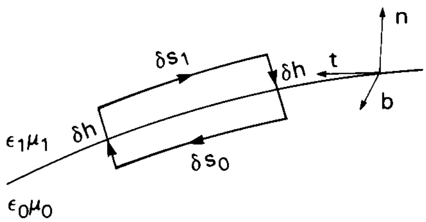

Fig. 1.5.2 Loop at the interface between two media. The vectors **n**, **t**, and **b** indicate the directions normal to the interface surface, tangential to the surface, but in the plane of the loop, and the orthogonal direction, also tangential to the interface, but normal to the loop area.

vector **b**. Applying Stokes' theorem to the loop, one finds

$$\int_{\text{loop area}} (\boldsymbol{\nabla} \times \mathbf{E}) \cdot \mathrm{d}\mathbf{A} = \int_{\text{contour}} \mathbf{E} \cdot \mathrm{ds}. \tag{1.5.5}$$

The integration path of the contour integral is along δs_1, δh, δs_0, and δh, as indicated in Fig. 1.5.2. By replacing the contour integral by its elements, the second of Maxwell's equations, Eq. (1.1.2), yields

$$-\int \dot{B}_b\, \delta s\, \delta h = E_{\mathrm{t}}^{(1)}\, \delta s_1 - E_{\mathrm{n}}\, \delta h + E_{\mathrm{t}}^{(0)}\, \delta s_0 + E_{\mathrm{n}}\, \delta h. \tag{1.5.6}$$

Upon once again letting δh approach zero, the integral over the area of the loop vanishes ($\dot{B}$ is assumed to be finite) and, considering that δs_1 and δs_0 are opposite in sign, we find

$$E_{\mathrm{t}}^{(1)} - E_{\mathrm{t}}^{(0)} = 0. \tag{1.5.7}$$

The tangential component of the electric field strength is continuous across the boundary. Following a similar procedure for the tangential component of **H** one finds:

$$H_{\mathrm{t}}^{(1)} - H_{\mathrm{t}}^{(0)} = \mathrm{j}_{\mathrm{surf}}. \tag{1.5.8}$$

The tangential component of the magnetic field strength changes abruptly in the presence of a surface current, but it is continuous in the absence of such a current.

1.6 Reflection, refraction, and the Fresnel equations

With the boundary conditions established, one may examine an electromagnetic wave striking the interface between two media. As before, we assume both media to be nonconductive and located above and below a flat surface, which we choose to be the x–z plane. The dividing surface between both media is assumed to be free of charges and currents, which implies that the normal components of **D** and **B** and the tangential components of **E** and **H** are continuous across the boundary. Medium 1 has the dielectric constants ε_1 and the permeability μ_1; medium 0 has the properties ε_0 and μ_0. We consider a plane wave with Poynting vector **S** incident on the interface; the plane containing **S** and the normal to the interface is called the plane of incidence. Here, we assume this is the x–y plane ($S_z = 0$), and that the electric field vector is perpendicular to this plane; later we consider the case where the electric vector lies in the plane of incidence. The incident wave will be split at the interface into a reflected and a transmitted (refracted) wave. In medium 1 the superposition of the incoming and the reflected wave is

$$E_z(y \geq 0) = B_1 \, \mathrm{e}^{\mathrm{i}k_1(x \sin\phi_1 - y\cos\phi_1)} + C_1 \, \mathrm{e}^{\mathrm{i}k_1(x \sin\phi_1' + y\cos\phi_1')}. \tag{1.6.1}$$

The refracted wave in the lower half-space is

$$E_z(y \leq 0) = B_0 \, \mathrm{e}^{\mathrm{i}k_0(x \sin\phi_0 - y\cos\phi_0)}. \tag{1.6.2}$$

The factor $\exp(\mathrm{i}\omega t)$ has been omitted for simplicity, but the amplitudes, B_1, C_1, and B_0 have been written explicitly. At $y = 0$, continuity of the tangential component, E_z, across the boundary requires

$$B_1 \, \mathrm{e}^{\mathrm{i}k_1 x \sin\phi_1} + C_1 \, \mathrm{e}^{\mathrm{i}k_1 x \sin\phi_1'} = B_0 \, \mathrm{e}^{\mathrm{i}k_0 x \sin\phi_0}. \tag{1.6.3}$$

Since this equation must be valid for all values of x, all exponentials must be the same, which leads to two conditions:

$$\phi_1 = \phi_1', \tag{1.6.4}$$

which expresses the law of reflection, and

$$k_1 \sin\phi_1 = k_0 \sin\phi_0 \tag{1.6.5}$$

or

$$\frac{\sin\phi_1}{\sin\phi_0} = \frac{k_0}{k_1} = \left(\frac{\varepsilon_0\mu_0}{\varepsilon_1\mu_1}\right)^{\frac{1}{2}} = \frac{n_0}{n_1} = n_{10}, \tag{1.6.6}$$

which is the law of refraction; n_{10} is the relative refractive index between media 1 and 0. For these conditions, Eq. (1.6.3) reduces to

$$B_1 + C_1 = B_0. \tag{1.6.7}$$

The tangential component of **H** provides another set of equations for the amplitudes B_1, C_1, and B_0. According to Eq. (1.3.32), the amplitude of **H** can be found by multiplying **E** by $\pm m$. The right-hand rule for the vectors **E**, **H**, and **S** determines the choice of sign of m.

$$H_x(y \geq 0) = m_1 \cos\phi_1\, \mathrm{e}^{\mathrm{i}k_1 x \sin\phi_1}(-B_1\, \mathrm{e}^{-\mathrm{i}k_1 y \cos\phi_1} + C_1\, \mathrm{e}^{\mathrm{i}k_1 y \cos\phi_1}) \tag{1.6.8}$$

and

$$H_x(y \leq 0) = -m_0 \cos\phi_0\, \mathrm{e}^{\mathrm{i}k_0 x \sin\phi_0} B_0\, \mathrm{e}^{-\mathrm{i}k_0 y \cos\phi_0}. \tag{1.6.9}$$

For $y = 0$ the tangential components of **H** must be the same for both media, which leads to

$$m_1 \cos\phi_1(-B_1 + C_1) = -m_0 \cos\phi_0 B_0, \tag{1.6.10}$$

where m_1 and m_0 are the conductances of medium 1 and 0, respectively [see Eq. (1.3.32)]. Combining Eqs. (1.6.10) and (1.6.7) permits elimination of B_0 or C_1. The relative amplitudes of the transmitted ($T_\perp = B_0/B_1$) and reflected ($R_\perp = C_1/B_1$) waves are

$$T_\perp = \frac{2m_1 \cos\phi_1}{m_1 \cos\phi_1 + m_0 \cos\phi_0} \tag{1.6.11}$$

and

$$R_\perp = \frac{m_1 \cos\phi_1 - m_0 \cos\phi_0}{m_1 \cos\phi_1 + m_0 \cos\phi_0}. \tag{1.6.12}$$

Now we consider the case of the magnetic vector normal to the plane of incidence; i.e., only H_z exists. **E** is orthogonal to **H** and, therefore, in the plane of incidence. The polarization of this wave is orthogonal to that of the first case. With similar considerations one finds $B_1 + C_1 = m_{10} B_0$ and $B_1 - C_1 = B_0 \cos\phi_0 / \cos\phi_1$. Solving for $T_\parallel = B_0/B_1$ and $R_\parallel = C_1/B_1$ yields in this case

$$T_\parallel = \frac{2m_1 \cos\phi_1}{m_0 \cos\phi_1 + m_1 \cos\phi_0} \tag{1.6.13}$$

and

$$R_\parallel = \frac{m_0 \cos\phi_1 - m_1 \cos\phi_0}{m_0 \cos\phi_1 + m_1 \cos\phi_0}. \tag{1.6.14}$$

The transmitted and the reflected fractional amplitudes of the incident radiation polarized perpendicular to the plane of incidence are $T_\perp$ and $R_\perp$, respectively. The components polarized in the plane are $T_\parallel$ and $R_\parallel$, respectively. Equations (1.6.11) through (1.6.14) are the Fresnel equations (Fresnel, 1816).

Since the emissivity of a surface is one minus the square of the amplitude ratio, $(R_\perp)^2$ or $(R_\parallel)^2$, the thermal emissivity depends also on the refractive index and the emission angle. Consider the case of a nonmagnetic homogeneous layer of refractive index $n_0 = n$ bounded by a vacuum, $n_1 = 1$. With the help of Eq. (1.6.6) we can eliminate ϕ_0 from the reflection ratios, Eqs. (1.6.12) and (1.6.14); calling $\phi_1 = \phi$ we obtain for the emissivities

$$\varepsilon_\perp = 1 - (R_\perp)^2 = 1 - \left[\frac{\cos\phi - (n^2 - \sin^2\phi)^{\frac{1}{2}}}{\cos\phi + (n^2 - \sin^2\phi)^{\frac{1}{2}}}\right]^2 \tag{1.6.15}$$

and

$$\varepsilon_\parallel = 1 - (R_\parallel)^2 = 1 - \left[\frac{n^2\cos\phi - (n^2 - \sin^2\phi)^{\frac{1}{2}}}{n^2\cos\phi + (n^2 - \sin^2\phi)^{\frac{1}{2}}}\right]^2. \tag{1.6.16}$$

The emissivities of substances with refractive indices of 2 or 4, bordered by a vacuum, are shown in Fig. 1.6.1 for both planes of polarization as functions of the emission angle, ϕ. The emitted radiation from a smooth surface is polarized, except for the case of normal incidence. The emission maximum of $\varepsilon_\parallel$ corresponds to the reflection minimum at the Brewster angle. To find the hemispherical emissivity one has to integrate $\varepsilon_\parallel$ and $\varepsilon_\perp$ over the whole hemisphere and average the results for both planes of polarization.

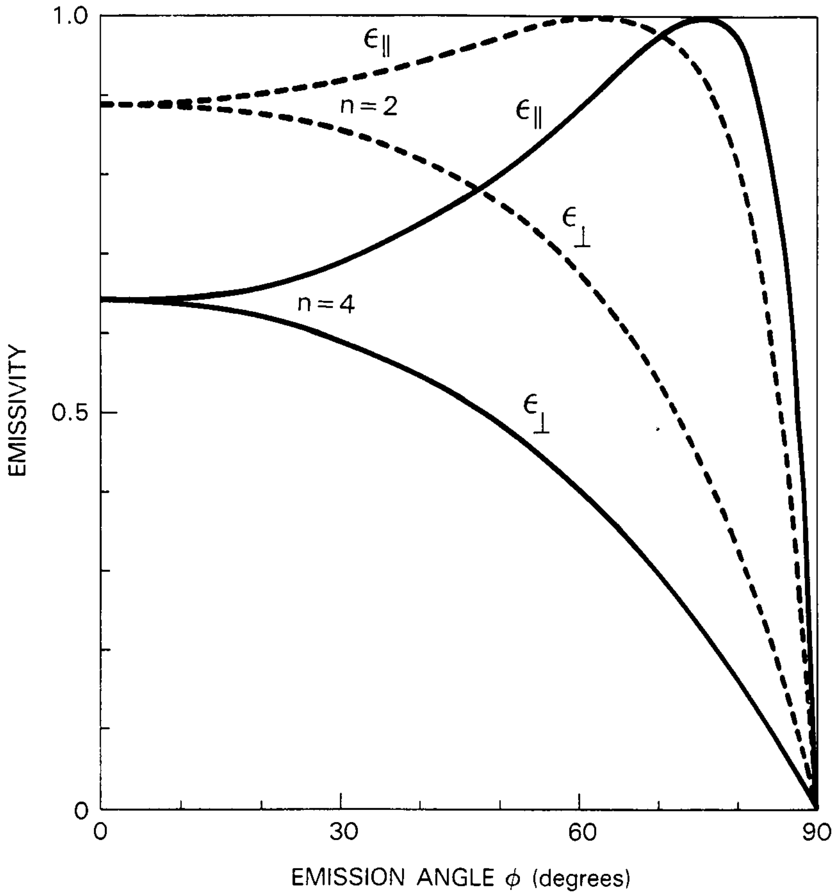

Fig. 1.6.1 Emissivity of a smooth flat surface with an index of refraction, n, as a function of emission angle. Both planes of polarization are shown for n equal to 2 (dashed lines) and 4 (solid lines).

For normal incidence ($\cos\phi_1 = \cos\phi_0 = 1$) and nonmagnetic materials ($m_1/m_0 = n_1/n_0$) the Fresnel equations simplify to

$$T_{\perp} = T_{\parallel} = \frac{2n_1}{n_1 + n_0} \tag{1.6.17}$$

and

$$R_{\perp} = -R_{\parallel} = \frac{n_1 - n_0}{n_1 + n_0}. \tag{1.6.18}$$

If the second medium is metal the same equations are valid; however, n_0 becomes complex [see Eq. (1.3.24)]. For $n_1 = 1$ and $n_0 = n_{\mathrm{r}} + \mathrm{i}n_{\mathrm{i}}$ the ratio of amplitudes