HYDROLOGIC OPTICS

Volume II. Foundations

R.W. PREISENDORFER

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC & ATMOSPHERIC ADMINISTRATION
ENVIRONMENTAL RESEARCH LABORATORIES

HONOLULU, HAWAII

1976

NOAA FORM 25-13 (2-73)

BIBLIOGRAPHIC DATA SHEET

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

1. NOAA ACCESSION NUMBER	2.	3. RECIPIENT'S ACCESSION NUMBER
NOAA-76082308		

4. TITLE AND SUBTITLE
Hydrologic Optics, Volume II. Foundations

5. REPORT DATE
1976

6.

7. AUTHOR(S)
R. W. Preisendorfer

8. REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Joint Tsunami Research Effort,
Honolulu, HI 96822

10. PROJECT/TASK NO.

11. CONTRACT/GRANT NO.

12. SPONSORING ORGANIZATION NAME AND ADDRESS
NOAA, Environmental Research Laboratories
Pacific Marine Environmental Laboratory
Boulder, CO 80302

13. TYPE OF REPORT AND PERIOD COVERED

14.

15. PUBLICATION REFERENCE
Hydrologic Optics, Vol II. Foundations, 1976. 410 p. numerous fig, tab, ref.

16. ABSTRACT
This second volume of a set of six volumes of the text on hydrologic optics deals with foundations: radiometric and photometric concepts and the interaction principle. Detailed treatment is given on: radiant flux; the meaning and fundamental geometric properties of radiant flux; irradiance and radiant emittance; radiance; an invariance property of radiance; scalar irradiance, radiant energy, and related concepts; vector irradiance; radiant intensity; polarized radiance; transition from radiometry to photometry; generalized photometries; the interaction principle; reflectance and transmittance operators for surfaces; applications to plane surfaces; applications to curved surfaces; reflectance and transmittance operators for plane-parallel media; applications to plane-parallel media; interaction operators for general spaces; applications to general spaces; derivation of the beam transmittance function; derivation of the volume attenuation function; derivation of path radiance and path function; derivation of apparent-radiance equation; derivation of volume scattering function; equation of transfer for radiance; integral structure of the interaction operators and summary of the interaction method. (OEIS)

17. KEY WORDS AND DOCUMENT ANALYSIS

17A. DESCRIPTORS

*Light, *Reflectance, *Water properties, *Physical properties

17B. IDENTIFIERS/OPEN-ENDED TERMS

*Hydrologic optics, Photoelectric effects, Irradiance, Scalar irradiance, Transmittance

17C. COSATI FIELD/GROUP

COSATI 20F, 7D COWRR 01A

PRICES SUBJECT TO CHANGE

18. AVAILABILITY STATEMENT
Released for distribution:
Sarah J. Knoll

19. SECURITY CLASS *(This report)*
UNCLASSIFIED

21. NO. OF PAGES

20. SECURITY CLASS *(This report)*
UNCLASSIFIED

* U. S. GOVERNMENT PRINTING OFFICE: 1975--665077/1124 REGION NO. 6

HYDROLOGIC OPTICS

Volume II. Foundations

R.W Preisendorfer

Joint Tsunami Research Effort
Honolulu, Hawaii

1976

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
Environmental Research Laboratories
Pacific Marine Environmental Laboratory

i-a

Nature Seen Through Half-Shut Eyes

A man who takes a magnifying glass into a picture gallery and examines the canvases at a distance of 3 inches may acquire much interesting information about the texture of paint, but he does not see the pictures. It is better to stand away. If trivial details still intrude, it is better to half-shut the eyes. As a final step, it is well to shut the eyes completely and think about what has been seen.

J.L. Synge

Science, 5 October 1962

Volume II

Chapter 2

Radiometric and Photometric Concepts

Chapter 3

The Interaction Principle

PREFACE

My first encounter with radiometry and photometry was as a student reading Sears' *Optics*.* The lucid exposition in Sears' book, ably illuminated by the lectures of Prof. S. Q. Duntley, awakened my interest in the subject. Soon afterward the geometer in me took over as I sought the foundations of the subject. Following graduation from Massachusetts Institute of Technology in 1952, and during my first years as a mathematics graduate student at Scripps Institution of Oceanography, I had the opportunity to develop my awakened interest in radiometry, and to find the foundations of this subject in measure theory. Eventually I found that radiometry, a beautiful union of Euclid's geometry and the axiomatized notion of radiant flux, is the ground on which radiative transfer theory could be built. For a few heady years I had the leisure to explore this foundation (see, e.g., [210], [211], [216]) much as the ancient geometers explored the world of euclidean geometry and the beginnings of mechanics. For one need not know much about the physical world beyond what his senses reveal in order to be qualified to pursue radiometry, radiative transfer, and their applications to problems of visibility and radiant energy flow in the sea. In this sense radiometry and probability theory are very much alike. While radiometry is the marriage of geometry and radiant flux, probability theory is the union of geometry and chance. In both disciplines, the mathematical vehicle for the physical concept is the notion of a *measure*. Indeed, the parallel between radiative transfer and that branch of probability known as 'Markov chains' is exact, as I showed sometime later in Chapter XIII of my monograph [251] on radiative transfer theory.

This volume, then, is the product of a labor of love, wherein very deep geometric predilections took over my first years of scientific research, years in which for better or worse I half-shut my eyes to the multifarious richness of the real world, and tried by thought alone to order my visual experiences in a suitable mathematical frame. I was successful in that effort. For if one carves out of the chaos of his experience a small enough piece, he can examine it and understand it, and eventually make its secrets part of himself once and for all. But the price of this victory is quite dear: the remaining portions of the world sweep by and onward while one remains anchored to a spot, examining a few grains of earth for order and meaning.

---

*Sears, F. W., *Optics*, Addison Wesley, Cambridge, Mass. (1949), 3rd ed.

The interested student of radiative transfer theory may take the following as a base on which to rest his own work. He will then be spared the necessity for remaining overly long at a relatively isolated point in the conceptual landscape of radiative transfer theory. If he can then see farther and clearer because of this work, my efforts will have done some good. If he feels that the final answer has yet to be found in the quest for the foundations, then I wish him good luck and a full measure of joy in pursuing that quest.

The final manuscript was typed by Ms. Judy Marshall.

R.W.P.

Honolulu, Hawaii
January 1974

CHAPTER 2

RADIOMETRIC AND PHOTOMETRIC CONCEPTS

2.0 Introduction

Having completed our introductory survey of hydrologic optics, we now embark on a theoretical reconstruction of all that we saw and did: we shall now start from scratch.

In this chapter we shall develop the concepts of radiometry and photometry needed in the study of radiative transfer. The mode of approach to these concepts is governed by the particular outlook of radiative transfer theory as it is applied to hydrologic optics. As we have defined it (Sec. 1.0) hydrologic optics is the study of radiative transfer in general hydrosols; and radiative transfer, in turn, employs a phenomenological viewpoint of light. Therefore, the approach to radiometry and photometry in hydrologic optics takes place on a phenomenological level. In other words, the concepts of radiometry and photometry which play a major role in hydrologic optics are those which are defined for natural light fields in which operations have been performed on a macroscopic level and with instruments which in certain key types of response are very similar, qualitatively, to the human eye. Thus, like the human eye, the sizes of the special instruments used in radiative transfer measurements are large compared to the sizes of the wave and particle structures of light and matter. Therefore these instruments do not ostensibly sense and record any of those features of light directly characterizable by its wave or particle structure, such as diffraction and interference features. However, the radiometric instruments used are designed to extend and amplify under precise quantitative control certain selected capabilities inherent in the human eye, foremost of which is the ability to sense and record the various brightnesses of monochromatic light in all directions about a given point in the observer's locale. This, after all is said and done, is the principal goal of classical photometry. The complete route to this goal is necessarily through an intricate maze of psychometric, radiometric, electromagnetic and quantum constructs. However, the paths we shall take to photometry in this chapter can fortunately bypass most of the usual detours along the route, detours which study the manifold conceptual and experimental aspects of the subject. The topics we select for discussion are mainly theoretical and on this level constitute the minimal number just sufficient to allow the logical establishment of those photometric concepts and various radiometric models of natural light fields used in

radiative transfer measurements. We shall stand away from electromagnetic complexity, and half-shut our eyes as we reconstruct radiometry. In the next chapter we shall shut our eyes completely and think about what we have seen in Vol. I.

The outline of this chapter is as follows. We begin in Sec. 2.1 with the operational definition of radiant flux. It is always good practice to give as many means of visualization of a newly defined concept as mutual consistency will allow. For this reason, and also to pave the way for a more versatile presentation of the concepts of hydrologic optics than that of Chapter 1, we develop in Sec. 2.2 the three main ways to conceptually view the notion of radiant flux. The principal properties of radiant flux, as they are used in geometrical radiometry, are developed in Sec. 2.3. Then, in close succession, the principal derived concepts of radiometry are developed: radiance and various forms of irradiance, along with theorems governing and examples illustrating their salient properties. Throughout our development we shall emphasize the *geometrical* aspects of radiometry rather than their physical aspects. The latter aspects, to the degree that we shall need to study them in this work, are reserved for discussion in Sec. 2.1. However, some notice must also be taken of the physical aspects of radiometry in preparing to construct the bridge between radiometric and photometric concepts. Therefore, in Sec. 2.12, we pause to develop those concepts of photometry which facilitate the operational definition of the notion of luminous flux--the photometric counterpart to radiant flux. With the radiometric discussions as a model, the various derived photometric concepts are then readily attained. The chapter closes with some remarks on generalized photometric concepts.

Our present viewpoint of geometrical radiometry and photometry may then be summarized in the following definitions of these disciplines, which we adopt: *Radiometry is the science of the measurement of radiant energy. Geometrical Radiometry is the union of euclidean geometry and Radiometry: it measures and describes the flow of radiant energy of given frequency through volumes, across surfaces, along lines, and at points in space.* With this in mind we can go on to say that: *Geometrical Photometry measures the visual, erythemal, photoelectric, or photographic response, by given receptors, to the quantities of geometrical radiometry, with respect to different frequencies of radiant energy.*

2.1 Radiant Flux

We now take up the details of an operational definition of radiant flux. The heart of the definition we shall adopt consists of the postulation of some physical device which can sense and record in quantitative detail the presence of light--or radiant energy in general--in a neighborhood of a point in space. There are several devices available for such a purpose. Of those currently available, the photoelectric devices are most satisfactory from the point of view of sensitivity and quantitative precision. We pause briefly to survey this class of devices.

Basic Photoelectric Effects

The class of light-measuring devices known collectively as *photoelectric cells* consists of three broad sets, each set being characterized by a distinctive mode of interaction of light with matter and the particular form of electrical response arising from that interaction. These responses are denoted by the terms *photoemissive*, *photoconductive*, and *photovoltaic*. A comparison of the characteristic features of these phenomena is readily made by means of Fig. 2.1.

Part (a) of the figure depicts the electrical essence of a *photoemissive cell* (or *phototube*). Light, indicated by the arrow, is incident on a negatively charged electrode. The impact of the incident light dislodges electrons from the surface of the electrode and these are drawn across the gap to the relatively positively charged electrode within the element. The seat of electromotive force is supplied by a battery or other means and so continuously replenishes the supply of electrons on the negative electrode. The net result of the incident light is a small but measurable current of electrons flowing through a current meter, as shown in the figure. The swarm of electrons, liberated at the electrode by the incident light, streams across the gap between the electrodes and thereby completes the circuit. If there is no incident light on the electrode, then under normal conditions, there are no electrons liberated from the electrode to complete the circuit, and there is consequently no current registered by the meter. Generally, the greater the amount of light incident on the receiving electrode, the correspondingly greater is the resultant current in the circuit. By a careful calibration, the meter can be made to read directly the rate of incidence of radiant energy on the receiving electrode. The photoemissive effect just described is the most recently discovered of the three effects. It was discovered in crude form in 1887 by Heinrich Hertz as a by-product of his classical researches on electromagnetism. Under subsequent refinements, over the years, it has become the principal effect used in photoelectric devices. The theory of the photoemissive effect was not evolved until about eighteen years after its discovery. The theory of the photoemissive effect itself forms a major epoch in the history of physics, for its completion eventually required the concept of the photon as introduced by Einstein in 1905.

A *photoconductive cell* is schematically depicted in part (b) of Fig. 2.1. It was found experimentally in 1873 by Willoughby Smith that the conductivity of the metal selenium increases when light is incident on it. This effect can therefore be put to use in sensing and recording the presence of light, in the manner shown in the figure. The greater amount of light incident on the selenium cell results in a correspondingly greater amount of current flowing through the current meter. When no light is incident on the photoconductive element, there is under normal conditions a small known amount of current (the *dark current*) flowing in the circuit. The full understanding of the photoconductive effect on a microscopic level was achieved only recently using the quantum-

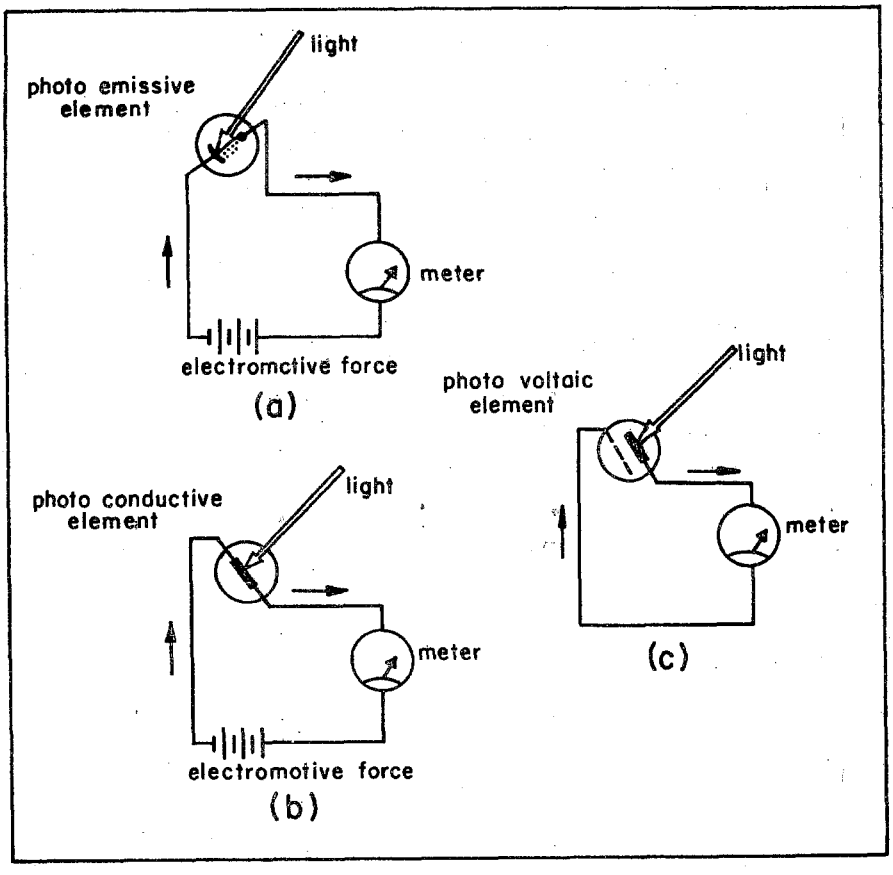


FIG. 2.1 The Basic types of photoelectric cells

based theory of semiconductors. On the basis of this understanding, one can test and use all manners of semiconductors as possible photoconductive materials.

A *photovoltaic cell* is schematically depicted in part (c) of Fig. 2.1. The photovoltaic element consists of two dissimilar substances in close contact (shown slightly separated, for clarity). Light incident on the photovoltaic element generates a difference of electric potential between the two basic parts of the element and as a consequence a current flows in the circuit. This current is measured by a current meter included in the circuit. When no light is incident on the element, no electromotive force is normally produced in the parts of the element, and consequently no current flows in the circuit. Generally, the greater the amount of incident light on the element of the cell, the greater the resultant potential, and the greater the ensuing current in the circuit. The photovoltaic effect antedates both other effects discussed

above. It appears that Edmond Becquerel first observed it in 1839 when a liquid electrolyte containing two immersed electrodes connected through a galvanometer was irradiated by sunlight. Becquerel eliminated the possibility of a thermal voltaic effect generated by differential heating of the electrodes and thereby was led to believe that the light itself gave rise to an electric potential between the electrodes which in turn gave rise to a current in the galvanometer.

The theory of the photovoltaic effect requires the quantum picture of the structure of matter for its complete formulation. However both the photoconductive and photovoltaic effects can be intuitively pictured as being something like weaker versions of the photoemissive effect: on the one hand, in the case of the photoconductive cell, instead of knocking electrons completely free of an area of selenium surface, the incident light on the surface merely gives them enough energy to skim through the lattice of the positive nuclei of the selenium atoms. If there is an existing voltage in the metal, the footloose electrons in the irradiated region are then more readily moved along in a more or less organized manner by the potential difference. On the other hand, the mechanism of the photovoltaic effect is relatively complex. For our descriptive purposes here it may be explained in terms of the effects generated by inherently different electromotive forces of the chemical elements. When two substances of different electromotive force are placed in close proximity (e.g., the dotted and solid elements schematically shown in part (c) of Fig. 2.1) the pull exerted by the positive nuclei of the atoms of one of the substances on electrons is greater than that of the corresponding pull by the other substance. As a result some electrons are swapped from the 'weaker' to the 'stronger' substance when the substances are placed into close contact. However, the electrons captured by the stronger substance can be relatively easily dislodged by irradiation of the boundary between the substances, and thus be caused to move in the resultant electric field naturally existing between the two substances. The magnitude of the potential of this field under irradiation is very nearly the difference in the electromotive forces of the substances.

Operational Definition of Radiant Flux

We now present the operational definition of radiant flux. The brief preliminary excursion into the basic photoelectric effects just completed will endow the definition procedures below with a measure of realism that perhaps may not have been possible had we not paused to make some contact with physical reality. However, the logical basis of the definition of radiant flux and its manifold properties discussed subsequently are quite independent of what radiation measuring devices are used in practice. Indeed, the concepts of radiometry as used in practice are all constructable in terms of the basic notion of radiant flux and appropriate geometrical notions such as surface areas and solid angles. The concept of radiant flux in turn and its few basic geometric

properties are now so well established that they can actually be axiomatized for the purpose of developing a self-contained discipline of geometrical radiometry. In the present development we shall steer a middle road between these extreme alternatives. We shall not go so far as to develop in complete detail an axiomatic theory of radiometry, but we shall indicate the fundamental properties of radiant flux that would occur in such a formulation. The notion of radiant flux will for the most part be handled as an empirically-based concept. However, we shall not, beyond the general suggestions given in the discussion of photoelectric devices above, fix in any detail the form of the device which is used to sense and record the incident flow of radiant energy. In sum, we shall henceforth agree that we have some light-sensitive device which can accurately, quickly, and repeatedly reproduce a quantitive measure of the instantaneous flow of radiant energy onto some well defined surface which acts as a collecting surface for the incident energy. Except for some suggestive remarks in Sec. 2.2, the notion of 'radiant energy' will remain undefined in this work. We take it as given.

Figure 2.2 depicts in more detail, and on a schematic level, the basic form of a widely used type of radiant flux meter. The sequence of events leading to a radiant flux measurement with the radiant flux meter is generally as follows. Radiant energy is incident on the filter of the meter. This energy is funneled in from the environment through a set D of directions. The filter ideally transmits a set F of frequencies of the incident energy and does not transmit any other frequencies. The transmitted frequencies then pass on to a plane collecting surface S. This surface acts to collect a representative amount of the transmitted flux from each

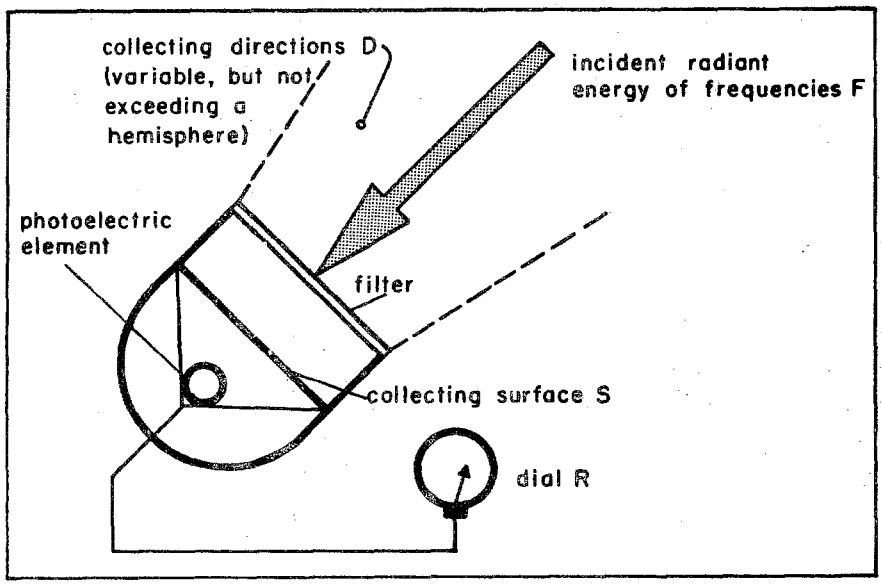


FIG. 2.2 Schematic detail of a radiant-flux meter

direction in D and to pass it on to the photosensitive element of some type of radiant energy sensor. The sensor is part of a circuit of a photoelectric cell, and the presence of the radiant energy flow on the filter thus becomes manifest in a dial reading R of the current meter in the photoelectric cell's circuit (see Fig. 2.1).

In order to obtain a usable measure of the flow of radiant energy, there is basically only one additional requirement on the radiant flux meter assembly, above and beyond the usual requirements on its components demanded by good mechanical and electrical engineering practice. The additional requirement is that its collecting surface S collect energy in a manner which is effectively independent of the direction of incidence of the energy on S. Thus, suppose a narrow beam of radiant energy is incident normally on S, and note the associated reading R of the dial of the meter. Then let the beam's incident angle vary slowly away from normal incidence, keeping the beam always to fall within the surface S. An ideal collecting surface will accept, diffuse, and pass on the energy of this varying beam to the sensor below so that the dial reading R remains fixed. When a collecting surface comes within some preassigned distance of this ideal, we shall call it a *cosine collector*. The reason for this terminology will become clear after the study of the concept of irradiance below. Briefly, it derives from the fact that if the collector is completely bathed in the flux of a homogeneous cylindrical beam, then the *recorded* flux will vary as the cosine of the angle the normal to the collector makes with the axis of the beam of flux. Henceforth, it will be assumed that the collecting surface of the radiant flux meter is a cosine collector.

We now can state the operational definition of radiant flux. We assume that the radiant flux meter, outfitted with a cosine collector S, has been calibrated against some radiometric standard with a known *rate of radiant energy* (*radiant flux*) output (see Chapter 6, Ref. [3]). Then we imagine that we have taken the meter into some radiometric environment such as the depths of some lake or ocean, or perhaps to some point in the atmosphere. The meter is then oriented so that at time t the surface S accepts through the set D of directions radiant flux comprised of a set F of frequencies, with a resultant associated reading R of the meter's dial. The calibration of the dial permits the assignation to this reading R of a *radiant flux* in the form of a nonnegative number denoted by "$\Phi(S,D,t,F)$". The reading R is thereby associated with this particular S,D,t, and F in the radiometric environment. Thus "$\Phi(S,D,t,F)$" denotes the *radiant flux* of frequencies in F which are incident on S, through D, at time t. The dimensions of radiant flux are *energy/time*, or synonymously, *power*, and convenient units are *joules/sec*, or synonymously, *watts*. This pairing process therefore generates a function, the *radiant flux function* denoted by "Φ", which assigns to each collection (S,D,t,F) of surface, direction, frequency and time parameters the nonnegative number $\Phi(S,D,t,F)$ in the manner just described.

The definition of radiant flux given above is an *operational definition* in the sense that it may be translated into a definite sequence of physical operations with a specific instrument in a given environment and which culminate in a unique nonnegative number $\Phi(S,D,t,F)$. This type of definition can be made to stand out in bold relief from still another type which may also be used as effectively as the operational definition in establishing the theory of radiometry. This alternative definition is known as the *constitutive definition* of radiant flux which uses only the concepts of the mathematical framework within which radiometry is modeled. In a constitutive definition there is no immediate appeal to physical operations with a specific instrument in a given environment. For an example of a constitutive definition of radiant flux and the other radiometric quantities, the reader may consult Secs. 109 and 131 of reference [251].

2.2 The Meaning of 'Radiant Flux'

It will be helpful during the discussions of this and subsequent chapters to have in mind some visualizable construct of radiant flux. By having the reader picture in a relatively concrete manner the meaning of the term 'radiant flux', the various principles and laws of radiative transfer used throughout this work will become more readily understood and applied. We have already given the term 'radiant flux' a relatively concrete meaning by adopting an operational definition of the term. In this section we shall go one step further and suggest three ways in which one may visualize radiant flux directly. What we shall offer, then, are conceptual frameworks within which to view the notion of radiant energy and which, especially during theoretical discussions of radiative transfer, one may use in a heuristic manner.

One manner in which radiant flux may be visualized is by a means similar to that used in geometrical optics. In order to discuss the theory of lenses within geometrical optics one may use the method of ray tracing. The heart of this method resides in the concept of the "light ray" and a few simple rules of construction of a ray of light through a lens. Corresponding to this notion we have in geometrical radiometry the notion of a *line of flux*. One may thus visualize $\Phi(S,D,t,F)$ as proportional to the number of straight or curved lines having directions lying within the set D where they terminate on the surface S. The time t and set F of frequencies are usually fixed or understood during a discussion so that the lines of flux constitute a representation of the geometric construct of $\Phi(S,D,t,F)$. In this representation, the magnitude of $\Phi(S,D,t,F)$ is proportional to the number of such lines of flux, the proportionality factor being some fixed number of lines per unit of radiant flux. One may thus imagine the radiant energy as a fluid travelling along the lines of flux. The closer together the lines are within some region, the greater the radiant flux (i.e., radiant energy flow) through that region. The lines of flux are to be determined using the same formulas of geometrical optics as used in ray tracing. Whenever scattering takes place, however, some lines

of flux undergo abrupt changes in direction. Between the points of these abrupt changes in direction the structure of the lines of flux are again governed by the ray tracing formulas of geometrical optics (and the lines are, in most practical instances, merely straight line segments).

Another manner in which radiant flux may be visualized is by means of the Poynting vector of electromagnetic theory. Besides serving to generate an alternative semantic dimension to the term 'radiant flux', such a visualization--if carefully done--serves to establish an analytic link between electromagnetic theory and the concepts of radiometry. Thus, consider the electric and magnetic vector fields $\mathbf{E}$ and $\mathbf{H}$. The vector product $\mathbf{E}\times\mathbf{H}$ is called the *Poynting vector field* and is usually denoted by "$\mathbf{P}$". A dimensional analysis of $\mathbf{P}$ shows that it has the dimensions of radiant flux per unit area. The direction of this flow is along the direction of $\mathbf{P}$, and the magnitude of the flow is that of $\mathbf{P}$, and takes place across a unit area normal to the direction of $\mathbf{P}$. Generally, $\mathbf{P}$ varies in both magnitude and direction many times a second at a given point. Thus the quantity $\Phi(S,D,t,F)$ may be viewed as a time average of the magnitude of the Poynting vector $\mathbf{P}$ confined to the directions within D over the surface S and during some short time period T around time t. The analytic details of the connection between $\mathbf{P}$ and $\Phi(S,D,t,F)$ beyond those alluded to here will have no application in the present work. The reader wishing to study this connection in more detail may consult Sec. 124 of Ref. [251].

One further manner of visualizing the concept of radiant flux is by means of the notion of photons, that is by means of 'particles' of light. Monochromatic radiant flux, i.e., radiant flux of a single frequency, say, ν may be associated with the flow of a set of photons each of energy $h\nu$, where "h" denotes Planck's quantum of action (per photon). In more detail, let* "$n(x,\xi,t,\nu)$" (the *phase-space density*) denote the number of ν-frequency photons per unit frequency interval and per unit volume at each point x over the surface S, and moving with speed v in a unit solid angle along the direction ξ within the bundle D of directions at time t. Then:

$$h\int_D\int_S\int_F n(x,\xi,t,\nu)v(x,\xi,t)\nu(x,\xi,t)\ \xi\cdot k(x)\ dA(x)d\Omega(\xi)dl(\nu) \quad (1)$$

gives the time rate of flow of radiant energy of frequencies in F crossing surface S, within directions D at time t. $k(x)$ is the unit inward normal to S at x. This is then the seat of the meaning of $\Phi(S,D,t,F)$ in terms of the photon concept. In the preceding integral, "A", "Ω" and "l" denoted area, solid angle, and frequency measures, respectively; and these

---

*For reference purposes, we observe here that the connection between the density n and the radiance N is: $N(x,\xi,t,\nu) = h\nu vn(x,\xi,t,\nu)$. (See (5a) of Sec. 2.5.)

will be explained in greater detail in the subsequent sections of this chapter.

As a special case of the preceding connection, let $n(x,\xi,t,\nu)$ be constant of magnitude n over S and over a narrow bundle of directions D normally incident on S and let F consist of discrete frequencies. Then, using the photon interpretation of $\Phi(S,D,t,F)$ just described, we can write:

$$\text{"}\Phi(S,D,t,F)\text{"} \quad \text{for} \quad h\nu n A(S)\Omega(D) \sum_{\nu \in F} \nu \qquad (2)$$

In summary, we have discussed three possible aids to visualizing the meaning of 'radiant flux'. There is the geometric-optics notion of *lines of flux*, the electromagnetic-theoretic construct of the *Poynting vector*, and the quantum-theoretic construct of the moving *photon*. A composite picture may be made by joining all three of the preceding concepts. Thus, one may visualize the photon not as a particle (i.e., a mathematical point) but rather as a spatially small wave train of electromagnetic waves of predominantly a single frequency and moving along the lines of flux. This concept allows light to have at least intuitively, the properties of both particles and waves.

2.3 Fundamental Geometric Properties of Radiant Flux

In this section we shall assemble the six properties of $\Phi(S,D,t,F)$ on which geometrical radiometry may be based. These six properties summarize precisely and explicitly those macroscopic properties of light which are customarily implicitly assumed in radiometry, and which are based on extended experience with the operational definition of radiant flux. By explicitly recognizing and isolating these six properties we may attain a unified and relatively rigorous development of geometrical radiometry. This fundamental group of six properties falls naturally into three pairs of properties, corresponding to the frequency, surface, and direction parameters occurring in $\Phi(S,D,t,F)$.

We begin with the properties of Φ associated with the frequency parameter F. For every two disjoint sets F_1 and F_2:

$$\Phi(S,D,t,F_1) + \Phi(S,D,t,F_2) = \Phi(S,D,t,F_1 \cup F_2) \qquad (1)$$

and

$$\text{if} \quad 1(F) = 0, \quad \text{then} \quad \Phi(S,D,t,F) = 0 \quad . \qquad (2)$$

These properties hold for arbitrary S,D, and t. The first of these is the F-*additivity property* of Φ. The symbol "$\cup$" will be used often below to denote the union of two sets of things. Here "$F_1 \cup F_2$" denotes the set of all frequencies in either F_1 or F_2. By "disjoint sets" we shall mean sets of things which have no elements in common. Thus by "two disjoint sets F_1 and F_2" we mean that F_1 and F_2 have no frequencies in common.

For example the set F_1 of frequencies in the interval from 10^9/sec to 10^{11}/sec inclusive is disjoint from the set F_2 of frequencies from 10^{11}/sec to 10^{12}/sec, inclusive. Thus $F_1 \cup F_2$ is now the set of all frequencies which are either in the interval F_1 or F_2. The second property (2) above is the (absolute) *F-continuity property* of Φ.

The meaning of the F-additivity property is quite simple: imagine radiant flux incident on S through D at time t and consisting of frequencies in $F_1 \cup F_2$. This flux could be irradiated onto S through D simultaneously from two separate sources of frequencies F_1 and F_2, or by means of suitably chosen filters. Alternatively, the flux can be presented first for the frequencies of F_1 and then for the frequencies of F_2. In essence, (1) states that on the macroscopic level there is no interference of two or more wave trains occupying the same region of space. The F-continuity property asserts that, on the macroscopic level and with S,D and t fixed and all other things being equal, the less the number of frequencies in F, the less the radiant flux amount of $\Phi(S,D,t,F)$. In particular, frequency intervals of zero length contain zero radiant flux.

From the F-additivity and F-continuity properties of Φ we derive the concept of monochromatic radiant flux. Thus let us write:

$$\text{"P(S,D,t,}\nu\text{)"} \quad \text{for} \quad \lim_{F \to \{\nu\}} \frac{\Phi(S,D,t,F)}{1(F)} \tag{3}$$

It is precisely the properties (1) and (2) which permit the limit $P(S,D,t,\nu)$ to exist.* When ν is understood, we may drop it from the notation to write:

$$\text{"P(S,D,t)"} \quad \text{for} \quad P(S,D,t,\nu)$$

and even further, we may write:

$$\text{"P(S,D)"} \quad \text{for} \quad P(S,D,t,\nu)$$

when both ν and t are understood. We call $P(S,D,t,\nu)$ the *monochromatic* (or *spectral*) *radiant flux* of frequency ν over S within D at time t, per unit frequency interval. The function P which assigns to (S,D,t,ν) the number $P(S,D,t,\nu)$ is

*The mathematical reader may consult Refs. [216] and [103] for the study of the existence of such limits and their use in radiative transfer theory. For simplicity in exposition we have displayed *finite additivity* of Φ, rather than the full *countable additivity* used in the cited references. Similar remarks pertain to subsequent additivity properties stated in this work.

the *monochromatic* (or *spectral*) *radiant flux function*. It follows from (3) and a theorem of calculus that:

$$\Phi(S,D,t,F) = \int_F P(S,D,t,\nu)\, dl(\nu) \ . \qquad (4)$$

The symbol "l" denotes the length measure along the frequency spectrum. Thus for the interval F consisting of all frequencies from frequency ν_1 to frequency ν_2, where $\nu_1 \leq \nu_2$, we have $l(F) = \nu_2 - \nu_1$. Thus in practical computations one can write "$d\nu$" for "$dl(\nu)$" and (4) is then understood to be a Riemann integral. This is the intended interpretation of (4) for use in this work. However, general discussions are occasionally greatly facilitated by the retention of the length measure, as shown in (4). The symbol "l" is also interpretable as the length measure along the wavelength spectrum. Furthermore, since both line spectra and continuous spectra are represented by the set of nonnegative real numbers, l can be used to denote either the Lebesgue or Riemann measure on that set if continuous spectra are envisioned, or the Stieltjes measure, if line spectra are considered. The particular choice of the nature of l will be clear by convention or from the context in each case. Thus, unless specifically stated otherwise, l is to be considered as the usual Riemann type length measure used in ordinary calculus, and we conventionally consider continuous spectra. For integrations over wavelength space, use the transformation (32) of Sec. 2.12.

The second pair of properties of Φ is associated with surfaces. For every two disjoint surfaces S_1 and S_2,

$$\Phi(S_1,D,t,F) + \Phi(S_2,D,t,F) = \Phi(S_1 \cup S_2,D,t,F) \qquad (5)$$

and

$$\text{If } A(S) = 0, \text{ then } \Phi(S,D,t,F) = 0 \qquad (6)$$

These properties hold for arbitrary D,t and F. The first of these is the S-*additivity property* of Φ. The second is the S-*continuity property* of Φ.

The S-additivity property is understood as follows. Suppose the radiant flux meter has a variable collecting surface S, so that at one time it can be of extent S_1 and at another time (very soon after) it can be of extent S_2, such that S_1 and S_2 are disjoint. Then (5) states that the sum of the two separate readings associated with S_1 and S_2 equals the reading associated with the union $S_1 \cup S_2$ of these surfaces. This experimental fact is generally valid, provided of course, that D, t, and F are fixed as closely as practicable throughout all three measurements. Statement (5) is the ideal indicated by accumulated empirical findings. Statement (6) is also intuitively clear: positive amounts of flux can only be recorded over surfaces of positive area. This relatively innocuous pair of properties of Φ comprises the logical root of the concepts of irradiance and radiant emittance, to be

considered later.

We now turn to the third and final pair of fundamental properties of the radiant flux function Φ. These are associated with the direction set D. For every two disjoint direction sets D_1 and D_2:

$$\Phi(S,D_1,t,F) + \Phi(S,D_2,t,F) = \Phi(S,D_1 \cup D_2,t,F) \tag{7}$$

and

$$\text{If} \quad \Omega(D) = 0, \quad \text{then} \quad \Phi(S,D,t,F) = 0 \tag{8}$$

These properties hold for arbitrary S,t,F. The first of these is the D-*additivity property* of Φ, the second is the D-*continuity property* of Φ. These properties along with the preceding four will lead to the rigorous basis for the discussions of radiance, irradiance and related radiometric concepts.

The meaning of the D-additivity property is perhaps the most interesting of all the additivity properties, for it shows most clearly that on the level of reality within which radiometry conventionally takes place, the interference phenomena of light are not discernable: the light fields are comprised of incoherent electromagnetic fields. Examples are abundant on the microscopic level of light phenomena which illustrate the negation of (7), namely that for some S, t and F, there exist disjoint sets D_1 and D_2 such that:

$$\Phi(S,D_1,t,F) + \Phi(S,D_2,t,F) \neq \Phi(S,D_1 \cup D_2,t,F) \quad .$$

Therefore, the left side can be either > or < the right side. Furthermore, the negation of (8) holds on the microscopic level, too. That is, for some S,D,t,F, we have:

$$\Omega(D) = 0 \quad \text{and} \quad \Phi(S,D,t,F) \neq 0 \quad .$$

The first of these inequalities may be illustrated by means of any diffraction arrangement; the second arises every time Maxwellian electromagnetic theory is applied to a *plane* electromagnetic wave. In such a case D consists of a single direction--the direction of propagation--and $\Phi(S,D,t,F)$ is computed by means of the Poynting vector (cf. Sec. 2.2). It is principally through the properties (5)-(8) that one may discern the differences between the electromagnetic and phenomenological views of light, as far as logical form is concerned. That is, if we assume (5)-(8), then with a few additional physical-process and logical requirements which are common to both the electromagnetic and phenomenological views of light, the fundamental equations of radiative transfer theory are logically deducible. This may be seen, for example, by studying the results of Ref. [251]. Thus the electromagnetic and the phenomenological views of light necessarily part ways in (5)-(8). In a similar manner the phenomenological and quantum views of light differ at the same two points as above and possibly also at property (2). For, in the quantum theory, as in electromagnetic theory, radiant flux of a

single frequency as carried by a single photon (or a pure monochromatic wave of light) exists in principle. Therefore, it is possible that $\Phi(S,D,t,\{\nu\})>0$ for some set F consisting of a single frequency ν, and that $l(F) = 0$ at the same time. This follows from use of the usual measure of length on the frequency domain. If one redefines length on the frequency domain by adopting a Stieltjes measure, e.g., so that isolated single frequencies are given nonzero (usually unit) length (instead of the zero length we conventionally assigned them by the usual continuum measure) then (2) would hold on the electromagnetic and quantum levels too, and (5)-(8) remain as the source of the fundamental distinctions between the microscopic and the macroscopic views of light.

2.4 Irradiance and Radiant Emittance

We now turn to the task of defining the radiometric concepts used in radiative transfer in general and hydrologic optics in particular. The first two of these are the concepts of irradiance and radiant emittance. These concepts describe the flow of radiant energy per unit area across a surface. That is, they describe the area-density of radiant flux. Irradiance describes the flow *onto* a unit area; radiant emittance describes the flow *from* a unit area. From a strictly geometric point of view, this is the only distinction between the two concepts. However, radiant emittance occasionally has an additional physical connotation associated with it, namely that of a flow of radiant flux from a unit area of surface which encloses an emitting source of radiant flux, i.e., a region manufacturing radiant energy. However, within the operational definitions of these concepts, this additional connotation does not exist; the connotation exists only in the mind of the experimenter. We now turn to the detailed definitions of these concepts.

Definition of Irradiance

We begin with the concept of irradiance. Imagine a radiant flux meter transported to a point x in a natural hydrosol, or in the atmosphere. Let the collecting surface S of the meter be placed so that x falls within its small expanse, and orient the set D of directions of the meter as desired. A filter is fitted on the meter so as to pass monochromatic radiant flux of given frequency ν. Hence, the meter can be made to read P(S,D) directly (with ν and t and their units understood). Let "A(S)" denote the area of the collecting surface S. Then we shall write:

$$\boxed{\text{"H(S,D)"} \quad \text{for} \quad P(S,D)/A(S)} \tag{1}$$

and call H(S,D) the (*empirical*) *irradiance* over S within D. In full notation for the unpolarized context, we would write:

$$\text{"H(S,D,t,F)"} \quad \text{for } \Phi(S,D,t,F)/A(S) \tag{2}$$

or

$$\text{"}H(S,D,t,\nu)\text{"} \quad \text{for} \quad P(S,D,t,\nu)/A(S) \quad . \tag{3}$$

However, in most discussions of radiative transfer in hydrologic and meteorologic optics the light field is steady in time, and is studied frequency by frequency. Hence we shall until further notice hold t and ν (or F) fixed and so exclude their symbols and units from the notation, as in (1).

Next, we let S become smaller and smaller, such that it always contains the point x and such that the flow of radiant energy is onto S along the fixed set D of directions. Then we write:

$$\text{"}H(x,D)\text{"} \quad \text{for} \quad \lim_{S \to \{x\}} H(S,D) \quad . \tag{4}$$

The existence of this limit is guaranteed by the S-additive and S-continuity properties of Φ postulated in Sec. 2.3. The irradiance H(x,D) is the (*theoretical*) *irradiance* at x within D. The dimensions of both empirical and theoretical irradiance are *radiant flux per unit area* (per unit frequency interval); convenient units are *watts/(meter)*^{2} (per unit frequency interval).

It is of interest to see the meaning of H(x,D) in terms of the radiant flux function Φ of Sec. 2.2. Thus, from (4) and (3) (making ν and t explicit for the moment):

$$H(x,D,t,\nu) = \lim_{S \to \{x\}} P(S,D,t,\nu)/A(S)$$

From Sec. 2.3 this becomes:

$$H(x,D,t,\nu) = \lim_{S \to \{x\}} \left[\lim_{F \to \{\nu\}} \frac{\Phi(S,D,t,F)}{1(F)\ A(S)} \right]$$

$$= \lim_{S \to \{x\}} \left[\lim_{F \to \{\nu\}} \frac{H(S,D,t,F)}{1(F)} \right] \tag{5}$$

It follows from (3) above and a theorem of calculus that:

$$P(S,D,t,\nu) = \int_S H(x,D,t,\nu)\ dA(x) \tag{6}$$

and hence from (4) of Sec. 2.3 that:

$$\Phi(S,D,t,F) = \int_S \int_F H(x,D,t,\nu)\ d1(\nu)\ dA(x) \tag{7}$$

It is easy to see that these integrals can be generalized to the case where D in $H(x,D,t,\nu)$ may vary with x, and we shall understand that (6) and (7) hold in such cases.

In actual practice, the size of the collecting surface S, which serves to accept, diffuse, and transmit the incident flux on to the photoelectric element of the meter, ranges from the size of a pinhead to that of a dinner plate. These extremes are not intended to be precise limits; rather they are representative of the extremes that may be encountered in natural radiometric environments under ordinary working conditions. The lower limit cited above begins to approach the size where, for very sensitive photoelectric elements, effects of diffraction may be noticeable. For example, an ordinary household stickpin or a human hair held in a pencil-thin shaft of sunlight will cast a shadow with a diffraction pattern clearly discernable by the unaided human eye. Hence a very small radiant-flux meter collecting-surface can pick up such irradiance variations over the shadow. The upper limit cited above is dictated by the fact that natural lighting variations become noticeable over such relatively large areal extents: changes of lighting with depth in ponds or oceans, shadows cast by leaves or fish, edges of dense cloud shadows on the ground or a sea surface, etc. Hence by staying within these limits and choosing the size of S accordingly, good empirical estimates of irradiance can usually be made using the definition (1).

The Meaning of 'Irradiance'

It is occasionally helpful in both theoretical and experimental considerations to keep in mind the various meanings of 'radiant flux' discussed in Sec. 2.2. These may be applied directly to the concept of irradiance. Thus $H(S,D)$ may be imagined as proportional to the number of lines of flux incident per unit area over S and whose directions at their points of intersection with S lie within the set D. Further, using the Poynting vector interpretation of radiant flux, we see that the dimensions of the vector are precisely those of irradiance. Finally, $H(S,D)$ may be viewed as a measure of the number of photons per unit area per unit time on S, funneling down onto S along the directions of D. In particular, using (2) above and (2) of Sec. 2.2 for a monochromatic set of n photons over a small collecting area, and incident within a small set D of directions normally on S, we have:

$$H(S,D,t,\{\nu\}) = \frac{\Phi(S,D,t,\{\nu\})}{A(S)} = h\nu\nu n\Omega(D) \qquad (8)$$

A further insight into the concept of irradiance is obtained by considering some of the typical magnitudes of irradiance encountered in natural environments. Table I lists some of these values. They are order-of-magnitude estimates and are not to be used beyond establishing an intuitive feeling for the meaning of irradiance (see also Sec. 1.2).

TABLE 1

ENVIRONMENT	TYPICAL ORDER OF MAGNITUDE OF IRRADIANCE
At sea level, on surface S normal to sun's rays, clear day*	10^3 watt/m^2
At sea level, slightly overcast days, horizontal surfaces	10^2 watt/m^2
At sea level, heavily overcast day, horizontal surface S (sunset)	10 watt/m^2
Lighted interiors: walls, ceilings, floors	1 watt/m^2
At sea level, clear night, high full moon, horizontal suface S	10^{-3} watt/m^2
At sea level, clear night, flux from 1st magnitude (highly visible) star, on surface S normal to star's rays	10^{-9} watt/m^2

As another base for comparison and also to extend our intuitive feeling for irradiance and its connection with the photon picture of light, let us calculate the number of photons per unit volume, of wavelength λ, required to produce H watt/m^2 at a point of some surface. To fix ideas, suppose a thin pencil of photons arrives at each point x of a surface S in the direction of its inward normal ξ, and that each pencil is of the same density comprised of photons of a single frequency ν. It follows that the photon density $n(x,\xi,t,\nu)$ has the form

$$n(x,\xi',t,\nu') = n_o(x,t)\ \delta(\xi'-\xi)\ \delta(\nu'-\nu)$$

where δ is the Dirac delta function and where ξ is the inward normal to S, and ν is the frequency associated with λ. When used in (1) of Sec. 2.2, this equation yields:

*According to Moon, Ref. [185], at sea level, for sun zenith, clear dry air, the irradiance is nearly 1200 watt/m^2. See also [296] for a survey of solar irradiation measurements.

$$h\nu\nu \int_D \int_S \int_F n_o(x,t)\delta(\xi'-\xi)\delta(\nu'-\nu)\, dA(x)\, d\Omega(\xi')\, d1(\nu')$$

$$= h\nu\nu \int_S n_o(x,t)\, dA(x) = h\nu\nu n_o A(S)$$

as the radiant flux crossing S normally at time t. Hence $h\nu\nu n_o$ is the irradiance produced by each pencil. Setting

$$h\nu\nu n_o = H \text{ watt/m}^2$$

we have:

$$n_o = \frac{H}{h\nu\nu} \frac{\text{photons}}{\text{m}^3}$$

or

$$n_o v = \frac{H}{h\nu} \frac{\text{photons}}{\text{sec} \times \text{m}^2}$$

or

$$n_o v = \frac{H\lambda}{h\nu} \frac{\text{photons}}{\text{sec} \times \text{m}^2}$$

For example, setting $H = 1$ watt/m^2, $\lambda = 550$ mμ, and recalling that $h = 6.6 \times 10^{-34}$ Joule sec/photon and $v = 3 \times 10^8$ m/sec, we have

$n_o v$ = number of photons of wavelength 550 mμ per sec. normally incident per square meter which produce one watt

$$= \frac{5.5 \times 10^{-7}}{6.6 \times 10^{-34} \times 3 \times 10^8}$$

$$= 2.8 \times 10^{18}$$

From Table 1 we see that the normal irradiance of a first magnitude star is on the order of 10^{-9} watt/m^2. If we assume this flux to be comprised of photons all of wavelength $\lambda = 550$ mμ, then the number of associated photons producing this irradiance is $2.8 \times 10^{18} \times 10^{-9} = 2.8 \times 10^9$ photons per second normally incident per square meter $= 2.8 \times 10^5$ photons per second normally incident per square centimeter. Now a human eye's pupil is on the order of a tenth of a square centimeter in area. Hence when our eyes are directed toward a first magnitude star such as the present one, about 2.8×10^4 photons per second enter each eye to produce the visual sensation of brightness in the brain.

Terrestrial Coordinate Systems

Irradiance measurements and other radiometric measurements of hydrologic and meteorologic optics during careful experimental investigations are usually made with respect to either one of two *terrestrially-based frames of reference*. Each reference frame uses the usual euclidean three-dimensional coordinate system with its familiar xyz-axes. The two terrestrially-based reference frames are primarily distinguished by the way they anchor the directions of the x and z axes in each case. See Fig. 2.3. The *sun-based frame* directs the plane determined by the x-axis and z-axis (the xz *plane*) so as to contain the center of the sun. (The *north-based frame* directs the xz plane so as to lie in the plane of the local meridian circle on the earth.) In each frame the z-axis is parallel to the local vertical direction, (i.e., the local gradient of the gravitational field). In meteorologic optics z is measured as increasing in the upward direction, i.e., the unit vector k along the z-axis. In hydrologic optics it is more convenient to measure z as increasing in the downward direction -k, as shown in Fig. 2.3. In meteorologic optics, "z" (or other symbols) denotes *altitude*, in hydrologic optics, "z" (or other symbols) denotes *depth*, when specific reference to terrestrial coordinate frames is made.

The concept of direction within a reference frame established for a natural optical medium such as the atmosphere or the sea is of central importance in hydrologic optics and ranks equally in importance with the notion of location. In view of this importance it will be well to define with care precisely what is meant by "direction", and to develop some of the more frequently occurring concepts associated with it.

Now, to locate an object within a terrestrially-based reference frame, it suffices to give the x,y and z coordinates in terms of meters, say. Thus, in the hydrologic optics reference frames, the triple of numbers (1, 10, 100) locates a point in a natural hydrosol by going one meter along the direction i from the origin, then 10 meters along the direction j, and then 100 meters vertically downward. (Recall that in natural hydrosols, z is measured positive in the downward direction, i.e., in the direction -k.) Now this point obviously lies in a well defined "direction" from the origin of the reference frame. We observe that this "direction", however, has nothing to do with the distance of (1, 10, 100) from the origin. Indeed, the points (1/2, 5, 50) and (2, 20, 200) which are, respectively, half and twice as far from the origin as the original point, all lie in the same "direction" from the origin. A convenient measure of this common "direction" of all three points then would be established if we chose a point some standard fixed distance from the origin and which shares the same "direction" as they do. The obvious choice is the point a unit distance from the origin. Thus, if (x,y,z) is a point in a terrestrial frame of reference, then $(x,y,z)/(x^2+y^2+z^2)^{1/2}$ is a point a unit distance from the origin. We call this latter point the *direction* of (x,y,z) from the origin.

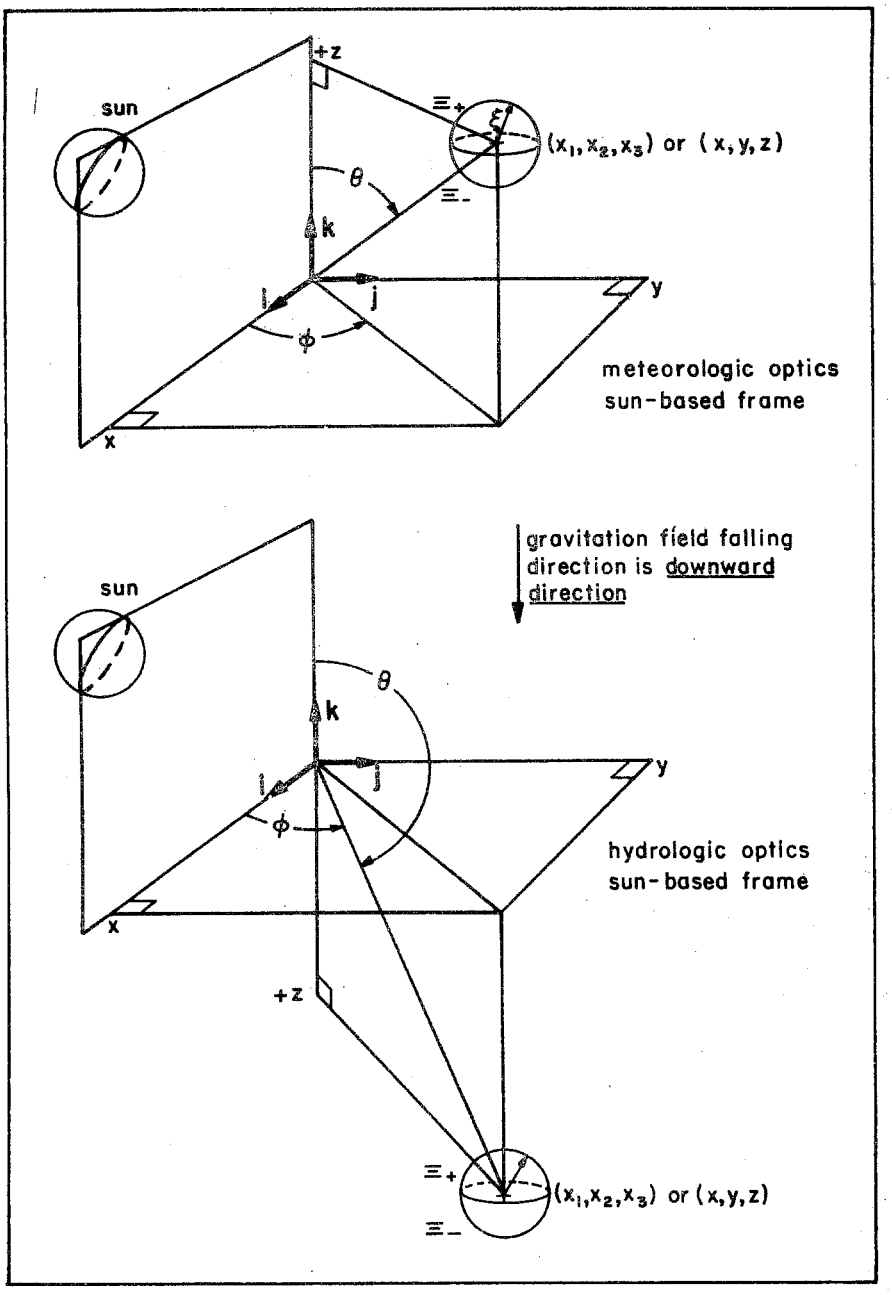


FIG. 2.3 Sun-based terrestrial frames of reference for meteorologic optics and hydrologic optics.

In many of our discussions we shall not need to specify explicitly the three coordinates of a point. In such cases we will simply write:

$$\text{"x"} \quad \text{for} \quad (x,y,z)$$

or

$$\text{"x"} \quad \text{for} \quad (x_1,x_2,x_3)$$

where the ordered triplets are the three coordinates of point x. Further we shall correspondingly write:

$$\text{"}\xi\text{"} \quad \text{for} \quad (x,y,z)/(x^2+y^2+z^2)^{1/2}$$

or

$$\text{"}\xi\text{"} \quad \text{for} \quad (x_1,x_2,x_3)/(x_1^2+x_2^2+x_3^2)^{1/2}$$

Hence, throughout this work the letter "x" (in either lightface or boldface type, as emphasis requires) is generally to designate a location and the letter "ξ" is generally to designate a direction. The denotation of the components of x and ξ will vary so as to permit simplicity and clarity of expression. We have already used the three special directions i, j, k, which we have agreed to be the points (1,0,0),(0,1,0), and (0,0,1), respectively.

We will also wish to consider collections of directions in addition to single directions. For example, certain sets D were already encountered in our discussions above. In particular, let us denote by "Ξ" the *set of all directions about the origin*. Clearly Ξ is a sphere of unit radius with origin as center. Observe that we use an upper case Greek Xi (the Greek counterpart to the English letter "X") to designate the set of all directions. There are two more sets of directions of very frequent occurrence in practice. First, there is the *set of all upward directions*, i.e., the set of all directions ξ such that ξ and **k** make an angle of less than ninety degrees. We shall designate this set by "Ξ_+". Second, there is the *set of all downward directions*, i.e., the set of all directions ξ such that ξ and **k** make an angle of greater that ninety degrees. We shall designate this set by "Ξ_-". The "+" and "-" are convenient mnemonics which help distinguish one set from the other. The reader may recall from vector analysis at this point that if ξ is in Ξ_+ then $\xi\cdot\mathbf{k} > 0$, i.e., the dot (or scalar) product of the vectors ξ and **k** is a positive number; and that if ξ is in Ξ_-, then $\xi\cdot\mathbf{k}$ is a negative number. This is the reason for the plus and minus signs in the names "Ξ_+" and "Ξ_-". Indeed, it would be well to recall that for every direction ξ,

$$\xi\cdot\mathbf{k} = \cos\theta$$

where θ is the angle between the lines along which ξ and **k** lie. See Fig. 2.3. For convenience we reproduce below the definition of the dot product of two unit vectors ξ_1 and ξ_2. Suppose we have written:

$$\text{"}\xi_1\text{"} \quad \text{for} \quad (a_1,b_1,c_1)/(a_1^2+b_1^2+c_1^2)^{1/2}$$

$$\text{"}\xi_2\text{"} \quad \text{for} \quad (a_2,b_2,c_2)/(a_2^2+b_2^2+c_2^2)^{1/2}$$

Then we write:

$$\text{"}\xi_1\cdot\xi_2\text{"} \quad \text{for} \quad \frac{a_1a_2+b_1b_2+c_1c_2}{(a_1^2+b_1^2+c_1^2)^{1/2}(a_2^2+b_2^2+c_2^2)^{1/2}}$$

From analytic geometry it is known that:

$$\xi_1\cdot\xi_2 = \cos \mathscr{v}$$

where $\mathscr{v}$ is the angle between ξ_1 and ξ_2.

The representation of a unit vector ξ as an ordered triple of numbers takes on deeper meaning when we observe the following geometric fact. Let "(a,b,c)" denote the ordered triple representation of ξ. Then compute the dot product of ξ with **i**, **j**, and **k** in turn. By the cosine law cited above we have:

$$\xi\cdot\mathbf{i} = \cos \mathscr{v}_1$$

$$\xi\cdot\mathbf{j} = \cos \mathscr{v}_2$$

$$\xi\cdot\mathbf{k} = \cos \mathscr{v}_3$$

where $\mathscr{v}_1$, $\mathscr{v}_2$, and $\mathscr{v}_3$ are, respectively, the angles between ξ and the positive x, y and z axes. Using the ordered triple representations of ξ, **i**, **j**, and **k**, and the definition of the dot product, we have:

$$\xi\cdot\mathbf{i} = a$$

$$\xi\cdot\mathbf{j} = b$$

$$\xi\cdot\mathbf{k} = c$$

Hence the components a,b,c of the direction ξ are simply the cosines of the angles that ξ makes with the positive x,y and z axes, i.e.:

$$a = \cos \mathscr{v}_1$$

$$b = \cos \mathscr{v}_2$$

$$c = \cos \mathscr{v}_3$$

This leads to the representation:

$$\xi = (\cos \mathcal{V}_1, \cos \mathcal{V}_2, \cos \mathcal{V}_3) .$$
$$= \mathbf{i} \cos \mathcal{V}_1 + \mathbf{j} \cos \mathcal{V}_2 + \mathbf{k} \cos \mathcal{V}_3 .$$

where we have written:

"i" for (1,0,0)

"j" for (0,1,0)

"k" for (0,0,1)

There is an alternate mode of representation of a unit vector ξ. This alternate mode attains its greatest utility in actual calculation. This is the representation of ξ by two especially constructed angles θ and ϕ, found as follows. By studying the schematic representation of Ξ in Fig. 2.4 it is clear that since each direction of Ξ has fixed known length (namely a unit length) it suffices to uniquely specify ξ by the angle it makes with the z-axis and the angle the plane determined by ξ and k makes with the xz plane. Suppose we designate the former angle by "θ", the latter by "ϕ", and agree to set $\theta = 0$ when $\xi = \mathbf{k}$. Further, we agree to have θ increase to π when $\xi = -\mathbf{k}$. Further, we agree to set $\phi = 0$ when ξ is in the xz plane and to have ϕ increase to $\pi/2$ as the plane of ξ and k rotates from the xz plane to the yz plane. We let ϕ increase in like manner through the next three quadrants, and finally have it measure 2π radians after one complete turn in this manner. To summarize this alternate mode

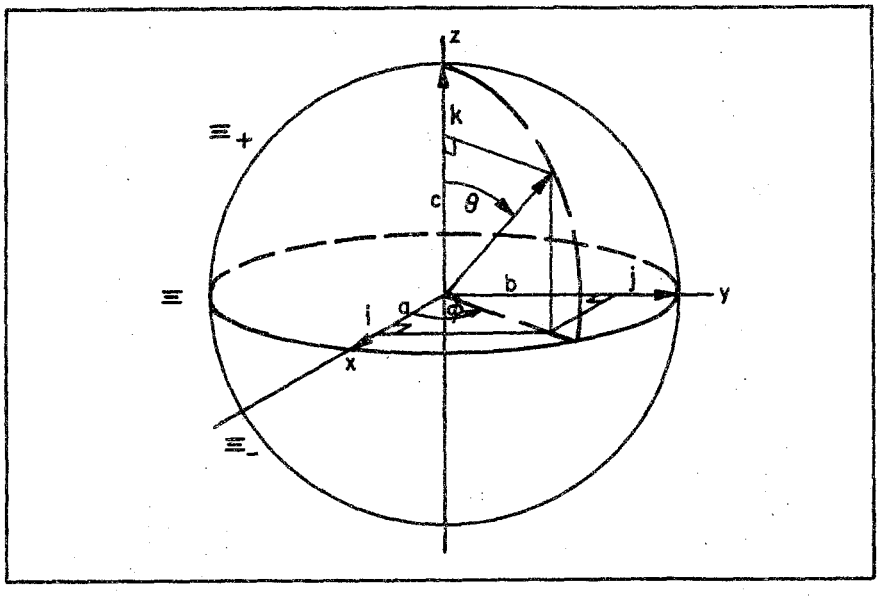


FIG. 2.4 Angle and direction definitions

of representation of ξ, we agree to write:

$$\text{"}(\theta,\phi)\text{"} \quad \text{for} \quad \xi$$

whenever ξ is in Ξ, and whenever $\xi = (a,b,c)$, θ = arc cos c, and ϕ = arc tan b/a, and where the quadrant containing ϕ is fixed by the signs of a and b. The angle θ is the *polar (or zenith) angle* of ξ, and ϕ the *azimuthal angle* of ξ.

Representation of Irradiance in Terrestrial Frames

Let us return now to apply these geometrical results to the task of specifying irradiance in natural optical media such as the atmosphere or the sea. It has become clear after much theoretical and experimental work in natural aerosols and hydrosols that the type of irradiance which is used most often in practice is the irradiance on a *horizontal* surface S at point x with a set D of directions which constitutes either the *hemisphere* Ξ_+ or Ξ_- of the unit sphere Ξ. To specify such irradiances, we return to the definition in (4), and replace "D" first by "Ξ_+" and then by "Ξ_-" (or by "+" and "-"). Thus $H(x,\Xi_+)$ (or $H(x,+)$) is the irradiance at point x induced by upward flowing radiant energy in the directions of Ξ_+, and $H(x,\Xi_-)$ (or $H(x,-)$) is the irradiance at point x induced by downward flowing radiant energy in the directions of Ξ_-.

A further specialization in notation can take place when the medium is stratified. Now, a natural optical medium (or a light field) with a terrestrially-based reference frame (Fig. 2.3) is said to be *stratified* if and only if the optical properties of the medium (or light field) as functions of coordinates x,y,z, are independent of the coordinates x and y. Thus for stratified light fields we may, for brevity and without loss of information, replace the general point name "x" in $H(x,\Xi_+)$ by "z", the depth-parameter name. Thus, let us agree henceforth in stratified natural optical media to write:

$$\text{"}H(z,+)\text{"} \quad \text{for} \quad H(x,\Xi_+) \tag{9}$$

and

$$\text{"}H(z,-)\text{"} \quad \text{for} \quad H(x,\Xi_-) \tag{10}$$

We call $H(z,+)$ the *upward irradiance* and $H(z,-)$ the *downward irradiance*.

The next most frequently occurring type of irradiance $H(x,D)$ after the types $H(z,\pm)$, is that for which D is an arbitrarily oriented hemisphere. Thus, let us denote by "$\Xi(\xi)$" that part of Ξ consisting of all unit vectors ξ' such that ξ' and ξ subtend an angle less than ninety degrees. Hence, after adapting definition (4) to the case where D is $\Xi(\xi)$, we have $H(x,\Xi(\xi))$ as the irradiance at point x on a collecting surface S with unit inward normal ξ, such that the irradiance is produced by radiant flux incident on S at x along the directions within $\Xi(\xi)$. See Fig. 2.5. Observe that the irradiance $H(x,\Xi(\mathbf{k}))$ is simply $H(x,\Xi_+)$ considered earlier, since $\Xi(\mathbf{k}) = \Xi_+$; and similarly $H(x,\Xi(-\mathbf{k})) = H(x,\Xi_-)$. Now a useful fact about such sets of directions as $\Xi(\xi)$ is that they are

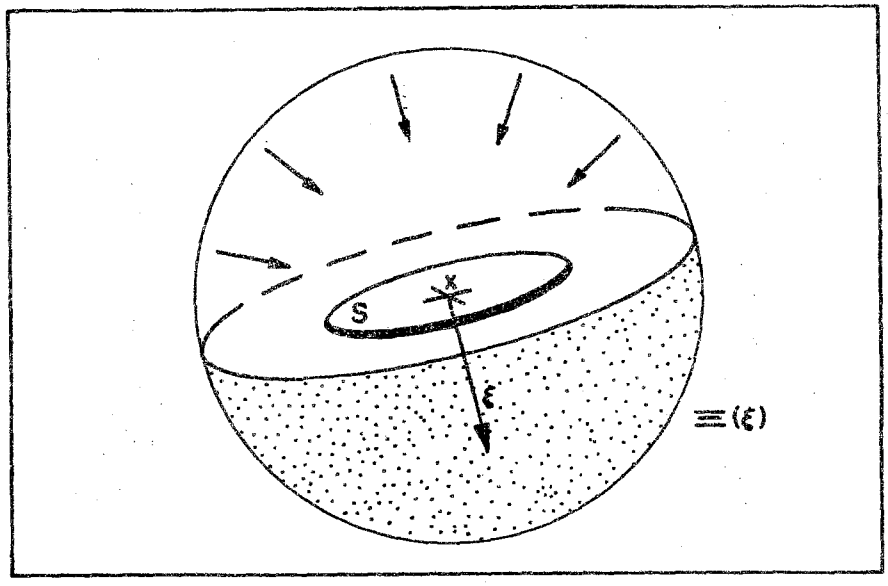


FIG. 2.5 Defining the hemisphere $\Xi(\xi)$ determined by the direction ξ

uniquely specified by giving the single vector ξ. We take advantage of this observation to shorten the irradiance notation by agreeing henceforth in (4), for the case $D = \Xi(\xi)$, to write:*

$$\text{"}H(x,\xi)\text{"} \quad \text{for} \quad H(x,\Xi(\xi)) \tag{11}$$

If we restrict attention to a fixed point x, then the totality of all values $H(x,\xi)$ as ξ varies over Ξ is called the *irradiance distribution* at x. If the light field is stratified we further agree to write:

$$\text{"}H(z,\xi)\text{"} \quad \text{for} \quad H(x,\xi) \tag{12}$$

Thus in stratified light fields, one may specify irradiances by giving a depth z and the *unit inward* normal ξ to a (hypothetical or real) collecting surface at that depth.

If one prefers to use the mode of representation of ξ by means of polar and azimuthal angles θ and ϕ, then it will be agreed to write:

---

*Whenever wavelength dependence and time dependence is to be shown explicitly we would use "$H(x,\Xi(\xi),\lambda)$", or "$H(x,\Xi(\xi),t)$" or "$H(x,\Xi(\xi),t,\lambda)$" as the case may be, and in contracted ξ-notation, as desired.

$$\text{"}H(x,\theta,\phi)\text{"} \quad \text{for} \quad H(x,\xi) \tag{13}$$

or

$$\text{"}H(z,\theta,\phi)\text{"} \quad \text{for} \quad H(z,\xi)$$

when the light field is stratified. It should be re-emphasized that the direction ξ (and hence (θ,ϕ)) refers to the unit *inward* normal to the collecting surface S in the operational definition of (13) *and that the flow of photons is onto* S *at* x *along the directions of* $\Xi(\xi)$. This is the convention we shall adopt when discussing irradiance measurements by collecting surfaces on a theoretical level; for the transport equations for $H(z,\pm)$ to be introduced later (Chapter 8) are written down in an intuitively natural manner if this convention is adopted. The convention may be altered if need be for empirical discussions. However, it is perhaps needless to point out that the fewer such conventions actually adopted for radiometers, the smaller will be the chance of conceptual chaos in practice.

One final definition, and then we shall be ready for a discussion of the cosine law for irradiance. We agree to write:

$$\text{"}\bar{H}(x,\xi)\text{"} \quad \text{for} \quad H(x,\xi) - H(x,-\xi) \tag{14}$$

and call $\bar{H}(x,\xi)$ the *net irradiance* at x in the direction ξ.

The Cosine Law for Irradiance

We now consider the property of irradiance which is its most important and most frequently used theoretical property. This is the cosine law for irradiance. The law is based on the simple geometric fact that the apparent area of a small plane surface at a fixed distance along one's line of sight varies as the cosine of the angle between the line of sight and the normal to the surface. If now we direct a swarm of photons along the line of sight toward the small surface then, all other things being equal, the area will intercept a number of photons proportional to the apparent area, i.e., proportional to the cosine of the angle between the direction of the beam of photons and the surface's normal. Hence the area density, i.e., the irradiance of the photons on the surface will vary as the cosine of this angle. The formal statement of this observation is the *cosine law* for irradiance. We now translate this verbal derivation of the cosine law into symbolic form.

In Fig. 2.6 a small plane surface is denoted by "S". An amount P(S,D) of radiant flux is incident on S and arrives at each point of S through a very narrow fixed conical solid angle D such that the central direction of D is normal to S. Since the radiant flux is limited to a relatively narrow bundle of directions, essentially all the lines of flux are confined to a cylindrical volume C in the immediate neighborhood of S. Let "S'" denote a section of C generated by a plane whose normal makes an angle α with the axis of C and such that

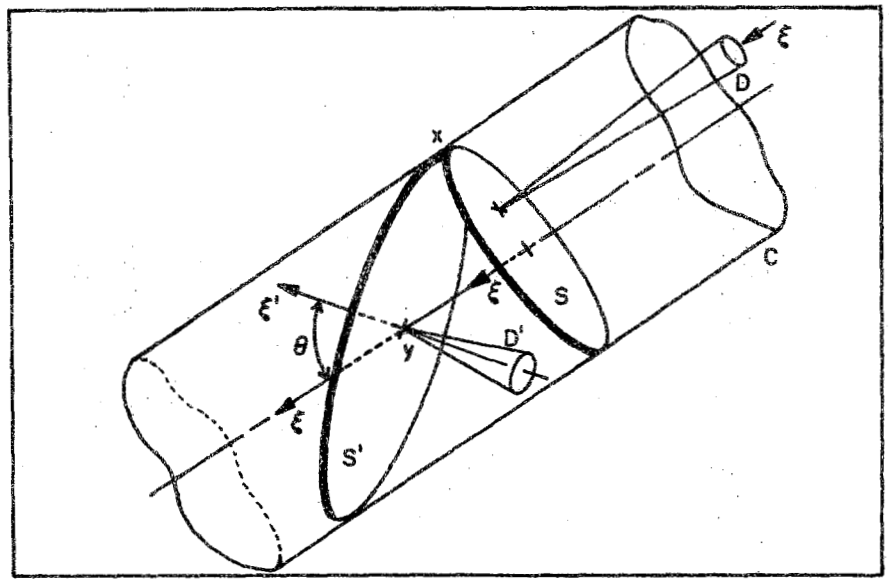


FIG. 2.6 Geometry of the cosine law for irradiance.

the plane goes through some point x on S. The area A(S') of S' is clearly related to the area A(S) of S by the relation:

$$A(S') = A(S) \sec \vartheta$$

Assuming no intervening sources or sinks of radiant flux in the region of C between S and S' the flux P(S,D) then also crosses S'. Thus we can write:

$$P(S,D) = P(S',D) \quad .$$

By definition, the area density H(S',D) of radiant flux across S' is:

$$H(S',D) = P(S',D)/A(S') \quad .$$

In view of the preceding flux conservation statement and the geometric relation between A(S') and A(S) we can write:

$$H(S',D) = P(S,D)/(A(S) \sec \vartheta)$$

By definition H(S,D) is P(S,D)/A(S) and we therefore arrive at the statement:

$$\boxed{H(S',D) = H(S,D) \cos \vartheta} \qquad (15)$$

This is the empirical form of the *cosine law for irradiance*. A theoretical form of the law is obtained by letting $S \to \{x\}$ (and hence $S' \to \{x\}$). The result is:

$$H(x,\xi') = H(x,\xi)\ \xi \cdot \xi' \qquad (16)$$

Here we have used the fact that D was sufficiently narrow so that in the limit $H(S,D)$ goes to $H(x,\xi)$ as S goes to the set $\{x\}$ consisting of point x. Further, $H(S',D)$ goes to $H(x,\xi')$ as S goes to $\{x\}$. Of course (16) is to be understood to apply to a set D of directions with a *small* but finite solid angle. The limiting case for $D \to \{\xi\}$ can be handled naturally only after the concept of radiance has been introduced. Further we have replaced "cos $\mathcal{V}$" by "$\xi \cdot \xi'$" in going from (15) to (16). After the introduction of the concept of vector irradiance (Sec. 2.8), (16) can readily be generalized to the case where the set of incident directions D is arbitrary.

Radiant Emittance

We close this section with a few comments on the concept of radiant emittance. As already noted in the introductory remarks to this section, the concept of radiant emittance is nearly identical to that of irradiance, differing from the latter geometrically only by the sense of flow of the radiant energy across a surface S. Fig. 2.7 schematically depicts the geometrical distinction between irradiance and radiant

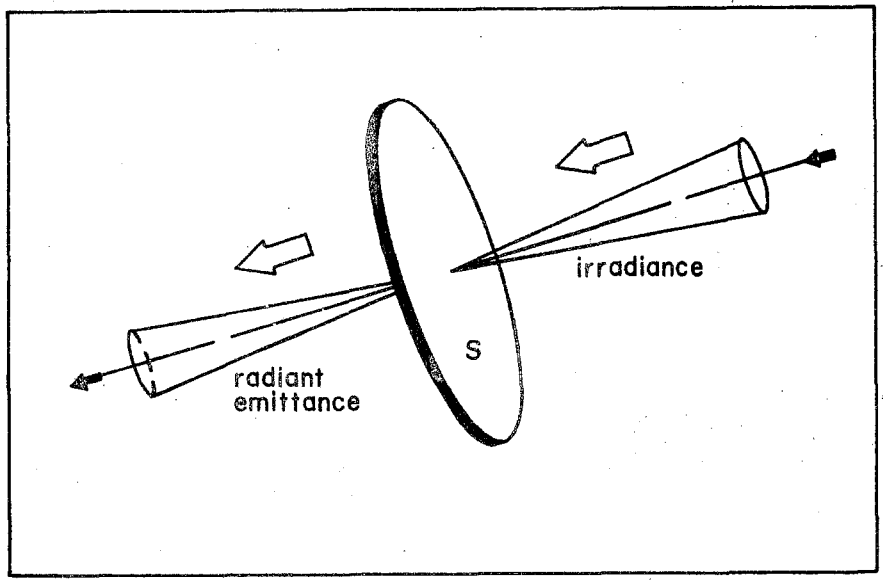


FIG. 2.7 Conceptual distinction between irradiance and radiant emittance.

emittance; a given parcel of radiant energy flowing *onto* a surface S generates *irradiance on* S: the same parcel flowing from the surface S *generates radiant emittance of* S. To emphasize this distinction and to have appropriate notation available when needed, we need only write "$\Phi^-(S,D)$" to denote radiant flux *onto* S and to write "$\Phi^+(S,D)$" for radiant flux *from* S. Then we extend this notation to radiant flux by means of "$P^-(S,D)$" and "$P^+(S,D)$". Thus, the definition (1) of empirical irradiance may be written as:

$$\text{"}H(S,D)\text{"} \quad \text{for} \quad P^-(S,D)/A(S) \tag{17}$$

for emphasis of the "onto" interpretation of the flux; and we now go on to write:

$$\text{"}W(S,D)\text{"} \quad \text{for} \quad P^+(S,D)/A(S) \tag{18}$$

for contrast of the two notions. We call W(S,D) the (empirical) *radiant emittance* over S within D. From consideration of Fig. 2.7 it is clear that in the context of that figure:

$$\boxed{P^+(S,D) = P^-(S,D)} \tag{19}$$

so that

$$\boxed{W(S,D) = H(S,D)\ .} \tag{20}$$

Another distinction between W(S.D) and H(S,D) for a given S and D lies on the physical rather than the geometric level. Indeed, it is on this level that the concept W(S,D) was originally conceived and arose in connection with the derivation of the complete (or Planckian) radiator wherein radiant flux is generated within a body and then emitted through its boundary. This interpretation will be used in Sec. 2.12 during the transition from radiometry to photometry.

We conclude by observing that every auxiliary geometric definition and geometric law considered above for irradiance now holds analogously for radiant emittance. We shall henceforth apply the analogous notation for W(S,D) (such as W(x,D), W(x,ξ), etc.) without further explicit definitions. Thus for example we write:

$$\text{"}W(x,D)\text{"} \quad \text{for} \quad \lim_{S\to\{x\}} W(S,D) \tag{21}$$

and

$$\text{"}W(x,\xi)\text{"} \quad \text{for} \quad W(x,\Xi(\xi)) \tag{22}$$

and so on.

2.5 Radiance

We now define the radiometric concept of radiance, discuss some of its various forms, and study some of its basic geometrical and physical properties of particular use in hydrologic optics.

For those who are studying the concept of radiance for the first time, we may introduce it by saying that radiance is designed to yield a simple mathematical representation of the percept of brightness experienced by the human eye as the eye is directed along various paths of sight. Some introspection will show that when one directs visual attention to a point in his environment, such as a point on a desk or a wall, the brightness sensations of neighboring points of the point under scrutiny can be willfully suppressed. The result is a possible conscious comparison of "brightness" of successive neighboring points in one's environment. Now when one attempts to simulate this sensation of brightness by means of a radiant flux meter, one must introduce a mechanical means of directing the 'attention' of the collecting surface along a narrow bundle of directions. The collecting surface by itself is obviously incapable of the complex and partly automatic process that takes place in the eye-brain circuits within a human head when visual attention is directed along a narrow bundle of directions. Some sort of "blinder", usually in the form of a long narrow circular cylinder, must be fitted around a circular collecting surface so that its axis is normal to the plane of the collecting surface. The result is a radiant flux meter with a relatively narrow conical set D of directions along which radiant flux may be incident on a plane circular collecting surface S. Such an assembly is depicted in Fig. 2.8, and is called a *radiance meter*.

The operational definition of the radiometric concept of radiance can be given in terms of a radiance meter as follows. The radiance meter is taken to a point x in a natural optical medium such as the atmosphere or a natural hydrosol. The center of the collecting surface is placed so as to be at point x. The axis of the cylindrical collecting tube of the meter is directed along a direction ξ so that radiant flux from the field of view is funneled along the set D in the general direction of ξ. The sensor component of the radiance meter records an incident radiant flux P(S,D) on the collecting surface. The area A(S) of S and the solid angle $\Omega(D)$ of D are known instrumental constants. The quotient:

$$P(S,D)/A(S)\Omega(D)$$

is called the (empirical) *radiance* at x along ξ. Radiance, therefore, is a nonnegative number which is paired with the dimensions of *power per area per solid angle* (per unit frequency interval), and with convenient units such as *watts per square meter per steradian* (per unit frequency interval). We will write:

$$\text{"N(S,D)"} \quad \text{for} \quad P(S,D)/A(S)\Omega(D) \tag{1}$$

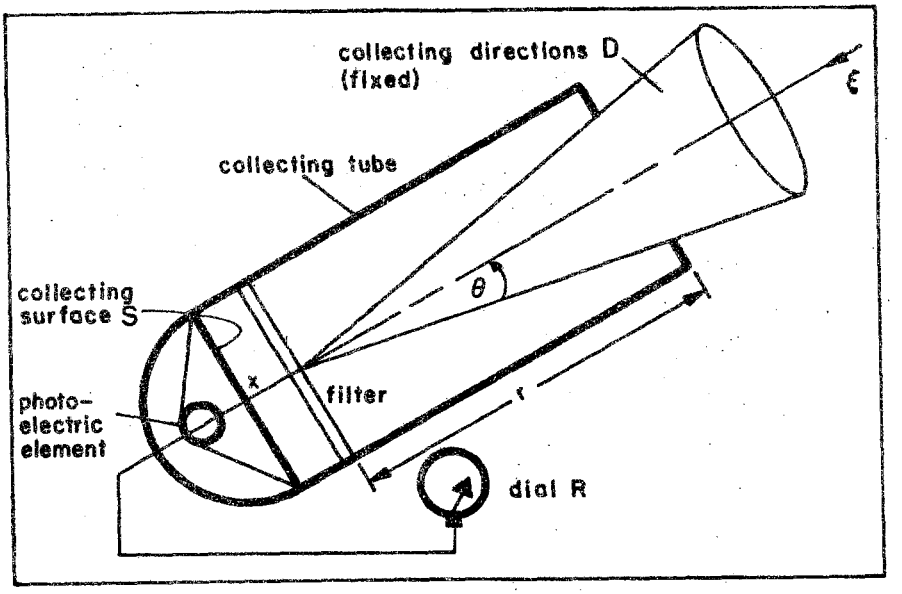


FIG. 2.8 Schematic details of a radiance meter

or in more complete notation:

$$\text{"}N(S,D,t,\nu)\text{"} \quad \text{for} \quad P(S,D,t,\nu)/A(S)\Omega(D) \quad .$$

Since $A(S)$ and $\Omega(D)$ are fixed numbers for a given radiance meter, the radiant flux reading can be calibrated directly in terms of $N(S,D)$. Experimentation with variously proportioned radiance meters indicates that those meters with solid angle magnitudes $\Omega(D)$ on the order of 1/30 steradians, and with collecting areas $A(S)$ on the order of that for a circular surface of a centimeter in diameter, are adequate for radiometry in most natural optical media. Of course, the smaller $A(S)$ and $\Omega(D)$, the sharper are the radiance maps obtainable (while still remaining above the level where diffraction and general interference effects set in).

Recall the definition of empirical irradiance $H(S,D)$ in (1) of Sec. 2.4. It follows from (1) above that:

$$\boxed{N(S,D) = H(S,D)/\Omega(D) \quad .} \tag{2}$$

Corresponding to (4) of Sec. 2.4 we shall write:

$$\text{"}N(x,D)\text{"} \quad \text{for} \quad H(x,D)/\Omega(D) \quad , \tag{3}$$

and

$$\text{"}N(S,\xi)\text{"} \quad \text{for} \quad \lim_{D \to \{\xi\}} H(S,D)/\Omega(D)$$

where the central direction of D is normal to the plane of S at x. We are then led to write:

$$\text{"}N(x,\xi)\text{"} \quad \text{for} \quad \lim_{D\to\{\xi\}} N(x,D) \quad . \tag{4}$$

We will occasionally use the more complete notations "N(S,D,t,F)", "N(x,ξ,t,F)", "N(x,ξ,t,ν)", etc., for radiance when time and frequency parameters are explicitly required. The time symbol "t" and the frequency symbol "ν" may be included or omitted as needed. In the case of the first of those listed above we agree to write:

$$\text{"}N(S,D,t,F)\text{"} \quad \text{for} \quad \Phi(S,D,t,F)/A(S)\Omega(D) \quad .$$

The quantity $N(x,\xi)$ is the (theoretical) *radiance* at x in the direction ξ. It exists as a mathematical entity by virtue of the D-additivity and D-continuity properties of Φ cited in (7) and (8) of Sec. 2.3.

It is instructive to disassemble the definition of theoretical radiance layer by layer until the primitive concept of radiant flux Φ is recovered. Thus, beginning with (4) and using (3):

$$N(x,\xi) = \lim_{D\to\{\xi\}} H(x,D)/\Omega(D) \quad .$$

Then by means of (4) and (1) of Sec. 2.4:

$$N(x,\xi) = \lim_{S\to\{x\}} \left(\lim_{D\to\{\xi\}} P(S,D)/A(S)\Omega(D) \right) \quad .$$

Finally, by means of (3) of Sec. 2.3 we have (in full notation):

$$N(x,\xi,t,\nu) = \lim_{S\to\{x\}} \left[\lim_{D\to\{\xi\}} \left(\lim_{F\to\{\nu\}} \frac{\Phi(S,D,t,F)}{A(S)\Omega(D)1(F)} \right) \right]$$

$$= \lim_{S\to\{x\}} \left[\lim_{D\to\{\xi\}} \left(\lim_{F\to\{\nu\}} \frac{N(S,D,t,F)}{1(F)} \right) \right] \quad . \tag{5}$$

This is the basis for the fact that, in the last analysis, *all radiometric concepts are reducible to the primitive physical concept of radiant flux embodied by Φ and the appropriate geometrical and analytical notions of limit and measure.* Hence all equations of pure and applied radiative transfer are resolvable into expressions containing only one primitive physical notion, namely Φ(S,D,t,F), and auxiliary geometrical and analytical concepts.*

---

* [Those who desire radiant energy as the most primitive physical notion, may then start with U(S,D,T,F) where T is a finite time interval, so that $\Phi(S,D,t,F) = \lim_{T\to\{t\}} U(S,D,T,F)/1(T)$. In Vol. I, U was taken as a primitive concept; in this and subsequent volumes, U will be derived from Φ as in (17) of Sec. 2.7, e.g.]

Now that the definition of radiance has been established it is an easy matter to return to the definition of the phase density $n(x,\xi,t,\nu)$ of photons in (1) of Sec. 2.2 and by energy and dimensional arguments conclude that

$$N(x,\xi,t,\nu) = h\nu\nu n(x,\xi,t,\nu) \tag{5a}$$

which together with (5) connects $N(x,\xi,t,\nu)$ with some of the more basic constructs of radiometry (Φ, and photons). Statement (5a) can be cast into terms of wavelength λ by using the transformation (32) of Sec. 2.12.

To gain some insight into the magnitudes of radiances found in nature, we append Tables 1,2, which are constructed from the graphs in parts III, IV of [26] and which form part of a four-part series of compilations of sky (field) radiance distributions. The skies in the present tabulation were morning (0800 hours) skies at sea level, covered 40% with scattered clouds. Two regions of the spectrum are considered: Table 1 gives orders of magnitude of field radiance in the wavelength interval [400, 500] mμ, and Table 2 is for the interval [580,700] mμ. The main purpose of the tables is to complete the statement: "daylight skies (away from the sun) have radiances on the order of 10^n watts/(m^2 × steradian), where n = ?" It is clear that the answer is around n= -1,0,1. By way of contrast, recall that the radiance over the sun's disk is on the order of 2×10^7 watts/(m^2 × steradian) as seen just outside the atmosphere, and for a wavelength interval $[0,\infty]$ mμ (cf., (3a) of Sec. 1.2). Hence N in the vicinity of the sun runs from 10^0 to 10^7 units of radiance. The data were taken June 21, 1958 in balloon flights over central Minnesota. For angle conventions, see part (a) of Fig. 2.3.

Table 1

Sample Radiances, Morning Skies

400-500 mμ, Sea level, sun zenith angle = 70°
watts/(m^2 × steradian)

Azimuthal ϕ / Polar θ	sun's azimuth $\phi = 0°$	$\phi = 80°$	$\phi = 180°$
$\theta = 0°$ zenith	1.5	1.5	1.5
$\theta = 45°$	5.3	2.1	2.0
$\theta = 90°$	7.0	2.2	1.5

Table 2

Sample Radiances, Morning Skies

580-700 mμ, Sea level, sun zenith angle = 70°
watts/(m² × steradian)

Azimuthal φ / Polar θ	sun's azimuth φ = 0°	φ = 80°	φ = 180°
θ = 0° zenith	.45	.45	.45
θ = 45°	3.0	.62	.45
θ = 90°	5.5	2.0	1.2

Radiance Distributions

We have seen how the operational definition of radiance leads to the theoretical radiance displayed in (4). This in turn leads to the construction of a function N which assigns to each point x in an optical medium and direction ξ at that point, a *radiance* $N(x,\xi)$ of the natural light field. $N(x,\xi)$ is the number of watts of radiant flux incident per unit solid angle, in the direction ξ normal to a unit area at x. Implicit in the notation is the time t of the measurement and the frequency ν of the energy passed by the filter of the meter. The totality of all values $N(x,\xi)$ paired to (x,ξ) as x ranges over all points of a selected optical medium X and as ξ ranges over the unit sphere Ξ is called the *radiance function* on $X \times \Xi$ and is denoted by "N". If attention is restricted to an arbitrary fixed point x and the totality of values $N(x,\xi)$ are considered for all ξ in Ξ, then that totality of values is called the *radiance distribution* at x and is denoted by "$N(x,\cdot)$".

The radiance function is the most important radiometric function in geophysical optics and in particular, in hydrologic optics. For an exhaustive empirical study of radiance distributions in a natural hydrosol, see the classic work of Tyler [298]. The importance of the radiance function rests in the fact that from knowledge of the radiance function alone, all other radiometric quantities are relatively easily calculable. This fact will become increasingly apparent as the discussion of this work proceeds, and we begin below with a first example of this fact. (See also Figs. 1.23-1.25)

Irradiance from Radiance

As an illustration of the use of the concept of radiance, and to aid the reader to fix in mind its definition, we shall derive the relation between irradiances of the form $H(x,\xi)$ introduced in Sec. 2.4 and radiances $N(x,\xi)$ just defined above. More detailed examples are reserved for Sec. 2.11.

We begin with the empirical connection between $H(S,D)$ and $N(S,D)$ established as a matter of course in equation (2). If $N(S,D)$ is known, we can compute $H(S,D)$ using:

$$H(S,D) = N(S,D)\Omega(D) \quad .$$

It should be recalled that D is a narrow conical set of directions associated with the radiance meter, and that the central direction of the cone is normal to the surface S.

We now apply this general relation to the following problem, which is formulated with the aid of Fig. 2.9. A surface S with inward normal is irradiated by n distinct sources of flux such that the i-th flux has radiance $N(S_i,D_i)$ and is incident on the points of S through a small conical set D_i of directions centered on direction ξ_i. The sets D_i are pairwise disjoint (i.e., no two overlap) and all lie on the same side of S. What is the resultant irradiance $H(S,D)$ produced by this given set of incident irradiances?

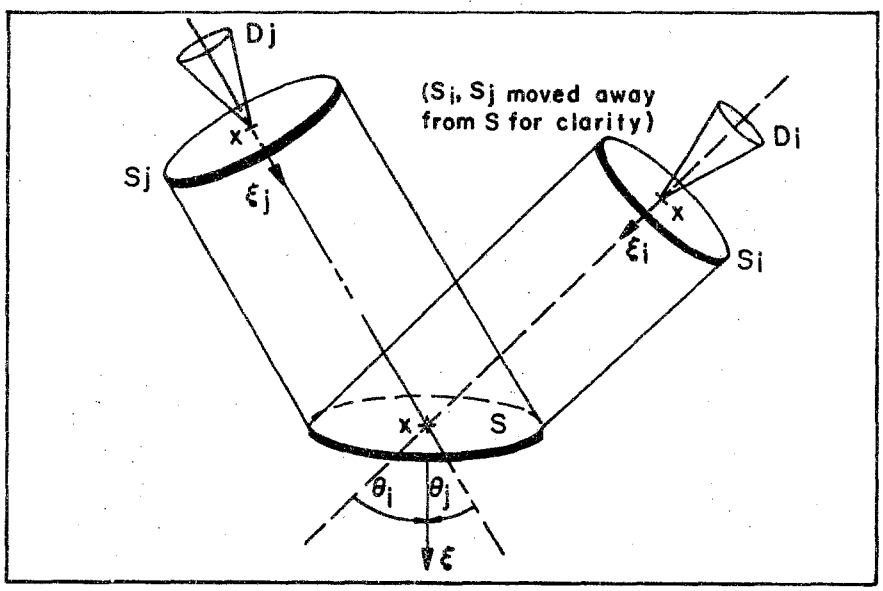


FIG. 2.9 Setting up the connection on going from radiance to irradiance

The first step in relating H(S,D) to the n radiances is to observe that by successive applications of the D-additivity property of radiant flux (equation (7) of Sec. 2.3) we can write:

$$H(S,D) = H(S,D_1) + H(S,D_2) + \ldots + H(S,D_n) \quad ,$$

where "$H(S,D_i)$" denotes the irradiance on S produced by radiant flux incident within the set of directions D_i. The second step consists in using the cosine law for irradiance (equation (15) of Sec. 2.4) to relate $H(S_i,D_i)$ and $H(S,D_i)$, for i = 1,..., n. Thus:

$$H(S,D_i) = H(S_i,D_i) \cos \vartheta_i \quad ,$$

where "ϑ_i" denotes the angle between the unit inward normal ξ to S and ξ_i. We have chosen S small enough so that the conditions of the derivation of the cosine law (15) or (16) of Sec. 2.4, are satisfied. Furthermore, we use (4) above to permit slight adjustments of the choice of the S_i as may be required to meet the cosine law derivation conditions without noticeably changing the value of the radiance of the flux on S_i through D_i. Thus, by definition, for every i = 1,...,n:

$$H(S_i,D_i) = N(S_i,D_i)\Omega(D_i)$$

and this constitutes the third and final step. By assembling the results of these three steps we have the desired connection:

$$H(S,D) = \sum_{i=1}^{n} N(S_i,D_i) \cos \vartheta_i \, \Omega(D_i) \quad . \qquad (6)$$

When n = 1, we have the intuitively useful special case of (6):

$$H(S,D) = N(S',D)\Omega(D) \cos \vartheta \quad ,$$

where we have written "S'" for S_1, and now $D = D_1$ in Fig.2.9.

The connection (6) is a useful relation in practical situations where knowledge of radiance distributions is applied to find irradiances on arbitrarily oriented surfaces. By using terrestrially based coordinate systems (Sec. 2.4) equation (6) can be translated into a workable standard computation procedure. This task is facilitated by first establishing the theoretical counterpart to (6). Thus, let $S \to \{x\}$, so that also $S_i \to \{x\}$. Then $H(S,D) \to H(x,D)$ and $N(S_i,D_i) \to N(x,D_i)$, according to (3). Equation (6) then becomes:

$$H(x,D) = \sum_{i=1}^{n} N(x,D_i) \cos \nu_i \, \Omega(D_i) \tag{7}$$

We now apply (7) to the case where Ξ is divided up into n disjoint pieces Ξ_i, and we then let the number n increase indefinitely so that each Ξ_i goes to zero. In this way we arrive at the integral counterpart to (7):

$$H(x,\xi) = \int_{\Xi(\xi)} N(x,\xi')\xi\cdot\xi' \, d\Omega(\xi') \tag{8}$$

Recall that the symbol "$\Xi(\xi)$" denotes that hemisphere of Ξ consisting of directions ξ' which make an angle less than ninety degrees with ξ. Further, ξ is the unit inward normal to the collecting surface S at x. Recall also that $\xi\cdot\xi'$, by the discussion of 2.4, equals the cosine of the angle ν between ξ and ξ'. Thus (8) may be rewritten in terms of ν. Before this can be done with complete clarity, we must express $d\Omega(\xi')$ in terms of polar and azimuth angles. This we shall do, taking the opportunity to explicate at the same time the notion of "solid angle".

Toward this end, let us consider a set D of directions on the unit sphere Ξ. Fig. 2.10 depicts a typical set D occurring in practice, i.e., one which consists of a single

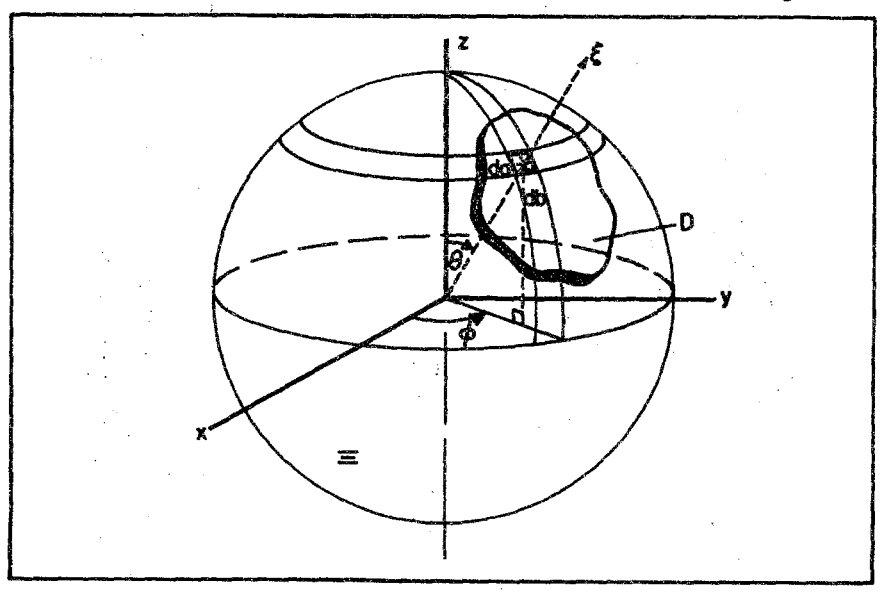


FIG. 2.10 The unit sphere of directions as the natural setting for solid angle measurements.

connected part of Ξ. It is quite natural to characterize the "amount of opening" of the set D by specifying the amount of area that D occupies on the unit sphere. Thus, we denote by "$\Omega(D)$" the number representative of the area of D on Ξ. This is the standard definition of the measure of a set of directions D, the one which we have been using informally up to this point. (For a further discussion of solid angle measure, see Note (h) to Table 3, Sec. 2.12.)

We can now go one step further and characterize $\Omega(D)$ in terms of the polar and azimuthal angles θ and ϕ (measured in radians). Clearly a small rectangular patch on Ξ about the point specified by (θ,ϕ) and of lateral extents da and db is very nearly of area da db. But since the sphere Ξ has unit radius, db = dθ and da = sin θ dϕ. Hence:

$$d\Omega(\xi) = \sin\theta \, d\theta \, d\phi \tag{9}$$

where (θ,ϕ) is associated with the direction ξ (see Sec. 2.4).

From (9) we obtain:

$$\Omega(D) = \int_D d\Omega(\xi) = \iint_D \sin\theta \, d\theta \, d\phi \tag{10}$$

It should be clear that the radius of Ξ plays no essential role in determining $\Omega(D)$. In general, if we write "$\Omega(D)$" for $A(D)/r^2$, where A(D) is the area determined by the set D on a sphere of radius r, then equation (10) results once again for $\Omega(D)$. We leave the ranges of integration in (10) undetailed, as the mode of specification of D varies widely. The number $\Omega(D)$ is customarily dimensionless. However, when dimensions of $\Omega(D)$ are needed, the system in Table 3 of Sec. 2.12 may be adopted. The standard *unit* of a solid angle is the *steradian*. It is important for a thorough understanding of solid angle, to make the distinction between the set D of directions and its measure $\Omega(D)$: D is a set of points on Ξ, $\Omega(D)$ is a number describing the size of that set.

As an example of the use of (10), consider the spherical cap D on Ξ consisting of all directions ξ with polar angles less than or equal to θ. See Fig. 2.11. Then:

$$\Omega(D) = \int_{\theta'=0}^{\theta} \int_{\phi'=0}^{2\pi} \sin\theta' \, d\theta' \, d\phi'$$

$$= 2\pi(1-\cos\theta) \quad . \qquad \text{(General } \theta\text{)} \tag{11}$$

This formula is frequently used. It is also the basis for the following well-known estimate of $\Omega(D)$ for *small* θ. In (11) let θ be small so that θ^2 is much smaller than

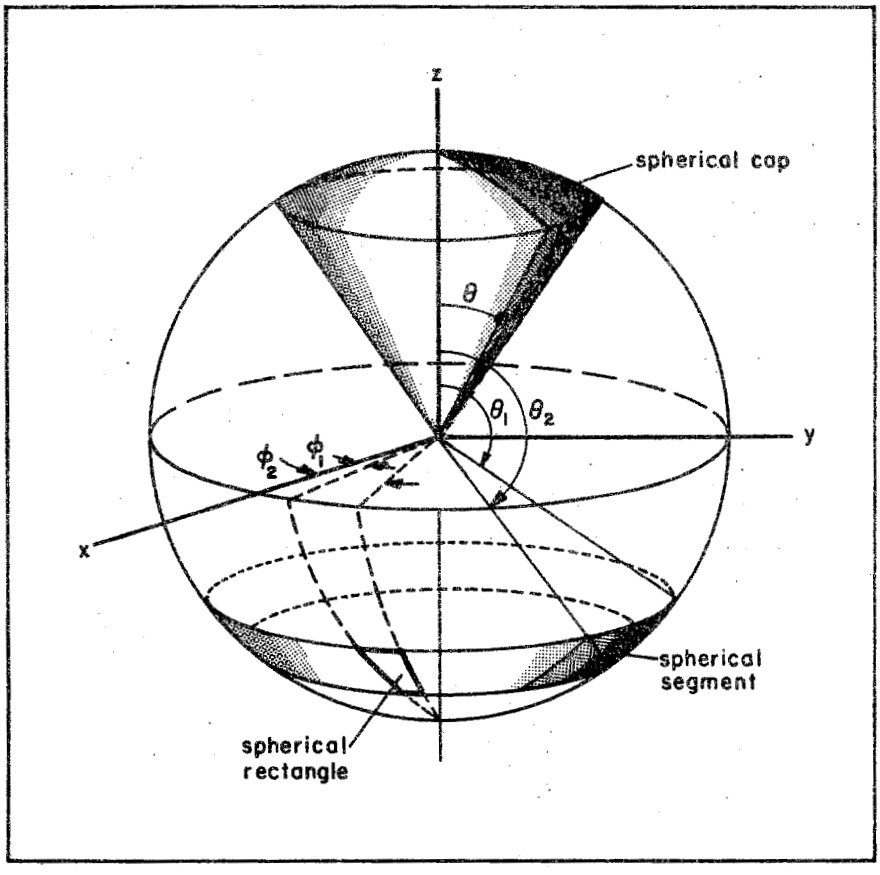


FIG. 2.11 Solid angle measures of some simple shapes

(negligible compared to) θ. Then we can approximate cos θ by $1-(\theta^2/2)$, by truncating the series expansion of cos θ at its second term. Under this assumption, (11) becomes:

$$\Omega(D) = \pi\theta^2 \qquad \text{(Small } \theta\text{)} \qquad (12)$$

From (11) we also obtain the solid angle measures for various special parts of Ξ which can be made up from spherical caps. Thus:

$$\Omega(D) = 2\pi \qquad (\theta = \pi/2, \text{ D is a hemisphere}) \qquad (13)$$

$$\Omega(D) = 4\pi \qquad (\theta = \pi, \text{ D is } \Xi) \qquad (14)$$

$$\Omega(D) = 2\pi(\cos\theta_1 - \cos\theta_2) \qquad (\theta_1 \leq \theta_2, \text{ D is a spherical segment}) \qquad (15)$$

This last formula is, incidentally, a generalization of (11), for (11) is recovered by setting $\theta_1 = 0$. Equation (15) can be used to obtain the solid angle measure of a spherical rectangle bounded by two latitude circles and two longitude circles of the unit sphere. Thus, if ϕ_1 and ϕ_2 are the bounding meridians with $\phi_1 \leq \phi_2$, then the rectangle bounded by them takes up the fraction $(\phi_2-\phi_1)/2\pi$ of the spherical segment area bounded by latitude circles at θ_1 and θ_2. Hence from (15):

$$\Omega(D) = (\phi_2-\phi_1)(\cos\theta_1-\cos\theta_2) \qquad \begin{matrix}(\theta_1 \leq \theta_2,\ \phi_1 \leq \phi_2,\ \text{(16)}\\ D \text{ is a spherical}\\ \text{rectangle)}\end{matrix}$$

Equation (16) is a further generalization of (11). The latter may be obtained by assuming $\phi_2 = \phi_1 + 2\pi$ and $\theta_1 = 0$. Of course (16) can also be obtained by direct appeal to (10).

We now return from the preceding digression on solid angles and conclude our discussion of the computation of irradiance $H(x,\xi)$, given a radiance distribution $N(x,\xi)$. It remains to cast (8) into θ,ϕ notation. Using (8), and (13) of Sec. 2.4 we have:

$$H(x,\theta,\phi) = \int_{\Xi(\theta,\phi)} N(x,\theta',\phi') \cos\varphi \sin\theta' \, d\theta' \, d\phi', \qquad (17)$$

where "$\Xi(\theta,\phi)$" is simply another name we shall use for $\Xi(\xi)$, when (θ,ϕ) is explicitly associated with ξ. Further, φ is the angle between ξ and ξ', where the latter direction is associated with (θ',ϕ'). Now, $\cos\varphi$ can be represented by means of (θ,ϕ) and (θ',ϕ') as follows. Recall first of all from Sec. 2.4 that if ξ is a unit vector, then:

$$\xi = \mathbf{i}\cos\varphi_1 + \mathbf{j}\cos\varphi_2 + \mathbf{k}\cos\varphi_3 \ ,$$

where φ_1, φ_2, and φ_3 are the angles between ξ and the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, respectively. Once again, now for ξ':

$$\xi' = \mathbf{i}\cos\varphi_1' + \mathbf{j}\cos\varphi_2' + \mathbf{k}\cos\varphi_3' \ .$$

Then by the observations in Sec. 2.4:

$$\cos\varphi = \xi\cdot\xi' = \cos\varphi_1\cos\varphi_1' + \cos\varphi_2\cos\varphi_2' + \cos\varphi_3\cos\varphi_3' \qquad (18)$$

By means of Fig. 2.4, or Fig. 2.10, we see that:

$$\left.\begin{aligned}\cos\varphi_1 &= \sin\theta\cos\phi\\ \cos\varphi_2 &= \sin\theta\sin\phi\\ \cos\varphi_3 &= \cos\theta\end{aligned}\right\} \qquad (19)$$

There are three precisely similar equations for the $\mathscr{V}_i'$ associated with ξ'. In this way (17) is cast into a well-defined analytical formula involving only θ,ϕ and θ',ϕ', in an integration over $\Xi(\theta,\phi)$. This completes the detailed unfolding of equation (8). The result is a formula often used in practice to compute $H(x,\theta,\phi)$, given $N(x,\cdot)$ at point x. We will return to illustrate equation (17), and other formulas, in Sec. 2.11. For the present we continue with the development of further properties of the concept of radiance.

Radiance from Irradiance

As a further illustration of the interconnections between the concepts of radiance and irradiance we now reverse the considerations of the preceding discussion and show that from a given irradiance distribution at a point x in an optical medium, one can compute the radiance distribution at that point. As a consequence of this fact and the results of the preceding discussion, we see that radiance and irradiance distributions share equal informational content. In addition to this theoretical consequence, there is also one of experimental import: it is possible, at least in principle, to measure irradiance distributions in natural hydrosols and aerosols and from this data deduce complete information about radiance distributions. In other words, one can in principle completely document the light fields in natural optical media solely by means of irradiance distributions.

We begin the illustration with the simplest possible case: we are given that the irradiance distribution $H(x,\cdot)$ at point x is generated by a radiance distribution $N(x,\cdot)$ which is of uniform radiance N over a small conical set D' of directions of solid angle magnitude $\Omega(D')$ with central direction ξ' and with $N(x,\xi')$ zero for all other directions. It is required to find N. Now from (8) we have very closely:

$$H(x,\xi) = N\xi\cdot\xi'\Omega(D') \quad ,$$

whence:

$$N = H(x,\xi)/\xi\cdot\xi'\Omega(D') \quad ,$$

where ξ is some specifically chosen vector such that $\xi\cdot\xi'>0$.

Suppose next that the given irradiance distribution is generated by a radiance distribution which is of uniform magnitude N_1 over a narrow set D_1' of directions with central direction ξ_1' and of uniform magnitude N_2 over a narrow set D_2' of directions with central direction ξ_2' and such that D_1' and D_2' are disjoint and $N(x,\xi')$ is zero for all other directions. From (8) we have now:

$$H(x,\xi) = N_1\xi\cdot\xi_1'\Omega(D_1')\chi(\xi,\xi_1') + N_2\xi\cdot\xi_2'\Omega(D_2')\chi(\xi,\xi_2') \tag{20}$$

where χ is a function with the property that $\chi(\xi,\xi') = 1$ or 0 according as ξ' is or is not in $\Xi(\xi)$, respectively, and where ξ is any direction in Ξ.

Now we may choose ξ at will from a large collection of possibilities, and use the given values $H(x,\xi)$ to try to determine the two radiances N_1 and N_2. Clearly we must generally choose two directions ξ_1 and ξ_2 in order to determine N_1 and N_2. This is readily seen if we write:

$$\text{"}C_{ji}\text{"} \quad \text{for} \quad \xi_i \cdot \xi_j' \Omega(D_j') \chi(\xi_i, \xi_j')$$

for each i = 1,2, and j = 1,2, and furthermore, if we write:

$$\text{"}H_i\text{"} \quad \text{for} \quad H(x,\xi_i)$$

for each i = 1,2. Then equation (20) yields, for $\xi=\xi_1$ and $\xi=\xi_2$ the two equations:

$$H_1 = N_1 C_{11} + N_2 C_{21}$$

$$H_2 = N_1 C_{12} + N_2 C_{22}$$

If we write:

$$\text{"C"} \quad \text{for} \quad \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} ,$$

then the preceding set of equations can be written:

$$(H_1, H_2) = (N_1, N_2)C \quad , \tag{21}$$

or, more succinctly, as:

$$H = NC \quad . \tag{22}$$

We have written:

$$\text{"H"} \quad \text{for} \quad (H_1, H_2)$$

and

$$\text{"N"} \quad \text{for} \quad (N_1, N_2) \quad .$$

The solution of Equation (22) is that N for which

$$N_1 = \begin{bmatrix} H_1 & C_{21} \\ H_2 & C_{22} \end{bmatrix} \cdot \Delta^{-1}$$

$$N_2 = \begin{bmatrix} C_{11} & H_1 \\ C_{12} & H_2 \end{bmatrix} \cdot \Delta^{-1} \tag{23}$$

where we have written:

$$\text{"}\Delta\text{"} \quad \text{for} \quad C_{11}C_{22} - C_{12}C_{21} \quad ,$$

and where the four C_{ij} are such that $\Delta \neq 0$. This solution may be put into the form:

$$N = HC^{-1} \quad , \tag{24}$$

where

$$C^{-1} = \begin{bmatrix} C_{22} & -C_{12} \\ -C_{21} & C_{11} \end{bmatrix} \cdot \Delta^{-1} \tag{25}$$

The pattern is now clear as to the means of obtaining a radiance distribution from the generated irradiance distribution. For, generalizing the two simple cases just considered, we now suppose that a given irradiance distribution at point x is generated by a radiance distribution at x which is uniform and of magnitude N_i over each of n narrow sets D_i' of directions such that D_i' and D_j' are disjoint and with central direction ξ_i' for each D_i', i= 1,...,n. Hence the set $\{D_1', D_2', \ldots, D_n'\}$ of subsets of Ξ is an arbitrary partition of Ξ into narrow bundles of directions. From (8) we have:

$$H(x,\xi) = \sum_{i=1}^{n} N_i \xi \cdot \xi_i' \Omega(D_i') \chi(\xi, \xi_i') \quad , \tag{26}$$

where ξ is any direction and $\chi(\xi,\xi_i')$ has the same meaning as before for the case n = 2.

We now choose n directions ξ_i, i= 1,2,...,n and write: "C_{ji}" and H_i exactly as before, but now with i and j ranging over the general finite set {1,2,..., n} of integers. With this notational convention (26) becomes:

$$H_j = \sum_{i=1}^{n} N_i C_{ij} \quad . \tag{27}$$

Writing:

$$\text{"H"} \quad \text{for} \quad (H_1, \ldots, H_n)$$

$$\text{"N"} \quad \text{for} \quad (N_1, \ldots, N_n)$$

$$\text{"C"} \quad \text{for} \quad \begin{bmatrix} C_{11} & C_{12} & \ldots & C_{1n} \\ C_{21} & C_{22} & \ldots & C_{2n} \\ \vdots & & & \\ C_{n1} & C_{n2} & \ldots & C_{nn} \end{bmatrix} \quad ,$$

we can then cast (27) into matrix form:

$$\boxed{H = NC} \tag{28}$$

The solution of (28) is given by:

$$\boxed{N = HC^{-1}} \tag{29}$$

where C^{-1}, the inverse matrix of C, generally exists upon suitable choice of the ξ_i.

The proof of the existence of the general continuous counterpart to C^{-1} is given in Example 15 of Sec. 2.11. There the full equivalence of $H(x,\cdot)$ and $N(x,\cdot)$ is established. The rigorous proof of the equivalence requires relatively advanced concepts and for that reason is deferred to Sec. 2.11. However, the present discussion has been designed so that the practical details involved in the determination of N by H require no tools beyond those of the elementary theory of algebraic equations.

As a result of the preceding discussion leading to (29), we can view in a new light the observation that "radiance is the most basic of radiometric concepts". The radiance concept is most basic in the sense that from it all other radiometric quantities can be *most conveniently* derived; it is not "most basic" in the sense that there is only a one-way computational path from it to every other radiometric quantity. This brings up the interesting question of: just which of the radiometric quantities discussed so far have informational content equivalent to radiance? and: just what, in the last analysis, characterizes a radiometric concept which has equivalent informational content to radiance? These questions will be briefly considered in Example 15 of Sec. 2.11.

Field Radiance vs Surface Radiance

There is a distinction that can be made in practice between two types of radiance, a distinction which is analogous to that made in Sec. 2.4 between irradiance and radiant emittance. This distinction is depicted with the help of Fig. 2.12 which shows radiant flux across an hypothetical surface S in the indicated direction and within a narrow conical set D of directions around a direction ξ normal to S.

Now, corresponding to the conceptual distinction established between W(S,D) and H(S,D) in (17) and (18) of Sec. 2.4, we can write:

$$\text{"}N^{+}(S,D)\text{"} \quad \text{for} \quad W(S,D)/\Omega(D) \tag{30}$$

and

$$\text{"}N^{-}(S,D)\text{"} \quad \text{for} \quad H(S,D)/\Omega(D) \ . \tag{31}$$

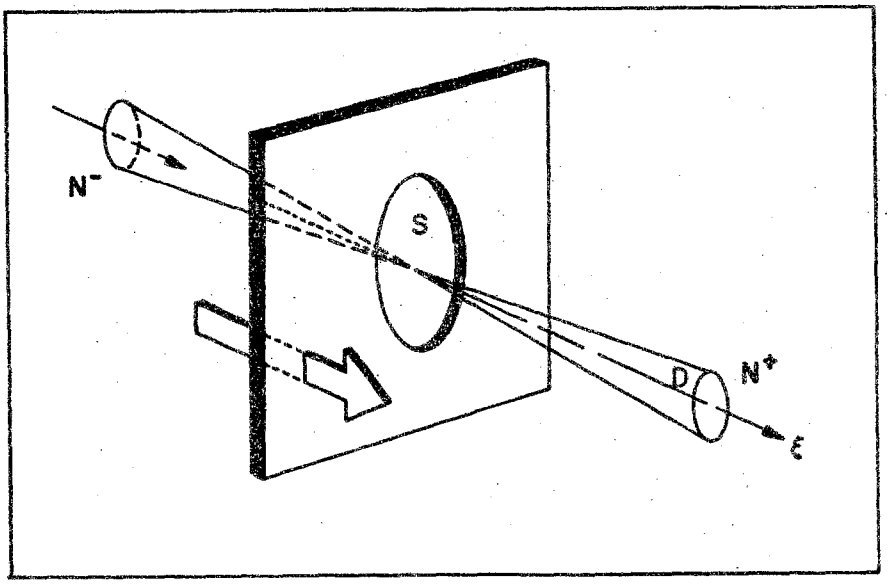


FIG. 2.12 Conceptual distinction between field radiance N^- and surface radiance N^+.

We call $N^-(S,D)$ the (empirical) *field radiance*, and $N^+(S,D)$ the (empirical) *surface radiance* (or *specific radiance*, or *specific intensity*). It is quite clear that, in the general context of Fig. 2.12:

$$N^+(S,D) = N^-(S,D) \tag{32}$$

Despite the numerical equality, the conceptual distinction between field and surface radiance is useful to maintain. Indeed, some need for a conceptual distinction inevitably forces itself on the attention of careful students of applied radiative transfer theory where on the one hand emitting surfaces, real or hypothetical, are characterized most naturally by surface radiance, and where measurements obviously result in field radiances. The term "surface" in "surface radiance" is a vestige of the days when surface radiance was associated with the radiant emittance of real surfaces enclosing sources of radiant energy. The present interpretation of "surface", however, includes the possibility of hypothetical surfaces anywhere in an optical medium. The term "field" in "field radiance" denotes the sense of "field of view". In practice, whenever possible, one of these two interpretations of radiance is usually fixed and agreed upon throughout a given discussion. Thus, we can omit the "+" (or "-") superscript from "N" when the type of radiance is understood.

2.6 An Invariance Property of Radiance

In this section we shall discuss a property of the concept of radiance which is of central importance from the point of view of radiative transfer phenomena. This is the so-called n^2-law of radiance which states that the quotient N/n^2 does not change along a path of sight through a transparent medium in which there is a generally variable index of refraction n. The importance of this law rests in the base line it establishes for comparison of the behavior of N/n^2 along lines of sight in *non*-transparent media, i.e., media that scatter and absorb radiant energy such as the atmosphere and the seas, and other natural optical media. The law also indicates a measure of success in our attempt to simulate the sensation of brightness by means of a simply defined radiometric concept. For it is a matter of daily experience that as one approaches or recedes from an object along a line of sight through a very clear homogeneous stretch of atmosphere (so that n is constant), the "brightness" of the object does not appear to change. For example, the brightness of a small part of a desk blotter does not change as we move away from it in a room, keeping attention constantly directed toward the patch. Of course, the *total flux* entering the eye and originating from the patch falls off rapidly with distance (very nearly as the square of the distance, as we shall eventually show); however, the *brightness* of the patch does not change with the observer's distance. This phenomenon is reproduced in the special form of the n^2-law where n is constant over the path of sight. We now show how the n^2-law for radiance follows from the definition of radiance. We shall divide the discussion into two main parts. The first part considers the important case in which n is constant along the path of sight. The second part considers the general case of a variable index of refraction.

The Radiance-Invariance Law

We begin the derivation of the n^2-law for the special case where n is constant along a line of sight through a transparent optical medium. This special case is of sufficient importance to be given a special name, the *radiance-invariance law*. We shall prove the radiance invariance law twice: first in as simple a way as possible so as to reveal the geometrical essence of the law; then the derivation will be repeated in slightly more detail, filling in steps and giving more explanations on the way.

The setting for the simple derivation is shown in Fig. 2.13. Two holes S and S' of arbitrary shape and about the size of collecting surfaces used in radiant flux meters are cut out of two large pieces of opaque cardboard. The pieces are then mounted so that they are parallel and separated a distance r which is large compared with the linear dimensions of the holes. Light is then directed through S which flows along straight lines in the transparent space between the cardboards and then on through S'. The holes are arranged so

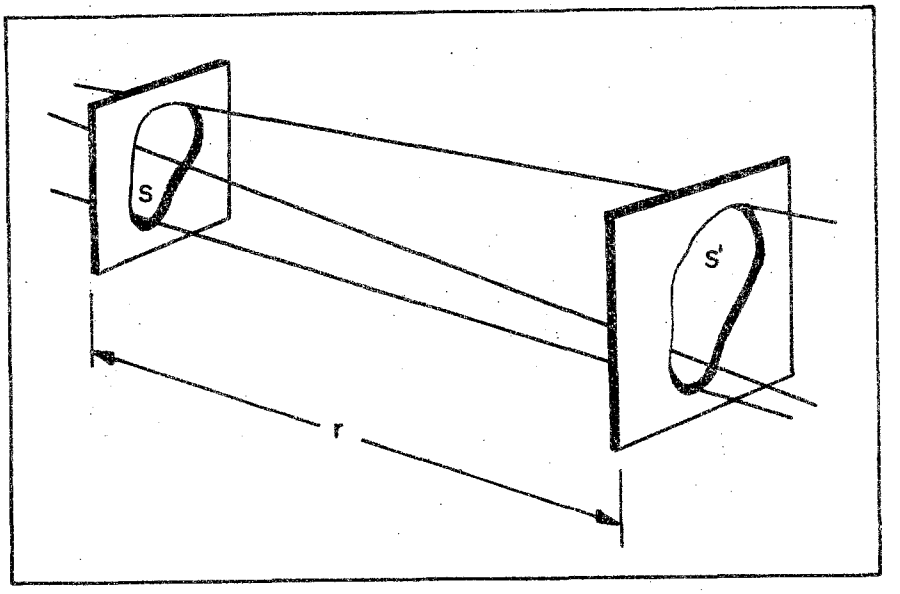


FIG. 2.13 Illustrating the invariance of the radiance of a narrow bundle of light rays in a vacuum.

that for the most part, the lines of flux through both openings are nearly perpendicular to the planes of the holes. The observation is now made that the amount P of radiant flux across S, associated with the common bundle of lines of flux through S and S', is the same as that across S'. Thus the same number of lines of flux go through both S and S'. With this in mind we consider the number:

$$\frac{P}{\frac{A(S)A(S')}{r^2}}$$

in two ways. First as:

$$\frac{P(S,D)}{A(S)\frac{A(S')}{r^2}} ,$$

and then as:

$$\frac{P(S',D')}{\frac{A(S)}{r^2}A(S')} .$$

In the first case we observe that $A(S')/r^2$ is essentially the solid angle $\Omega(D)$ subtended by S' at each point of S. In the

second case $A(S)/r^2$ is the solid angle $\Omega(D')$ subtended by S at each point of S'. "P(S,D)" and "P(S',D')" both denote the common radiant flux P, but now in an obviously suggestive way by recalling the meanings of S, S', D, D'. Therefore we have:

$$N(S,D) = N(S',D') \tag{1}$$

This is the empirical form of the *radiance-invariance law*. The form of the law is "empirical" because it is couched in terms of empirical radiances--radiances directly measurable by real radiance meters.

A somewhat more detailed derivation of the radiance-invariance law will now be given. Part (b) of Fig. 2.14 depicts a radiance meter G directed at a surface S at the end

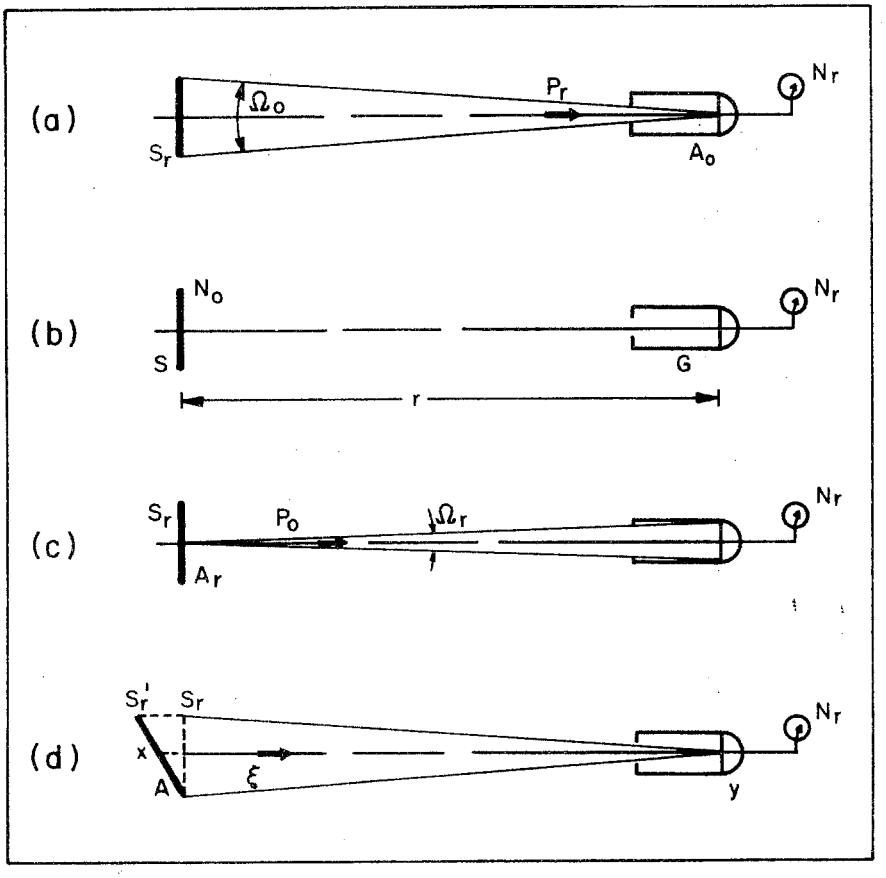


FIG. 2.14 A more detailed, and operational, study of the radiance-invariance law

of a clear path of sight of length r. The surface S is normal to the line of sight and has a uniform surface radiance N_O over its extent in the direction of G. The meter G has its field of view completely filled by S. The resultant radiance reading is N_r. We will show that, under these conditions we have, for every r, $N_O = N_r$. The basic idea of the proof is to examine the same diagram (b) from two distinct points of view. These points of view are schematically depicted in parts (a) and (c) of Fig. 2.14. We consider part (a) first. Here the radiance meter's reading N_r is seen to be the quotient $P_r/A_O\Omega_O$, where P_r is the radiant flux originating on S and incident on the collecting surface of area A_O, and which has funneled through the solid angle of magnitude Ω_O, defined as shown. On the other hand, part (c) views this flux as an amount P_O sent to area A_O in G and as emitted from those points of S_r within G's field of view. The emitting surface S_r comprising G's field of view is of variable magnitude A_r, and the emitted flux from each point of S_r is within the bundle of directions of solid angle defined as shown. Hence, N_O is the quotient $P_O/A_r\Omega_r$. The definitions of P_O and P_r imply at once (as in the previous proof) that:

$$P_O = P_r \quad .$$

Further, we have the geometric observation that:

$$\frac{A_O A_r}{r^2} = \Omega_r A_r = A_O \Omega_O \quad .$$

On the basis of these two facts, we see that, by virtue of the defining equations:

$$N_r = P_r/A_O\Omega_O$$

$$N_O = P_O/A_r\Omega_r$$

we have:

$$\boxed{N_O = N_r} \qquad (2)$$

Observe how we have implicitly used the distinction between field and surface radiance and the connection (32) of Sec. 2.5 in order to interpret N_O operationally at surface S, which then is N_r for $r = 0$.

The Operational Meaning of Surface Radiance

One final matter must be resolved before the radiance-invariance law is fully established. This is the matter of assigning a meaning to the surface radiance of a surface at a point x in a direction other than the normal direction to S at x. Observe that this problem does not arise with field radiance, since field radiance is defined by the convention of

using the fixed collecting area of a radiance meter, which is assembled so that it is perpendicular to the axis of the meter. In the case of surfaces such as a portion of the earth's surface, a desk top or a wall, or a given cloud boundary, how shall we assign a surface radiance to radiant flux leaving a point on such surfaces in directions other than the perpendicular direction to the surface at that point? The path to the answer is guided by the manner in which such surface-radiance information is first of all to be interpreted and secondly how it is to be used. In the first case we really have very little choice as to the manner of interpretation of the radiance information. We have already committed ourselves to work solely with operational concepts: measurable fluxes, areas and solid angles. Hence, if we heard someone say: "The surface radiance of flux of wavelength 550 mμ at point x on wall A is 2 watts/(m^2 × steradian) in every direction 30° from the normal to A at x", our first impulse, after this data has been mentally assimilated, would be to attempt a verification by directing a radiance meter toward x on A so that the axis of the meter makes an angle of 30° with the normal to A at x. If we were challenged to defend such a procedure, we would cite the argument leading to the radiance-invariance law above. However, if the challenger were particularly tenacious, he would point out that the argument establishing the law holds only for directions of sight *normal* to A at x. At this juncture we must concede that he is right.

The preceding objection to our justification for assigning an operational meaning to oblique surface radiance is logically unassailable. However, we have one more matter to consider which will add strength to the justification. We now consider the second aspect of the question posed above, namely: how is the information of oblique surface radiance to be used? The answer, based on considerable practical and theoretical experience, is that such oblique surface radiance information is to be used to calculate irradiances,scalar, vector, or of the ordinary variety, at points optically accessible to the surface which emits the surface radiance. Or again, the surface radiance information will be used to obtain path functions, and various attenuation functions used in hydrologic optics or meteorologic optics and these determinations will be made at points optically accessible to the surface. The pertinent fact that emerges as these uses of the surface radiance information are paraded before the mind's eye is the following: *without exception, the information used can always be in the form of field radiance values of the radiometric field in the direction of point* x *on surface* A. In short, surface radiance *per se* while of great conceptual and theoretical worth, is never really used in actual practical calculations--only the directly observable field radiance values are used in such calculations. We are therefore motivated to *assign an operational value of surface radiance to a surface* A *at* x *in the general outward direction* ξ *by means of the corresponding field radiance reading* $N(y,\xi)$ *obtained when the radiance meter is at some point* y *and is directed at* x *so that the unit inward normal to the collecting surface of the meter is* ξ. *The point* y *is to be anywhere along a clear path of sight from* x *in the direction* ξ *so that the radiance-invariance law holds.*

For conceptual definiteness in the preceding convention, one can imagine an hypothetical surface S_Y normal to ξ as in (d) of Fig. 2.14, which is assigned the surface radiance $N(x,\xi)$, (as N_0 in the derivation of the radiance-invariance law). Now while the radiance-invariance law allows us to conclude that the radiance reading will remain unchanged as distance r varies from 0 at S_Y to larger values from S_Y, there still is a conceptual gap that must be filled between the surface radiance of S_Y and the surface radiance of S_Y', the projection of S_Y onto the oblique surface under consideration (as, e.g., A above). And this gap, we have agreed, is to be filled by means of the preceding convention. Equation (2) and every result deduced from it, shall henceforth be interpreted with this convention implicitly understood.

The n^2-Law for Radiance

The intuitive basis for the n^2-law, to which we now turn, becomes clear upon consideration of Fig. 2.13. This figure shows a narrow bundle of lines of flux coursing through empty space. The two holes in the cardboard arrangement used above were so much inessential material scaffolding which can be removed now that the idea of the derivation has been explained. What is left after this is done is the concept of a narrow bundle of lines of flux coursing through space in such a way that at each section the product $A\Omega$ of the normal cross sectional area A of the bundle and solid angle Ω of the bundle is a fixed quantity. This invariance of $A\Omega$ is a purely geometric concept. Physical considerations enter subsequently at the point where we assert the invariance of the radiant flux through a variable section of the bundle of lines of flux. By combining these physical and geometric considerations, the desired radiance-invariance law is obtained for a light beam in a vacuum. We now inquire: how are these physical and geometrical considerations to be modified in the case of a light beam coursing through matter such as air or water? The physical considerations governing the radiant flux content of the beam must take into account the scattering and absorption phenomena all along the extent of the beam. These phenomena affect the radiant flux content of the beam in complex and subtle ways. The full study of these effects is reserved for the theory of Part Two of this work. We shall limit our present inquiry to sets of adjoining transparent media. Any alterations of the radiant flux content of the beam are then limited to the interfaces of these media. If we now repeat our query above for the case of contiguous transparent media which are distinguished from each other only by their various indices of refraction, then the answer to the query is given in the form of the n^2-law for radiance. The derivation of this law for the simplest case will now be given.

Figure 2.15 depicts a beam of radiant flux lines incident on the interface Y between two transparent optical media X_1 and X_2. Let us agree that the central axis of the beam is normally incident on the interface, that it arrives from medium X_1, and that the beam passes on through the interface Y and enters medium X_2. For example, X_1 may be a part of the

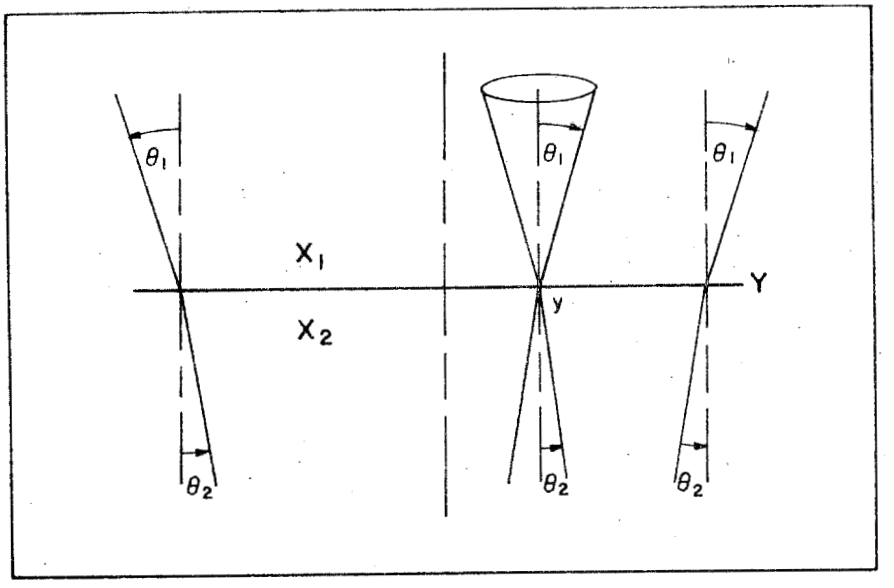


FIG. 2.15 When a bundle of light rays is suddenly squeezed into a narrower bundle--without changing its flux content--the radiance of the bundle increases proportionately. This essentially is what happens, e.g., at the air-water surface of natural hydrosols (flux losses to one side).

atmosphere, and X_2 a part of the hydrosphere, so that Y is the air-water interface. In general, X_1 has some index of refraction n_1 and X_2 an index of refraction n_2 in the immediate neighborhood of the interface. Our current goal is to relate the radiance of the beam in X_1 to the radiance of the beam in X_2 in the immediate vicinity of Y. Before going into the details of the derivation it is instructive to anticipate the result intuitively. We ask: which of the three main quantities P, A or Ω of the definition of radiance N will change from one side of Y to the other? Clearly, the radiant flux content P will be affected to some extent by reflection of some of the lines of flux back into X_1. The area A of the beam will remain essentially unchanged arbitrarily close to each side of Y. Finally, the solid angle magnitude Ω of the beam will change from one side to another because of refraction of the lines of flux transmitted across the surface Y. For the moment we ignore the change of P, this change yielding a relatively small change in N and one which will eventually vanish as the derivation proceeds to its final stages. Hence the principal change in N that is wrought on the radiant flux is traceable to the abrupt change in the Ω of the beam as it crosses Y. For example, if X_1 and X_2 are respectively, air and water, then as a glance at Fig. 2.15 would show, Ω_1, the solid angle magnitude of the beam in air, is greater than Ω_2, the solid angle magnitude of the transmitted beam in water. Since P and A are essentially unchanged during the passage of

the flux across Y, the radiance N_2 of the beam in water exceeds its radiance in air. This yields the general observation that, aside from reflection losses, *as a narrow beam of radiant flux goes from a medium of smaller to greater index of refraction, its radiance changes from a smaller to a greater amount.* The reverse change occurs for a reverse traversal. The n^2-law gives a precise quantitative description of this observation.

Now to the particulars. At each point y of the interface Y within the beam, there is a cone of incident lines of flux on y. This cone has a small half angle θ_1. The associated solid angle magnitude Ω_1 of the incident cone is, by (12) of Sec. 2.5, $\pi\theta_1^2$. Similarly $\pi\theta_2^2$ is the solid angle magnitude Ω_2 of the beam of transmitted lines of flux. Snell's law ((5) of Sec. 12.1) gives the connection between θ_1 and θ_2 :

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad .$$

For small θ_1 and θ_2, this becomes, very nearly:

$$n_1\theta_1 = n_2\theta_2 \quad .$$

Squaring each side, we obtain:

$$n_1^2\theta_1^2 = n_2^2\theta_2^2 \quad .$$

Multiplying each side by π we obtain:

$$n_1^2(\pi\theta_1^2) = n_2^2(\pi\theta_2^2)$$

That is:

$$n_1^2\Omega_1 = n_2^2\Omega_2 \quad . \tag{3}$$

Now starting with the equal irradiances H_1 and H_2 of the beam on each side of Y, we have, for the reasons discussed above:

$$\frac{P_1}{A_1} = H_1 = H_2 = \frac{P_2}{A_2} \quad .$$

We now use (3) with this to get:

$$\frac{P_1}{A_1(n_1^2\Omega_1)} = \frac{P_2}{A_2(n_2^2\Omega_2)}$$

That is:

$$\boxed{\frac{N_1}{n_1^2} = \frac{N_2}{n_2^2}} \tag{4}$$

This is the desired form of the n^2-*law for radiance*.

It is now an easy matter to gradually generalize (4) to the following more general settings: (a) passage through an arbitrary finite number of transparent contiguous media and going on to the limit of continuously varying n; (b) oblique rather than normal incidence of the beam on Y; (c) inclusion of a transmittance factor to allow for scattering and absorption losses from P. Generalizations (a) and (b) result in no change in the form of (4). Generalization (c) results in a multiplicative factor T included on the left side of (4). This factor will be considered at great length in Sec. 3.10 in the discussions of beam transmittance. Specific suggestions for these generalizations are given in Sec. 12 of Ref. [251]. See also [98]. *Henceforth, whenever radiances are related within media of differing indices of refraction, it will be understood that* N/n^2 *rather than* N *will be used, even though "N" only appears in the equations.*

2.7 Scalar Irradiance, Radiant Energy, and Related Concepts

The radiometric concepts studied in this section are those of scalar irradiance, radiant energy, and related radiometric concepts. The first of these concepts is designed to quantitatively describe the volume density of radiant energy in a way which is amenable to operational methods of determination. In addition to the notion of scalar irradiance, we shall develop in this section several closely related notions which together with scalar irradiance comprise a set of useful measures of the volume density of radiant energy. The first of these is radiant density.

Radiant Density

The notion of radiant density is one of several concepts designed to give a measure of the radiant energy per unit volume at a point. Consider a steady beam of radiant flux normally incident on surface S at point x at time t, as shown in Fig. 2.16. Let the field radiance of the beam at this instant be N, its cross sectional area be A, and its solid angle be Ω. The amount of radiant flux incident on S at time t is then $NA\Omega$. An instant t later, the flux of the beam will have moved on a distance $r = vt$, and the flux will have swept out a cylindrical volume of magnitude $V = Avt$. During this time the beam has been steadily pouring an amount of radiant energy into the volume at the rate of $NA\Omega$ watts. Hence the radiant energy content of the beam is $NA\Omega t$, and its average content per unit volume is $NA\Omega t/Avt = N\Omega/v$.

Suppose that point x were simultaneously irradiated at time t by an arbitrary finite number of narrow beams of radiance N_i, $i = 1, \ldots, n$, and corresponding solid angles Ω_i. Then the radiant energy u(x,t) per unit volume at x is given at time t, by means of the D-additivity of Φ (equation (7), Sec. 2.3):

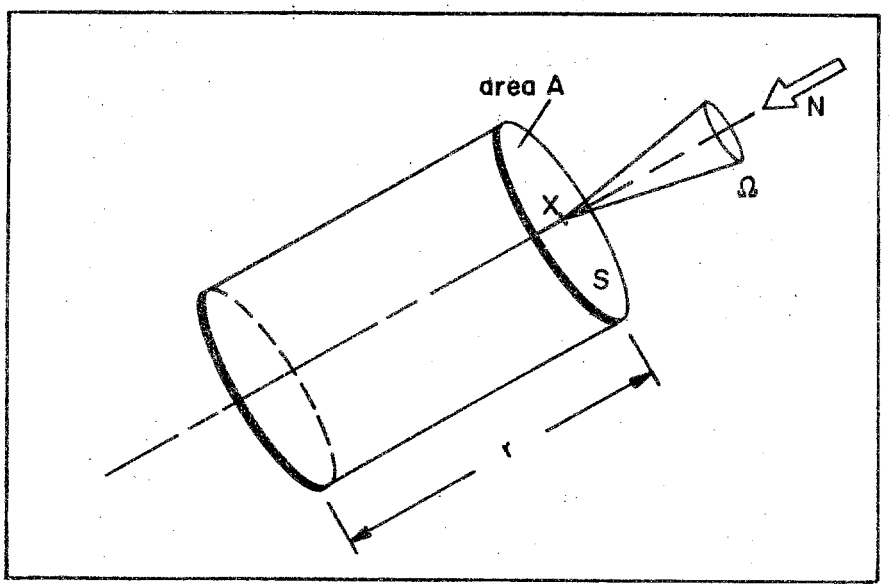


FIG. 2.16 Setting up the connection between radiance and radiant density

$$u(x,t) = \frac{1}{v}\sum_{i=1}^{n} N_i \Omega_i \tag{1}$$

The transition to the continuous case is immediate. Toward this end, let us continue to write "u(x,t)" for the *radiant density*, i.e., we shall also write:

$$\text{"u(x,t)"} \quad \text{for} \quad \frac{1}{v(x,t)} \int_{\Xi} N(x,\xi,t)\, d\Omega(\xi) \ . \tag{2}$$

The units of u(x,t) are *joules/m*^{3}. We may use either the field or surface interpretation of radiance in this definition.

Scalar Irradiance

Let us go on to write:

$$\text{"h(x,t)"} \quad \text{for} \quad \int_{\Xi} N(x,\xi,t)\, d\Omega(\xi) \ . \tag{3}$$

$h(x,t)$ is the *scalar irradiance* at x and time t. The field radiance interpretation of N is most often used in (3), and this interpretation will be in force unless specifically noted otherwise. The reason for singling out $h(x,t)$ for special consideration will be made clear in a moment. For the present it suffices to note that in general:

$$u(x,t)v(x,t) = h(x,t) \tag{4}$$

By virtue of (3) it follows that in this equation the field interpretation of $u(x,t)$ is to be understood, and that while the units of $u(x,t)$ are joules/m^3, those of $h(x,t)$ are watts/m^2. Hence, the term "irradiance" in the name "scalar irradiance" is appropriate. The reason for the modifier "scalar" will also become clear subsequently after vector irradiance has been defined in (2) of Sec. 2.8. A generalization of (3) is obtained by replacing Ξ by a subset D of Ξ. In that case we would write:

$$\text{"}h(x,D,t)\text{"} \quad \text{for} \quad \int_D N(x,\xi,t)\, d\Omega(\xi) \quad .$$

The radiant density associated with $h(x,D,t)$ is $u(x,D,t)$ and (4) holds for these two quantities.

Spherical Irradiance

We shall now show why scalar irradiance is singled out as an alternate (and an actually preferred) description of the radiant density at a point in a radiant flux field. Consider the light field at a point x in a natural optical medium at time t. Let $N(x,\cdot)$ be the radiance distribution at x. Now imagine a small spherical collecting surface S of radius r in the field so that its center is at x. We then ask: *what is the average amount of radiant flux incident per unit area over* S?

To answer this question it is useful to conceptually decompose the great number of radiant flux streams at x into a discrete set of flows. Two such flows are shown in Fig. 2.17. The lines of flux of one of these flows along the direction ξ_i have been fitted with little direction cones of solid angle magnitude Ω_i. Suppose the radiance at x in the direction ξ_i is N_i. Then the irradiance at x on a plane normal to ξ_i is $N_i\Omega_i$. If the sphere is small, say the size of a ping pong ball, then for most natural light fields in the air and sea, N_i will not vary in the region of space taken up by the volume of the sphere. From this we see that we can treat the radiance function N as a constant with respect to location in the vicinity of the sphere and of value N_i for the direction ξ_i. It follows that the amount of radiant flux incident on the sphere contributed by the stream of flux in the direction ξ_i is $(n_i\Omega_i)\pi r^2$. This estimate is based on the

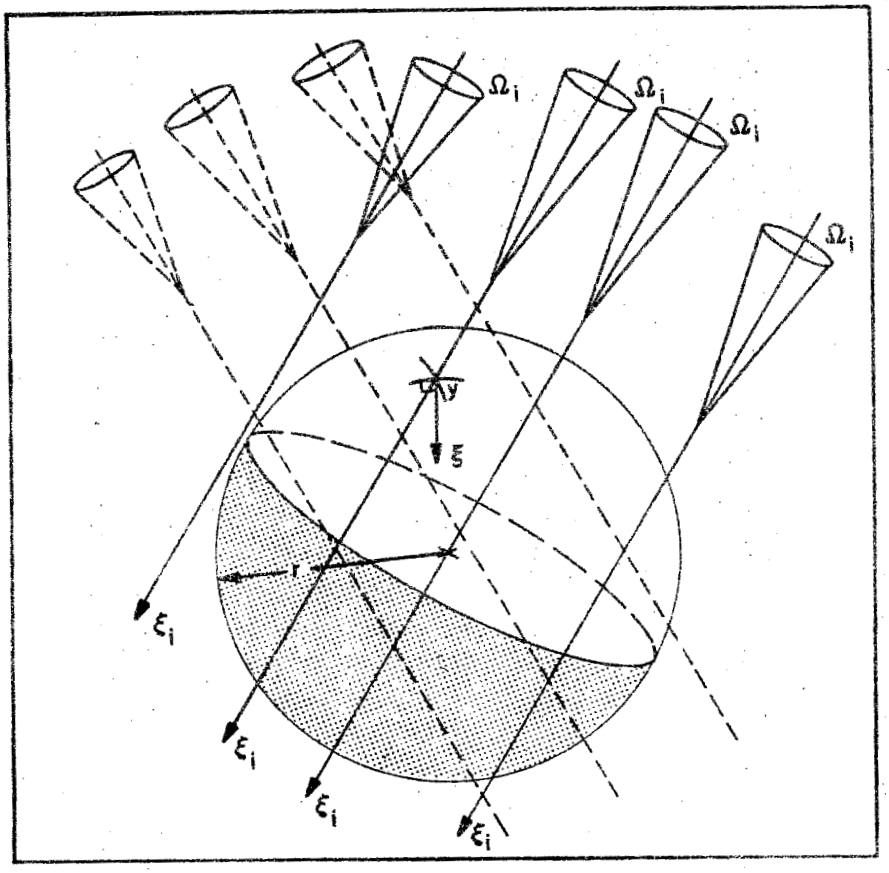


FIG. 2.17 Computing the radiant flux intercepted by a spherical collector in a general light field

assumption that the amount of flux of a narrow beam intercepted by a curved hemispherical surface is the same as the amount intercepted by the great-circle area associated with the hemisphere. The assumption is rigorously defensible for transparent media using the concepts of vector analysis and Stokes Theorem. For the present the reader's intuition will readily allow this assumption to stand even for the case of turbid media as long as r is kept very small. The "line of flux" interpretation will help the intuition considerably in this matter.

The main task in answering the above question has now been dispatched. It remains only to add up all the contributions by the various beams of flux, using as justification Equation (7) of Sec. 2.3. The result is:

$$\pi r^2 \sum_{i=1}^{n} N_i \Omega_i$$

The average radiant flux per unit area of the sphere S is then obtained by dividing this quantity by $4\pi r^2$. Let us designate this average by writing:

$$\text{"}h_{4\pi}(x,t)\text{"} \quad \text{for} \quad \frac{1}{4} \sum_{i=1}^{n} N_i \Omega_i \tag{5}$$

and agree to call it *spherical irradiance*. We shall retain this terminology and notation for the continuous formulation. That is, we shall write:

$$\text{"}h_{4\pi}(x,t)\text{"} \quad \text{for} \quad \frac{1}{4} h(x,t) \tag{6}$$

Definition (6) is the basis for an operational determination of scalar irradiance using a spherical collecting surface S. For the average radiant flux per unit area on S is readily measurable and this amount differs multiplicatively from u(x,t) by a fixed numerical factor. Hence, by only slight changes in optical design, the same photoelectric devices used to determine H and N can be directed to obtain scalar irradiance h. Therefore it is spherical irradiance or scalar irradiance which is directly measurable by photoelectric devices. The concept of radiant density u(x,t) is by way of contrast a theoretical concept related to the empirically-based concept h(x,t) by means of (4).

Hemispherical Irradiance

One of the most useful mathematical models of light fields in natural waters is the exact two-flow model to be considered in detail in Chapter 8. A radiometric concept which arises in that theory, and one which also has been found of intrinsic interest to experimenters, is the concept of hemispherical scalar irradiance. We now discuss this concept.

Figure 2.18 (a) depicts a small spherical collecting surface S with center x which is exposed to flux from only one hemisphere of Ξ. Let $N(x,\cdot)$ be the radiance distribution at x. Let us say that light is incident on the sphere in the direction of $\Xi(\xi)$. We ask: what is the average amount of radiant flux incident per unit area over S? Clearly every point of S is in principle exposed to the light field over $\Xi(\xi)$. Fig. 2.18 (b) shows how an obliquely incident beam with a direction in $\Xi(\xi)$ can come close to illuminating the "north pole" of the little spherical surface. If we divide up $\Xi(\xi)$ into pieces analogously to the manner used in deriving the expressions above for spherical and scalar irradiance, then it becomes clear that the integral of $N(x,\cdot)$ over $\Xi(\xi)$ yields the appropriate scalar or spherical irradiance component.

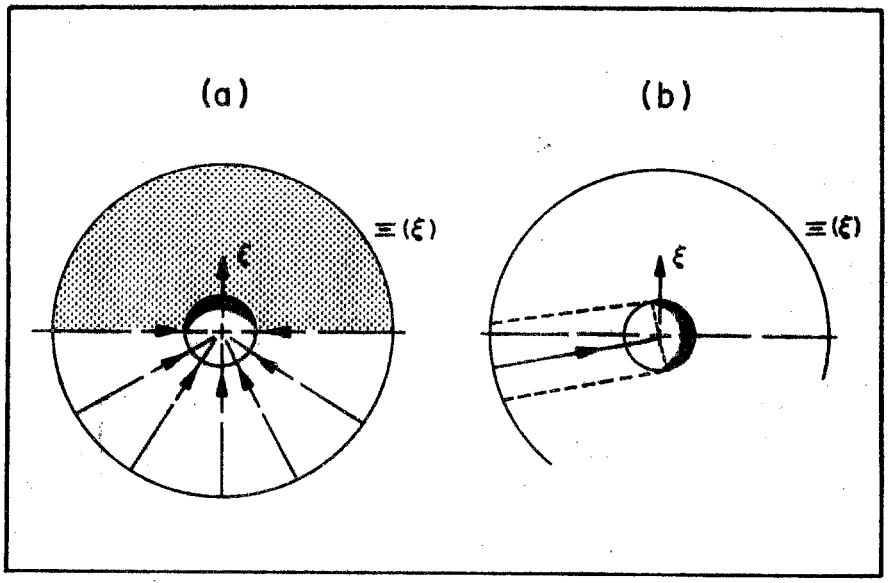


FIG. 2.18 Details for a shielded spherical radiant flux collector

Thus, using field radiance let us write:

$$\text{"}h(x,\xi,t)\text{"} \quad \text{for} \quad \int_{\Xi(\xi)} N(x,\xi',t)\, d\Omega(\xi') \tag{7}$$

and analogously, we write:

$$\text{"}h_{4\pi}(x,\xi,t)\text{"} \quad \text{for} \quad \frac{1}{4}\int_{\Xi(\xi)} N(x,\xi',t)\, d\Omega(\xi') \tag{8}$$

We call $h_{4\pi}(x,\xi,t)$ the *hemispherical irradiance* at x, over the hemisphere $\Xi(\xi)$, at time t. Further, $h(x,\xi,t)$ is the associated *hemispherical scalar irradiance*. It is clearly a special case of $h(x,D,t)$ defined after (3) above. Methods of measuring hemispherical irradiances will be discussed in Chapter 13. It follows immediately from (3) and (7) that:

$$h(x,t) = h(x,\xi,t) + h(x,-\xi,t) \quad . \tag{9}$$

An analogous connection to that displayed in (9) also holds between $h_{4\pi}(x,\pm\xi,t)$ and $h_{4\pi}(x,t)$. The introduction of $h_{4\pi}(x,\xi,t)$ into the family of radiometric concepts is motivated exactly for the empirical reasons that motivated the introduction of its full spherical companion $h_{4\pi}(x,t)$.

When we are working in stratified light fields (Sec. 2.4) then it is possible to drop without loss of generality the "x" and "y" coordinate symbols from the notation and retain only the depth coordinate symbol "z" in the notation. In such contexts we agree to write:

$$\text{"h(z,}\xi\text{,t)"} \quad \text{or} \quad \text{"h(z,}\theta\text{,}\phi\text{,t)"} \quad \text{for} \quad h(x,\xi,t) \quad . \tag{10}$$

In particular, if ξ is $\mathbf{k}$ or $-\mathbf{k}$, which occurs in the important case of the two-flow theory (Sec. 8.3), then we agree further to write:

$$\text{"h(z,}\pm\text{,t)"} \quad \text{for} \quad h(z,\pm\mathbf{k},t) \quad , \tag{11}$$

where we read upper signs together and then lower signs together to obtain two separate definitions. As usual, when the light field does not appreciably change in time, or when time is understood, we shall drop "t" from the notation. Applications of these concepts are taken up in Sec. 13.9.

Radiant Energy over Space

The discussion of this section is now continued by officially noting two interpretations of the term "radiant energy". The first interpretation centers on the simple connection that exists between scalar irradiance and radiant energy. Suppose X is a subset of an optical medium over which at time t there is defined a scalar irradiance function h for a given frequency ν. Let "U(X,t)" denote the radiant energy content of X at time t. That is, by the definition of u(x,t), we agree to write:

$$\text{"U(X,t)"} \quad \text{for} \quad \int_X u(x,t)\, dV(x) \tag{12}$$

and from (4):

$$U(X,t) = \int_X \big(h(x,t)/v(x,t)\big)\, dV(x) \tag{13}$$

where "V" is the volume measure of the optical medium. As a special case, if v(x,t) and u(x,t) are independent of x and t then (13) becomes:

$$U(X) = (h/v)\, V(X) \tag{14}$$

where, for this case, we have written:

$$\text{"U(X)"} \quad \text{for} \quad U(X,t)$$

$$\text{"h"} \quad \text{for} \quad h(x,t)$$

$$\text{"v"} \quad \text{for} \quad v(x,t) \quad .$$

It is clear from (12) that $U(\cdot,t)$ is V-*additive* and V-*continuous*. That is, for every two disjoint parts X_1 and X_2 of an optical medium:

$$U(X_1,t) + U(X_2,t) = U(X_1 \cup X_2,t) \quad , \tag{15}$$

and for every X and t:

$$\text{If} \quad V(X) = 0, \quad \text{then} \quad U(X,t) = 0 \quad . \tag{16}$$

Radiant Energy over Time

There is still one more interpretation that can be made of the term "radiant energy". The preceding interpretation of (12) is associated with the energy content of a given region X at time t. There is a complementary interpretation of the total energy incident on or leaving a surface S over an interval T of time. For this interpretation we write,e.g.:

$$\text{"}U^-(S,T)\text{"} \quad \text{for} \quad \int_T \int_S H(x,\Xi(\xi),t)\, dA(x)\, dt \tag{17}$$

i.e., $U^-(S,T)$ is the radiant energy incident on S over the time interval T. The hemisphere of incident radiant flux at each x is $\Xi(\xi)$, with ξ normal to S at x, in the inward sense. A complementary definition can be made for $U^+(X,T)$ using radiant emittance.

It is worthwhile isolating the important concept, occurring in (17), of radiant flux across a *general* surface S rather than just a collecting surface of the kind encountered in the sections above. Thus we write:

$$\text{"}P^-(S,t)\text{"} \quad \text{for} \quad \int_S H(x,\Xi(\xi),t)\, dA(x) \quad , \tag{18}$$

where ξ is the unit inward normal to S at x. A similar definition of $P^+(S,t)$ can be phrased. As usual, the signs "+" and "-" can be dropped whenever no confusion results, and also the "t" can be omitted for brevity.

Scalar Radiant Emittance

We conclude this section with the definition of the notion of scalar radiant emittance. This concept is the surface-counterpart to scalar irradiance h defined in (3). Thus, let us write:

$$\text{"}w(x,t)\text{"} \quad \text{for} \quad \int_\Xi N^+(x,\xi',t)\, d\Omega(\xi') \tag{19}$$

$w(x,t)$ is the scalar radiant emittance at x at time t. This concept is useful in describing certain sources of radiant flux distributed continuously over some region of an optical medium. The emittance counterparts to hemispherical scalar irradiance emittance can now be defined for $w(x,t)$. These definitions would exactly parallel those in (5), (6), (7), (8), (10), (11), and therefore need not be given in detail at this time.

2.8 Vector Irradiance

The radiometric concept of vector irradiance, which will now be considered, constitutes an interesting and useful complementary concept to that of scalar irradiance. Whereas scalar irradiance in essence measures the volume density of radiant energy at a point and does so without emphasis on the directions of incidence of the component flows but only their magnitudes, vector irradiance in contrast gives a measure of the direction of the preponderant flow of radiant energy at the point without emphasis on the magnitude of the various component flows. Besides serving to complement the geometric properties of scalar irradiance in this way, vector irradiance forms a rigorous tool in deriving the transfer equations for scalar irradiance, and also a powerful means of measuring precisely and directly the absorption properties of real optical media. The basis for the latter means (the divergence relation for **H**) is considered in Chapter 8 and some of its applications are discussed in Chapters 6 and 13. In this section emphasis will be on introducing and explicating the geometric and physical meanings of vector irradiance.

A Mechanical Analogy

The notion of vector irradiance can be introduced by means of an analogy with the vectorial treatment of forces in static mechanics. Figure 2.19 (a) depicts a force diagram familiar to beginning students in static mechanics. A particle at point P is subject simultaneously to two steady forces of magnitude F_1 and F_2 along directions ξ_1 and ξ_2. In order to establish equilibrium of the particle--i.e., to balance out F_1 and F_2 so that the particle is stationary, another force of magnitude F_3' must be applied along direction ξ_3'. The magnitude F_3 and direction ξ_3 of the equivalent force that may replace F_1 and F_2 is found by means of the familiar parallelogram of forces shown in Fig. 2.19 (b). The required balancing magnitude is then $-F_3'$ and its direction is $-\xi_3'$, which follows directly from Newton's Third Law. The central observation to be made here is that, for the purpose of static equilibrium, two forces $\mathbf{F}_1 (= F_1\xi_1)$ and $\mathbf{F}_2 (= F_2\xi_2)$ can be replaced by a single force $\mathbf{F}_3$ $(= F_3\xi_3)$ which serves as a *mechanical* equivalent of the set of forces consisting of F_1 and F_2 together. Thus, $\mathbf{F}_3$ is, for the purposes of an equilibrium computation, equivalent to $\mathbf{F}_1+\mathbf{F}_2$.

Consider now a point P irradiated by two beams of radiant flux which are flowing along directions ξ_1 and ξ_2 with

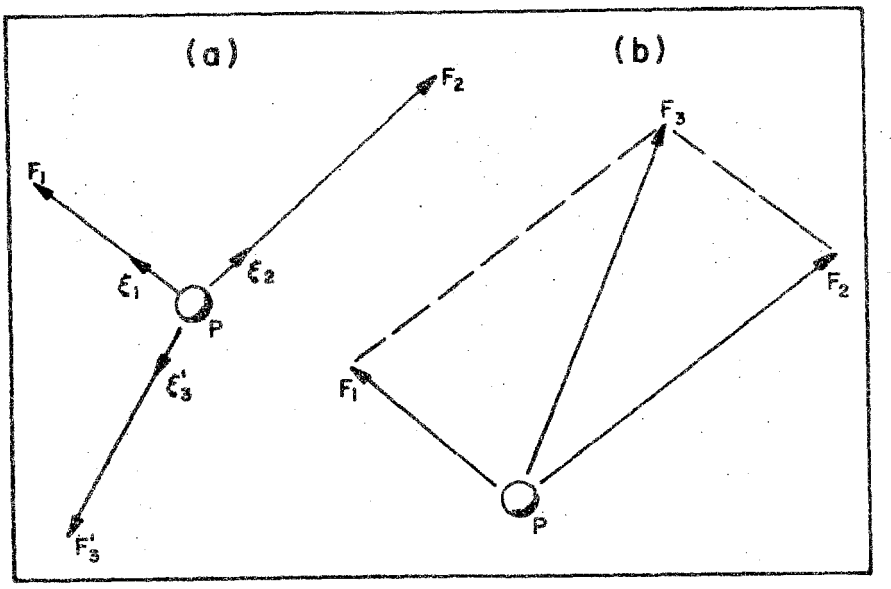


FIG. 2.19 The parallelogram law in mechanics

radiance N_1 and N_2, respectively, as in Fig. 2.20 (a). Each radiance has a fixed small solid angle Ω. Now whereas the mechanical context of Fig. 2.19 (a) is meaningful in terms of sets of directed forces and equivalent single forces, the context of Fig. 2.20 (a) is meaningful in terms of sets of directed radiances and equivalent single radiances. In the mechanical setting, a single force could, for the purpose of an equilibrium computation, replace the two given forces by a single force F_3. We now ask: can we replace the two directed radiances N_1 and N_2 by a single equivalent radiance N_3?

Some thought will show that, before the preceding question can even be entertained, the sense of "equivalent" must be defined. Clearly, the replacing radiance can be "equivalent" in any one of several desirable ways. For example, if it is required that the replacing radiance produce the same scalar irradiance at P, then there are many possible candidates for N_3. If on the other hand it is required that the replacing radiance produce the same net irradiance on an arbitrary collecting surface at P, then there is generally one and only one radiance N_3 that can replace N_1 and N_2 in this sense. Observe that the replacing radiance N_3 must be equivalent to N_1 and N_2 in this sense *not just for one* fixed position of a collecting surface at P; if that were the case, then N_3 could be chosen from any of an infinite number of radiances. Rather, N_3 is to produce the same effect for *all* possible orientations of a collecting surface at P. The analogy here with the mechanical context is essentially exact: in the mechanical context $\mathbf{F}_3$ establishes the same net force on

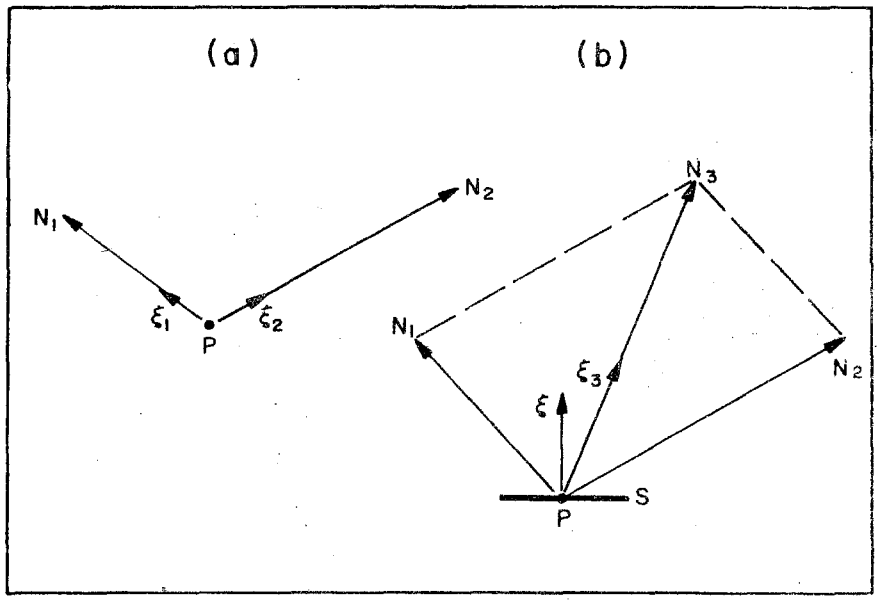


FIG. 2.20 The parallelogram law in radiometry

any particle at P as does $\mathbf{F}_1$ and $\mathbf{F}_2$; in the radiometric context N_3 is to establish the same net irradiance on any surface at P as does N_1 and N_2. It is a simple matter now to prove that the parallelogram law may also be used for the radiometric context to solve the analogous problem in that setting. Thus, the requisite replacing radiance N_3 and its associated direction ξ_3 of flow follow from a parallelogram construction as in Fig. 2.20 (b). In particular, if we write:

$$\text{"}\mathbf{N}_1\text{"} \quad \text{for} \quad \xi_1 N_1$$

and

$$\text{"}\mathbf{N}_2\text{"} \quad \text{for} \quad \xi_2 N_2$$

then $\mathbf{N}_1+\mathbf{N}_2$ is the requisite vector radiance provided it has a solid angle Ω. For if ξ is the inward unit normal to a collecting surface S at P, then we have by (6) of Sec. 2.5:

$$\xi\cdot\xi_1 N_1 + \xi\cdot\xi_2 N_2$$

as the expression for the total net irradiance on S produced by the two beams. This sum may be written:

$$\xi\cdot\mathbf{N}_1\Omega + \xi\cdot\mathbf{N}_2\Omega$$

or, as:

$$\xi\cdot(\mathbf{N}_1 + \mathbf{N}_2)\Omega \quad .$$

This representation suggests that if we direct a radiance beam of solid angle Ω at P and along the *direction* of $\mathbf{N}_1 + \mathbf{N}_2$

and with the *magnitude* of $\mathbf{N}_1 + \mathbf{N}_2$, then this single beam will produce the same net irradiance across S at P as the two given beams. The vector $\mathbf{N}_1 + \mathbf{N}_2$, which we have denoted by "$\mathbf{N}_3$" in Fig. 2.20 (b), is found exactly as in any vectorial addition operation.

The observations just made can be generalized to the case of any finite set of beams irradiating a point x in a radiometric environment. Toward this end, suppose that the various beams have radiances $N_1,\ldots, N_k$ along directions $\xi_1,\ldots, \xi_k$ and that, for generality, they have generally distinct solid angles, $\Omega_1,\ldots, \Omega_k$, respectively. Then by repeated use of (6) of Sec. 2.5, the *net* irradiance produced on a surface S with unit inward normal ξ at x is:

$$\xi\cdot\xi_1 N_1\Omega_1 + \xi\cdot\xi_2 N_2\Omega_2 + \ldots + \xi\cdot\xi_k N_k\Omega_k \quad .$$

Suppose we write:

$$\text{"H(x)"} \quad \text{for} \quad \sum_{j=1}^{k} \xi_j N_j \Omega_j \quad .$$

Clearly $\mathbf{H}(x)$ is a vector and its magnitude $|\mathbf{H}(x)|$ has dimensions of irradiance. Furthermore $\mathbf{H}(x)$ has the property that:

$$\xi\cdot\mathbf{H}(x)$$

is the *net* irradiance on the surface S at x produced by the set $N_1,\ldots, N_k$ of radiances at x. In the introductory example considered above $\mathbf{H}(x)$ was $\xi_1 N_1\Omega + \xi_2 N_2\Omega$. We shall call $\mathbf{H}(x)$ the *vector irradiance* associated with the discrete radiance distribution $N_1,\ldots, N_k$.

General Definition of Vector Irradiance

We now can go one step further in the development of the idea of vector irradiance. Instead of a discrete finite set of radiances $N_1,\ldots, N_k$ at x, we consider a general radiance distribution $N(x,\cdot)$. Instead of the finite summation over the sets of directions of the radiances in (1), we use the continuous counterpart to the sum, namely an integral over all the directions ξ at x. Thus let us write:

$$\text{"H(x)"} \quad \text{for} \quad \int_{\Xi} \xi N(x,\xi)\, d\Omega(\xi) \tag{2}$$

and where, in turn, the integral uses field radiance and is to be understood as an ordered triple of integrals, as is customary in vector analysis. That is, we have written:

$$\text{"}\int_{\Xi} \xi N(x,\xi)\, d\Omega(\xi)\text{"} \quad \text{for}$$

$$\left(\int_{\Xi} N(x,\xi)\cos\vartheta_1\, d\Omega(\xi), \int_{\Xi} N(x,\xi)\cos\vartheta_2\, d\Omega(\xi), \int_{\Xi} N(x,\xi)\cos\vartheta_3\, d\Omega(\xi)\right) \tag{3}$$

and where ϑ_1, ϑ_2, and ϑ_3 are as defined in Sec. 2.5 (cf., e.g., (18) of Sec. 2.5). We call $\mathbf{H}(x)$ the *vector irradiance* at x. The alternate form (3) of the integral in (2) is the form in which $\mathbf{H}(x)$ is computed in actual practice. The integral in (2) is a compact symbol for the ordered triple of integrals in (3). A researcher requiring the direction and magnitude of $\mathbf{H}(x)$ at x knowing $N(x,\cdot)$ at that point, computes the three components of $\mathbf{H}(x)$ in accordance with (3). Thus, if we write:

$$\text{"}H_i(x)\text{"} \quad \text{for} \quad \int_{\Xi} N(x,\xi)\cos\vartheta_i\, d\Omega(\xi) \tag{4}$$

for $i = 1,2,3$, then:

$$\mathbf{H}(x) = (H_1(x), H_2(x), H_3(x)) \quad . \tag{5}$$

The magnitude of $|\mathbf{H}(x)|$ of $\mathbf{H}(x)$ is:

$$(H_1^2(x) + H_2^2(x) + H_3^2(x))^{1/2} \tag{6}$$

and the direction of $\mathbf{H}(x)$ is the unit vector:

$$(H_1(x), H_2(x), H_3(x)) \Big/ (H_1^2(x)+H_2^2(x)+H_3^2(x))^{1/2} \quad . \tag{7}$$

The General Cosine Law for Irradiance

The cosine law for irradiance was introduced in Sec. 2.4 in a rather special context. Our purpose here is to show how the law can be given the status of a general theorem in radiometry. Thus we will free the cosine law in (16) of Sec. 2.4 from the restrictions placed on it in that section. The means by which the generalization can be accomplished is the notion of vector irradiance. The law may be stated as follows: *Let* $N(x,\cdot)$ *be a radiance distribution at point* x *in an optical medium. Let* ξ *be the unit inward normal to a surface* S *at* x. *Then the vector irradiance* $\mathbf{H}(x)$ *at* x, *as defined in (2) has the property that:*

$$\boxed{\xi\cdot\mathbf{H}(x) = |H(x)|\cos\vartheta} \tag{8}$$

where "$|H(x)|$" *denotes the magnitude of* $\mathbf{H}(x)$, *as given by (6), and* "ϑ" *denotes the angle between* ξ, *and the direction of*

$\mathbf{H}(x)$ *as given by (7). Furthermore,*

$$|\mathbf{H}(x)| = \max_{\xi} \bar{H}(x,\xi) \tag{9}$$

i.e., $|\mathbf{H}(x)|$ *is the maximum of the set of all net irradiances* $\bar{H}(x,\xi)$ *at* x*, where the net irradiance* $\bar{H}(x,\xi)$ *at* x *across* S *in the direction of* ξ *is as defined in (14) of Sec. 2.4.* The proof of (8) is immediate, since (8), stripped of all physical connotations, simply constitutes an elementary theorem in vector analysis. Equation (9) is the more deep of the two and follows from the observation that each $H_i(x)$, $i = 1,2,3$, can be written out in full form as:

$$\begin{aligned} H_1(x) &= \int_{\Xi(\xi)} N(x,\xi) \cos \mathcal{V}_1 \, d\Omega(\xi) \\ &\quad + \int_{\Xi(-\xi)} N(x,\xi) \cos \mathcal{V}_1 \, d\Omega(\xi) \\ &= H(x,\mathbf{i}) - H(x,-\mathbf{i}) = \bar{H}(x,\mathbf{i}) \quad , \end{aligned} \tag{10}$$

where the first equality results from writing Ξ as the union of two disjoint hemispheres $\Xi(\xi)$ and $\Xi(-\xi)$ and where the second equality follows from two applications of (8) of Sec. 2.5. In a similar way we show that:

$$H_2(x) = H(x,\mathbf{j}) - H(x,-\mathbf{j}) = \bar{H}(x,\mathbf{j}) \tag{11}$$

$$H_3(x) = H(x,\mathbf{k}) - H(x,-\mathbf{k}) = \bar{H}(x,\mathbf{k}) \quad . \tag{12}$$

In this way we uncover the physical significance of the three components $H_1(x)$, $H_2(x)$, and $H_3(x)$ of $\mathbf{H}(x)$. For example, $H_1(x)$ is the *net irradiance* across a plane at x whose inward normal is the coordinate unit vector $\mathbf{i}$ along the x axis. Continuing on our way to establish (9), we now examine $\xi \cdot \mathbf{H}(x)$ directly:

$$\begin{aligned} \xi \cdot \mathbf{H}(x) &= \xi \cdot \left(\int_{\Xi} \xi' N(x,\xi') \, d\Omega(\xi') \right) \\ &= \int_{\Xi} \xi \cdot \xi' N(x,\xi') \, d\Omega(\xi') \\ &= \int_{\Xi(\xi)} \xi \cdot \xi' N(x,\xi') \, d\Omega(\xi') + \int_{\Xi(-\xi)} \xi \cdot \xi' N(x,\xi') \, d\Omega(\xi') \\ &= \int_{\Xi(\xi)} \xi \cdot \xi' N(x,\xi') \, d\Omega(\xi') - \int_{\Xi(-\xi)} (-\xi) \cdot \xi' N(x,\xi') \, d\Omega(\xi') \\ &= H(x,\xi) - H(x,-\xi) = \bar{H}(x,\xi) \end{aligned} \tag{13}$$

This sequence of five equations is justified analogously to the sequences culminating in (10), (11), and (12). Now, however, we have included more detailed steps. Clearly, (13) subsumes (10)-(12). In view of (13), we may write (8) as:

$$\bar{H}(x,\xi) = |\mathbf{H}(x)| \cos \vartheta \tag{14}$$

and now (9) follows immediately from this. The maximum value of $\bar{H}(x,\xi)$ occurs when $\vartheta = 0$. From this, we have (9). Thus $|\mathbf{H}(x)|$ is simply the net irradiance across that surface S at x whose unit inward normal is the same as the direction of $\mathbf{H}(x)$.

The results (8)-(14) are of importance in both theoretical and experimental radiative transfer. An intuitive feeling for $\mathbf{H}(x)$ and for equations (8) and (14) may be obtained by imagining an experimental device of the kind schematically depicted in Fig. 2.21. The device has two collecting surfaces S_+ and S_- placed so that S_+ and S_- together receive radiant flux from every direction in Ξ. Further, the unit inward normal ξ to S_+, may be represented by a wire with a pointer welded to one end, and the whole arrow fastened to the material collecting surfaces as shown schematically in Fig. 2.21. The meter for the device is wired to read $H(x,\xi)-H(x,-\xi)$, i.e., the recorded irradiance on S_+ minus the recorded irradiance on S_-. A device so constructed is called a *subtracting janus plate*, (where "janus" has the same etymology as "January") and may be used to empirically determine $\mathbf{H}(x)$ in natural optical media. To operate the device, one

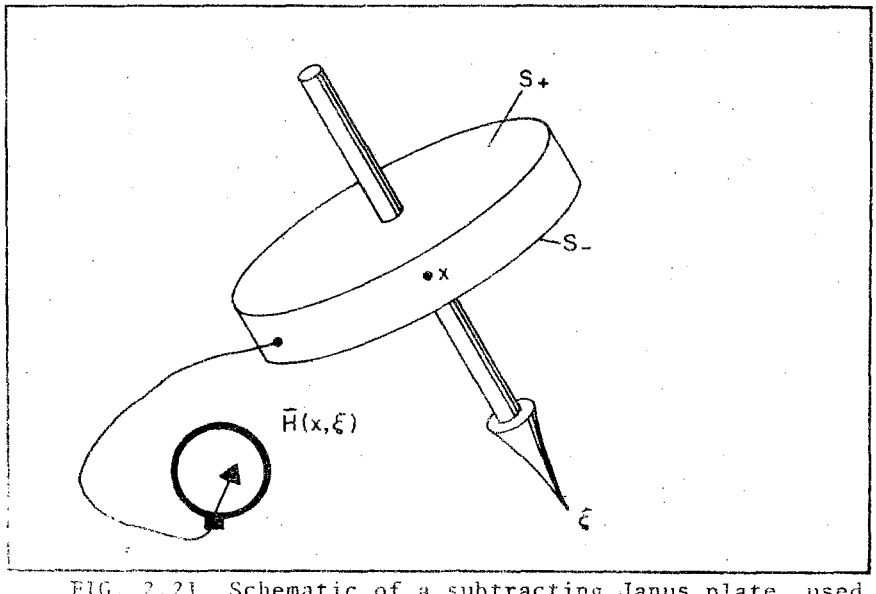


FIG. 2.21 Schematic of a subtracting Janus plate, used in measuring the vector irradiance field.

orients it at a point x so that the reading $\bar{H}(x,\xi)$ attains a maximum. Then by (9) the magnitude of H(x) is this maximum reading, and the direction of H(x) is the direction of the arrow fastened to the device. The full geometric and physical significance of (8) and (14) can now spring forth: regardless of the complexity of the radiance distribution $N(x,\cdot)$ --and as the eye is a witness, such complexity can be subtle and of infinite variety in natural optical media--the meter reading of the janus plate varies precisely in a sinusoidal fashion as the meter's direction is varied and as $\mathscr{V}$ increases from 0 to π. By means of H(x) we can develop a theory of the light field which is similar in many respects to certain classical fluid flows in hydrodynamics. This analogy has been explored by Gershun [98], and by Moon and Spencer [187].

Equations (8) and (14), aside from their intrinsic mathematical interest appear to have in store potential practical applications. For example, (14) may be of use in facilitating the practical task of computing and tabulating irradiances $H(x,\xi)$ over all directions ξ at given points x in optical media. Such a task is encountered, for example, when contrast calculations in submarine environments are desired. Specifically, by (14), it is clear that at a point x a full tabulation of $H(x,\xi)$ over all ξ in Ξ need not require a correspondingly full computation. Indeed, appropriate use of Equation (14) cuts the requisite work just about in half. Thus, one can compute $H(x,\xi)$ for each direction ξ over some pre-selected hemisphere, say $\Xi(\xi_0)$, and also compute H(x), whence $|H(x)|$ is obtained. Then from (14) we have:

$$H(x,-\xi) = H(x,\xi) - |H(x)| \cos \mathscr{V} \tag{15}$$

where ξ is in $\Xi(\xi_0)$, and so $-\xi$ is in $\Xi(-\xi_0)$, the complement of $\Xi(\xi_0)$ with respect to Ξ. Thus for every $-\xi$ in $\Xi(-\xi_0)$ one computes $H(x,-\xi)$ using the already tabulated value $H(x,\xi)$, the angle $\mathscr{V}$, and $|H(x)|$.

We conclude the present discussion of the general cosine law for irradiance by casting its basic form (8) into one which comes as close as possible to its special counterpart (16) of Sec. 2.4. Thus let "**m**" denote the unit vector associated with H(x), i.e., **m** is the direction of H(x) as computed by (7). Then, by (14), we have $\bar{H}(x,\mathbf{m}) = |H(x)|$ which is the maximum net irradiance at x. Further, in (14), $\cos \mathscr{V} = \xi \cdot \mathbf{m}$; hence (14) becomes:

$$\boxed{\bar{H}(x,\xi) = \bar{H}(x,\mathbf{m})\mathbf{m}\cdot\xi} \tag{16}$$

Clearly (16) of Sec. 2.4 is a special case of (16) above when radiant flux is incident on x in accordance with the restrictions on the earlier equation.

2.9 Radiant Intensity

The concept of radiant intensity, the last of the set of basic radiometric concepts to be introduced in this chapter, is designed to give a measure of the solid angle density of radiant flux. Thus radiant intensity is a dual concept to irradiance in the sense that the latter gives a measure of the *area density* of radiant flux while the former gives a measure of the *solid angle density* of radiant flux. At one time the concept of radiant intensity enjoyed the place now occupied by radiant flux. In the days of the candle and gas lamp there were very few artificial extended light sources. The controlled artificial point source, such as a candle's flame, was the sole basis for radiometric standards and its radiant output was conveniently measured in terms of intensity. However, with the passing of years the numbers of extended artificial light sources increased and the classical mode of use of radiant intensity has become correspondingly less frequent than that of radiant flux. Eventually radiance for the most part usurped the once centrally disposed radiant intensity concept. Despite this displacement of radiant intensity's status, it appears that there will always exist times when its use arises naturally. For example when emitting 'point sources' are considered, the use of radiant intensity seems automatically indicated. This useful aspect of radiant intensity will be discussed during its systematic development, to which we now turn.

Operational Definition of Empirical Radiant Intensity

In presenting the concept of radiant intensity we shall be guided by operational considerations so as to give the concept a secure footing relative to the other radiometric concepts already defined. Thus our first encounter with the notion of radiant intensity is in the following context: in the operational definition of P(S,D) (Sec. 2.3), a radiant flux meter with collecting surface S and collecting directions D and monochromatic filter passing a single frequency ν records an associated amount P(S,D) of radiant flux incident on S through the set of directions D. Once the datum P(S,D) is obtained, then one conceptual path leads, as in Sec. 2.4, to irradiance H(S,D), i.e., the area density of P(S,D) over S; another path leads to J(S,D), i.e., the solid angle density of P(S,D) over D, where we have written:

$$\text{"J(S,D)"} \quad \text{for} \quad P(S,D)/\Omega(D) \quad . \tag{1}$$

We call J(S,D) the (empirical) *radiant intensity* of P(S,D) over D on S. The dimensions of radiant intensity are *radiant flux per solid angle (per unit frequency interval)*, and convenient units are *watts/steradian (per unit frequency interval)*. In full notation for the unpolarized context, we would write:

$$\text{"J(S,D,t,F)"} \quad \text{for} \quad \Phi(S,D,t,F)/\Omega(D)$$

or:

$$\text{"}J(S,D,t,\nu)\text{"} \quad \text{for} \quad P(S,D,t,\nu)/\Omega(D) \quad . \tag{2}$$

However, we shall need only to employ the briefer notation in most of our discussions.

An examination of the operational definition of empirical radiant intensity, summarized in (1), will show that there is no restriction on the set of directions D. That is, D may be an arbitrary fixed set of directions along which the radiant flux funnels down onto the points of the collecting surface S. In practice, however, a radiance meter is the device used to estimate the radiant intensity of the light field at a point in an optical medium. In such an instrument, the set D is a relatively narrow conical bundle of directions whose axis is perpendicular to the collecting surface S of the meter. The connection between field radiance N(S,D), and radiant intensity in such a context, follows from (1) and is readily stated:

$$J(S,D) = N(S,D)A(S) \tag{3}$$

The connection between J(S,D) and N(S,D) can be generalized to take into account radiant flux which crosses S obliquely within the narrow set of directions D. The geometric setting is essentially that depicted in Fig. 2.6, the setting for the cosine law for irradiance.

To establish the generalized version of (3), we return to (1) and within the setting of Fig. 2.6, compute P(S',D) i.e., the radiant flux over S':

$$P(S',D) = P(S,D) = N(S,D)A(S)\Omega(D) \quad .$$

The reason for the equality of P(S',D) and P(S,D) stems from the hypothesized setting of Fig. 2.6, and the arguments presented earlier. Therefore, from (1):

$$J(S',D) = P(S',D)/\Omega(D)$$

$$= N(S,D)A(S) \quad .$$

But:

$$A(S) = A(S') \cos \mathscr{A} \quad .$$

Hence:

$$J(S',D) = N(S,D)A(S') \cos \mathscr{A} \quad .$$

By the radiance invariance law:

$$N(S',D) = N(S,D) \quad .$$

Hence:

$$\boxed{J(S',D) = N(S',D)A(S') \cos \mathscr{A} \ ,} \tag{4}$$

whenever the inward unit normal ξ' to S' makes an angle ν with the central direction ξ of D, as in Fig. 2.6, Eq. (4) should be compared with the special case (6) of Sec. 2.5. It is worthwhile re-emphasizing that relations (3) and (4) are relations among *empirical* radiometric quantities, i.e., radiometric quantities obtained with the use of a radiance meter of finite solid angle opening $\Omega(D)$, and finite collecting surface area A(S). The more finely-honed *theoretical* radiometric concepts come later with the help of the various additive and continuity properties of Φ postulated in Sec. 2.3. The empirical concepts serve to establish the bridge between theoretical constructs and the immediately given physical realities. The empirical concepts serve also to block out in rough form the incipient analytical structures of the theoretical relations.

Field Intensity vs. Surface Intensity

There is a distinction that can be made in practice between two types of radiant intensity, a distinction that is exactly analogous to the distinction made in Sec. 2.5 between field and surface radiance. Indeed, by referring to Fig. 2.12 wherein is depicted the two types of radiance, $N^+(S,D)$ and $N^-(S,D)$, which in turn are defined as in (30) and (31) of Sec. 2.5, we are led to write:

$$\text{"}J^+(S,D)\text{"} \quad \text{for} \quad P^+(S,D)/\Omega(D) \tag{5}$$

and

$$\text{"}J^-(S,D)\text{"} \quad \text{for} \quad P^-(S,D)/\Omega(D) \quad , \tag{6}$$

in complete analogy to the definitions of W(S,D) and H(S,D) in (17) and (18) of Sec. 2.4. We call $J^-(S,D)$ the *field intensity* and $J^+(S,D)$ the *surface intensity* over D within S. The utility of this distinction and the basis for the names of these concepts rest once again on the remarks for $N^+(S,D)$ and $N^-(S,D)$ in Sec. 2.5. In actual practice in natural optical media it is the surface intensity which is used with greatest frequency. However, in these settings it is the field intensity (or rather radiance) which, in the final analysis, must be measured before the surface intensity is obtained. The basic quantitative connection between the two types of radiant intensity is analogous to that between surface and field radiance in (22) of Sec. 2.5:

$$\boxed{J^+(S,D) = J^-(S,D) \quad .} \tag{7}$$

It follows from (30) of Sec. 2.5 and (5) above that:

$$N^+(S,D) = W(S,D)/\Omega(D) = J^+(S,D)/A(S) \tag{8}$$

and from (31) of Sec. 2.5 and (6) above that:

$$N^-(S,D) = H(S,D)/\Omega(D) = J^-(S,D)/A(S) \quad . \tag{9}$$

Henceforth we shall drop the "+" and "-" superscripts from the symbol "J" when it is clear from the context (or immaterial) which interpretation of radiant intensity is to be used in reading a statement using the concept of radiant intensity. Occasionally, however, especially for the purpose of emphasizing a delicate point in a discussion, the plus and minus appendages will be reattached to "J". In general, the following rule may be observed in regard to the base symbols "J" and "N": whenever "+" and "-" are omitted from "J" and "N", then the associated statement or term in which "J" and "N" appear is valid under both surface and field interpretations.

Theoretical Radiant Intensity

Suppose now that in the operational definition (1) the set of directions D becomes smaller and smaller, such that it always contains the direction ξ and such that the flow of radiant energy is onto S. Then write:

$$\text{"}J(S,\xi)\text{"} \quad \text{for} \quad \lim_{D\to\{\xi\}} J(S,D) \tag{10}$$

The existence of this limit is guaranteed by the D-additive and D-continuity properties of Φ postulated in Sec. 2.3. The radiant intensity $J(S,\xi)$ is called the (theoretical) *radiant intensity* in the direction ξ on S. It is important to note that non-zero values of $J(S,\xi)$ are necessarily associated with surfaces S which have non-zero area A(S). This fact is based on the S-continuity property of Φ recorded in Sec. 2.3. Thus, by S-continuity and S-additivity of Φ we have:

$$\lim_{S\to\{x\}} J(S,\xi) = 0 \quad , \tag{11}$$

for every x in S. However, once again by S-continuity and S-additivity of Φ, we have from (1), (4), (10) and the definition of $N(x,\xi)$:

$$N(x,\xi) = \lim_{S'\to\{x\}} J(S',\xi)/\left[A(S')\xi\cdot\xi'(x)\right] \tag{12}$$

where $\xi'(x)$ is the unit inward (or outward)* normal to S' at x. See Fig. 2.6.

From (10) and the fundamental theorem of calculus we obtain:

$$P(S,D) = \int_D J(S,\xi)\, d\Omega(\xi) \quad . \tag{13}$$

---

*Recall our convention on field intensity and surface intensity stated above.

From (4) and (12) we have for similar reasons:

$$J(S,\xi) = \int_S N(x,\xi)\,\xi\cdot\xi'(x)\,dA(x) \qquad (14)$$

where $\xi'(x)$ is the unit inward (or outward) normal to S at x and is in $\Xi(\xi)$.

At this point it would be instructive to view (12) in terms of (5) of Sec. 2.5. Furthermore, one can compare (14) with its 'dual' in (8) of Sec. 2.5. This 'duality' stems from a comparison of what is held constant and what is varied in (8) of Sec. 2.5 and (14) above. In (8) of Sec. 2.5, x in S is held fixed while ξ varies over all directions in $\Xi(\xi)$. In (14), ξ in $\Xi(\xi)$ is held fixed while x varies over all points in S. Furthermore, while the integration in (8) of Sec. 2.5 was limited for physical reasons to a hemisphere $\Xi(\xi)$ at x, the integration in (14) is limited, for similar reasons, to an S over which $\xi'(x)$ also stays within $\Xi(\xi)$. Hence the duality between $H(x,\Xi(\xi))$ and $J(S,\xi)$ is quite deep and complete.

Radiant Intensity and Point Sources

As noted in the introductory statements of this section, radiant intensity first arose as a measure of the directional radiant flux output of spatially very small emitters of flux. We shall now show that this feature of radiant intensity can still be employed within the operational point of view adopted in the present development of geometrical radiometry. The net result of this observation will be the recovery of the original conceptual feature of radiant flux but in a manner which will, it is hoped, now be operationally meaningful.

We begin by defining the notion of a point source of radiant flux. A part Y of an optical medium X is a (radiometric) *point source* with respect to point x in X if the set D(Y,x) of directions subtended at x by the points of Y is such that $\Omega(D(Y,x)) \leq 1/30$. The basis for this definition rests in two facts, one empirical, and the other theoretical.

The empirical fact is that radiance meters with solid angle openings such that $\Omega(D) \leq 1/30$ have been found to be adequate for the practical purposes of geophysical optics to distinguish the radiance variations occurring in natural optical media. Hence any part Y of an optical medium X which can be encompassed by the field of view of a radiance meter located at point x in X is radiometrically a 'point source' of flux. It might be that Y is a ship or an extensive wheat field, or a large patch of ocean surface, or a great cumulus cloud. As long as these objects (they can be either opaque solids, surfaces, or certain well-defined nearly transparent volumes of water or air) fall within the field of view of a

standard radiance meter, they are considered 'point sources' with respect to that meter.

The second fact on which the definition of 'point source' is based is that a part Y of X, such that $\Omega(D(Y,x)) \simeq 1/30$, has the property that the irradiance from Y on a surface about point x will vary, to within one percent, inversely as the square of the distance from x to Y whenever Y is some definitely localizable object such as a ship, or patch of sky or ocean surface, etc. In short, according to the preceding definition, Y will be a point source of flux only if the inverse square law and cosine law for irradiance holds with respect to it to within one percent. (See Example 5, Sec. 2.11.) It might be of interest to take note of the logical structure of the preceding statement. In particular, we *do not* assert that "if a part Y of X is such that the inverse square law and cosine law hold with respect to it, then Y is a radiometric point source". By considering a spherical body Y, the reason for this may be seen (cf., Example 4, Sec. 2.11). Finally, we shall henceforth assume that in the determination of the surface radiance of a point source Y, the solid angle opening of the radiance meter can be adjusted so as to fit exactly the set D(Y,x).

Consider now a radiometric point source Y in a medium X. For definiteness, let the point source Y be a spherical region of radius a within a vacuum and which steadily emits radiant flux. Further, Y is such that it can be observed from all directions. Suppose it is required to estimate the radiant flux output of Y but the measurements are constrained for various reasons to take place a distance r not less than a units from the center y of Y. Figure 2.22 (a) depicts the present situation. By adjusting the meter's solid angle

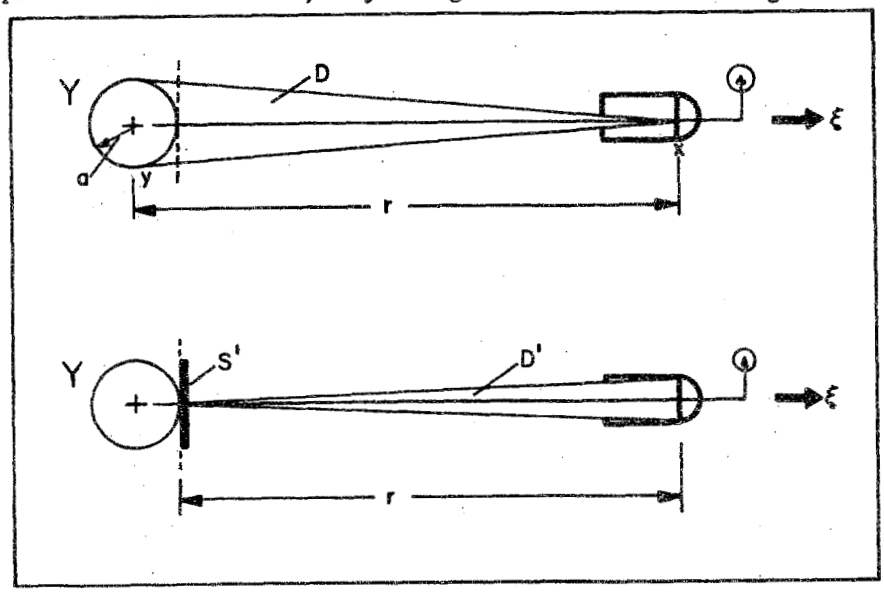


FIG. 2.22 Operational definition of a point source

opening so that the set D of directions from x to Y just fills the field of view, the field radiance N(S,D) associated with Y is read directly from the meter. Here S is the collecting surface of the meter. By the radiance invariance law (1) of Sec. 2.6, it follows that N(S,D) = N(S',D') where S' and D' are as shown in Fig. 2.22 (b), and are completely analogous to the observed surface and direction sets shown in Fig. 2.14. Hence the radiant flux output of Y across the projected surface S' of Y and within the set D' of directions is estimable as:

$$N(S',D')A(S')\Omega(D')$$

after using the measured radiance N(S,D) for N(S',D').

Now we have agreed to write:

$$\text{"}N(S',D')\text{"} \quad \text{for} \quad P(S',D')/A(S')\Omega(D')$$

where P(S',D') is the desired radiant flux output of Y in the direction D'. Since this radiance may be written as:

$$J(S',D')/A(S') \quad ,$$

we can now set:

$$P(S',D') = J(S',D')\Omega(D') \quad .$$

At this juncture the reader should first observe how the number N(S',D') 'belongs' to Y; that is, it is (by the radiance invariance law) independent of the mode of measurement. Secondly, it should be noted that of the two numbers A(S') and Ω(D'), the area A(S') 'belongs' to Y whereas Ω(D') depends on the mode of measurement (i.e., the distance r between x and y). It follows that the product N(S',D')A(S') 'belongs' to Y. But this product is simply J(S',D'), the radiant intensity (watts per steradian) of S' in the directions within D' from x to y. Hence the number J(S',D') is an intrinsic property of Y in the sense that it is independent of the mode of measurement. Finally, by recalling that the dimensions of J(S',D') contain no linear (i.e., length) terms, it becomes manifest that J(S',D') can be conceptually associated with the radiant flux output of the *point* y (the center of Y) in the direction ξ (the central direction of D'). In this way we arrive at the classical conception of radiant intensity as the radiant flux emitted by a point x per unit solid angle about a given direction ξ.

We can now use (13) as a basis for the classical formula relating the radiant intensity and radiant flux output of the point source Y. Since Y is a sphere, the projected area A(S') of Y on a plane normal to a direction ξ is independent of the direction ξ. More generally, in the point source context, we will agree to write:

$$\text{"}J(x,\xi)\text{"} \quad \text{for} \quad J(S',\xi)$$

$$\text{"}P(x,D')\text{"} \quad \text{for} \quad P(S',D')$$

whenever S' is the projection of part (or all of) the boundary of the point source Y on a plane normal to ξ, and x is some point within Y. Then, with this understanding, (13) becomes:

$$P(x,D') = \int_{D'} J(x,\xi)\, d\Omega(\xi)$$

$$= \int_{D'} J(x,\theta,\phi) \sin\theta\, d\theta\, d\phi \qquad (15)$$

where (for terrestrially-based coordinate systems) we have used (9) of Sec. 2.5, and have written "(θ,ϕ)" for ξ. If the radiant intensity output of Y is independent of ξ over D', then we can make the following statement:

If $D' = \Xi$, then,

$$P(x,D') = J(x)\Omega(D') \qquad (16)$$

i.e.,

$$P(x) = 4\pi J(x) \quad , \qquad (17)$$

where we have written:

"P(x)" for $P(x,\Xi)$

and

"J(x)" for $J(x,\xi)$.

Equation (17) is the customary form of the connection between the radiant intensity J(x) of a (directionally) uniformly emitting point source at x and its total power output P(x). By retracing the definitions of "P(x)" and "J(x)" the reader will see that the emitting object referred to is not a geometric point but rather a small finite part Y of an optical medium X, and that x is a point of X in or near Y. In this way we conceptually simplify the description of point sources to the form exhibited in (17) without contradicting the basic tenets of radiometry, in particular the S-continuity of Φ in Sec. 2.3.

Cosine Law for Radiant Intensity

The cosine law for radiant intensity (Lambert's law) can be stated as follows (cf., Fig. 2.6): *If the surface radiance* $N(S',\xi)$ *of point source surface* S' *is independent of direction* ξ *in* $\Xi(\xi')$, *where* ξ' *is the unit outward normal to* S' *at y, then the surface intensity* $J(S',\xi)$ *of* S' *varies as the cosine of the angle between* ξ' *and* ξ, *i.e.:*

$$\boxed{J(S',\xi) = J(S',\xi')\xi'\cdot\xi} \qquad (18)$$

The proof of statement (18) rests on (4). For by hypothesis we now can write:

$$N(S',D) = N(S',D')$$

where D' is a narrow conical set of directions whose central direction ξ' is normal to S', as in Fig. 2.6. Hence (4) becomes:

$$J(S',D) = J(S',D') \cos \mathscr{U} \quad .$$

Letting D and D' become smaller and smaller with limit $\{\xi\}$ and $\{\xi'\}$, respectively, we arrive at (18).

We have deliberately retained the notation of Fig. 2.6, despite the fact that (18) can be written with less primes adorning "S" and "ξ", for the purpose of encouraging a detailed comparison of (16) of Sec. 2.4 and (18) above. Close study will again reveal the interesting duality between intensity J and irradiance H already discerned by a comparison of (8) of Sec. 2.5 and (14) above. By dwelling on this recurrent duality between J and H, one is moved to inquire whether the cosine law for radiant intensity can be generalized to a form which would constitute a dual statement to the generalized cosine law for irradiance in the form of (8) or (16) of Sec. 2.8. It turns out that an exact dual statement to (8) of Sec. 2.8 can indeed be made for radiant intensity. Now, since the basis for the generalized cosine law for irradiance can be viewed as embodied in (8) of Sec. 2.5, we should expect the basis for the generalized cosine law for radiant intensity to rest in (14) above. We now show that this expectation is correct. We begin with deriving a result, of intermediate generality, from (14), a result which provides an interesting insight into the structure of the classical Lambert law.

Let Y be a region of an optical medium X. The region Y may be of arbitrary shape. Suppose further that from vantage point x, Y is a point source and that the observed surface radiance of its boundary surface is independent of the direction of observation of Y. For simplicity, we assume that the paths of sight from x to points of Y lie in a vacuum. The current geometric situation is depicted in Fig. 2.23. Let N(S,D) and N(S',D') be the observed surface radiances seen from two arbitrary vantage points x and x' at both of which Y is a point source. The surfaces S,S' and direction sets D and D' are as shown in the figure. Thus S is the projection of Y on a plane normal to the axis of the radiance meter located at x. Similarly for S'. Then by hypothesis and by the radiance invariance law:

$$N(S,D) = N(S',D') \quad .$$

This radiance equality can be written in terms of radiant intensity:

$$\frac{J(S,D)}{A(S)} = \frac{J(S',D')}{A(S')} \qquad (19)$$

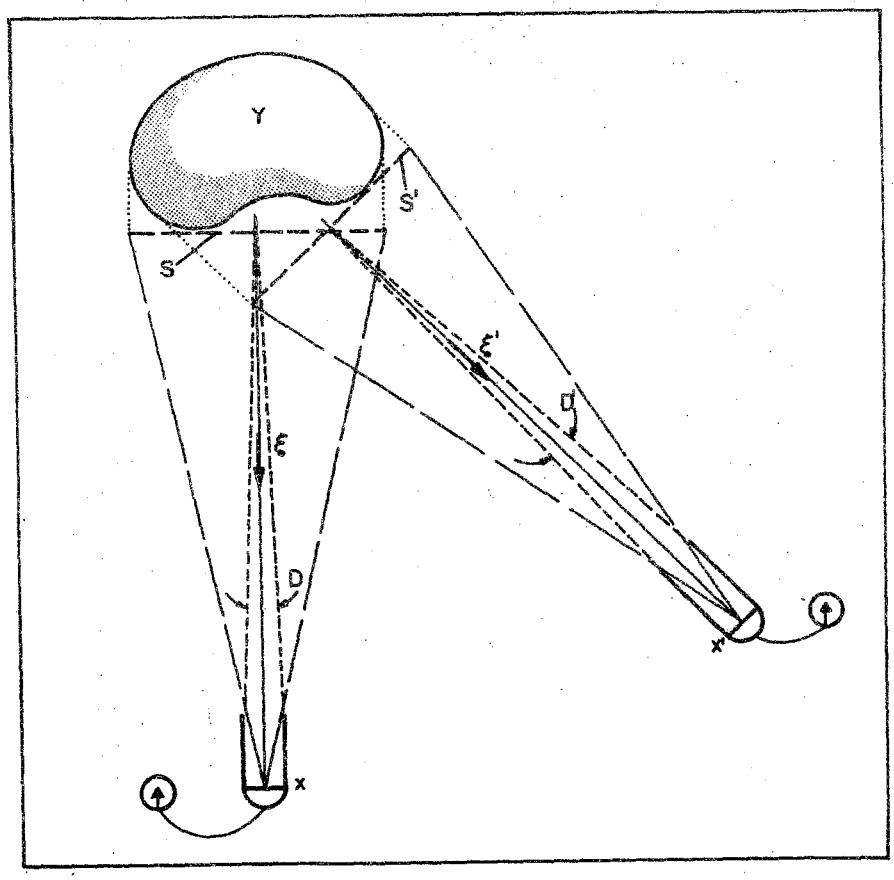


FIG. 2.23 Establishing the Cosine law for radiant intensity in the context of point sources.

Since the x and x' are arbitrary locations subject only to the requirement that Y is a point source with respect to these points, we arrive at the following slight generalization of the cosine law for radiant intensity:

If a part Y *of an optical medium* X *has uniform surface radiance for all directions and all points on the boundary of* Y, *and* Y *is a point source with respect to points* x *in some subset* X_0 *of* X, *and if the paths of sight from points of* X_0 *to* Y *lie in a vacuum, then the quotient* J(S,D)/A(S) *is invariant for every point* x *in* X_0, *where* S *and* D *are defined as in Fig. 2.23.*

It is clear how the classical form of Lambert's law (18) follows from this new statement and its analytic form (19); one now lets Y be a plane surface and lets X_0 be all the

appropriate points of X lying to one side of S.

It is of interest to note still one more variant of (19), one which has considerable intuitive value. Let "N" denote the hypothesized fixed radiance associated with Y. Then (19) implies that:

$$\boxed{J(S,D) = NA(S) \quad ,} \qquad (20)$$

i.e., that J(S,D) varies directly as the projected area A(S) of Y on a plane perpendicular to the central direction ξ of D. This may be compared with (3). From (20) we can deduce by inspection the *direct square law*--or *area law*--for radiant intensity which states that: *for areas which are radiometric point sources with respect to some observation point, the associated intensity varies directly as the apparent (projected) area of the surface as seen from that point.* If the area is compared with geometrically similar areas, then the associated radiant intensity varies directly as the square of a common transverse dimension of these areas. This observation brings to light still another facet in the duality between irradiance and radiant intensity, the dual law for irradiance in this case being the *inverse square law.*

Generalized Cosine Law for Radiant Intensity

The preceding discussions on point sources and radiant intensity lead us to formulate several useful alternative versions of the classical Lambert law for radiant intensity. During those discussions it was observed how the concepts of irradiance and intensity played the roles of dual concepts in a sense made clear in those discussions. This duality of irradiance and intensity is capable of being expressed in a precise fashion and on a level of generality comparable to that established for the general cosine law for irradiance (8) of Sec. 2.8. We now pause briefly in our developments of geometrical radiometry to establish this interesting generalization of Lambert's law. In doing so we round out and make formal the recurrent theme of duality between surface intensity J and irradiance H encountered throughout this section.

Let Y be a region of an optical medium X such that at each point x of the closed boundary surface S of Y the surface radiance is uniform over the hemisphere $\Xi(\xi'(x))$ where $\xi'(x)$ is the unit outward normal to S at x. Let "N(x)" denote the common value of the uniform radiance distribution at x on X over the set $\Xi(\xi'(x))$. Observe that the variation of the values N(x) over S is left to be quite arbitrary. For the present discussion the only restriction on the radiance function is that it be uniform over $\Xi(\xi'(x))$ at each point x of S. Some approximate physical realizations of such a region Y are: an opaque irregularly shaped body painted with matte paints such that the paints have an arbitrary spatial pattern over S; a luminous, dense region of space such as the sun which, for practical purposes, has a directionally nearly uniform radiance

distribution at each boundary point, but which still may be mottled with lighter or darker regions; the moon's surface forms still another example. However, when each of these objects is examined with extreme accuracy of radiometric detail in mind, a wealth of departures from these ideals is encountered.

Now returning to Equation (14) and using the present radiance function in the integral, we consider the particular integral:

$$\int_S N(x)\xi\cdot\xi'(x)\ dA(x) \quad . \tag{21}$$

Our studies of the duality between J and H led us to believe that we may be able to do for J what we did for H when going from (8) of Sec. 2.5 to (2) of Sec. 2.8. Therefore, we are led to take (21) as a base and write:

$$\text{"}\mathbf{J}(S)\text{"} \quad \text{for} \quad \int_S N(x)\xi'(x)\ dA(x) \quad . \tag{22}$$

We call $\mathbf{J}(S)$ the *vector intensity* for S. The definition of the integral is based on the notion of an ordered triple of integrals, using the form (3) of Sec. 2.8 as a model.

Now $\mathbf{J}(S)$ is a *bona fide* vector. As such it has three real numbers as x, y, and z components, and so a non zero magnitude $|\mathbf{J}(S)|$ and a direction $\mathbf{J}(S)/|\mathbf{J}(S)|$. This observation will allow us to state succinctly the radiant intensity analog to (8) of Sec. 2.8. However, before doing so, we explore one further facet of $\mathbf{J}(S)$.

Figure 2.24 depicts a typical region Y with boundary S for which $\mathbf{J}(S)$ is defined. If a direction ξ is chosen, then the boundary S of Y can be partitioned into two parts $S(\xi)$ and $S(-\xi)$ (or "S_+" and "S_-" for short) with the properties that $S(\xi)$ consists of all points x of S such that $\xi\cdot\xi'(x) > 0$, and $S(-\xi)$ consists of all points x of S such that $\xi\cdot\xi'(x) < 0$. There is generally, for all surfaces of use in practical radiometry, a closed curve C on S such that $\xi\cdot\xi'(x) = 0$ for every x on C. C is the boundary between S_+ and S_-. Observe how $S(\xi)$ and $S(-\xi)$ in the present context have their counterparts in the sets $\Xi(\xi)$, $\Xi(-\xi)$ used in the vector irradiance context.

Returning now to (14) we choose a ξ, determine the associated S_+ and S_- as just described, and then evaluate:

$$J(S_\pm,\xi) = \int_{S(\pm\xi)} N(x)\xi'(x)\ dA(x) \quad .$$

Suppose we write:

$$\text{"}\overline{J}(S,\xi)\text{"} \quad \text{for} \quad J(S_+,\xi)-J(S_-,\xi) \quad .$$

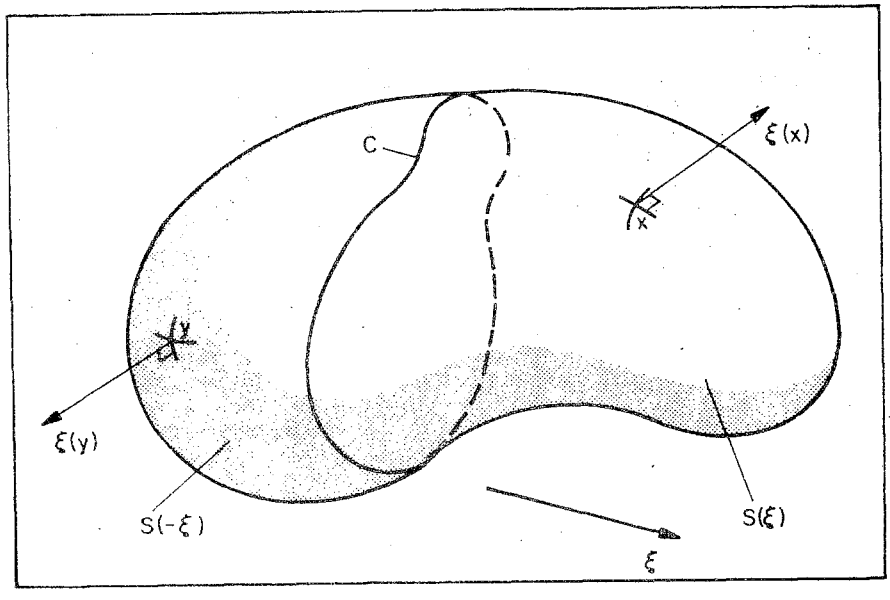


FIG. 2.24 Establishing the general Cosine law for radiant intensity.

Then from the definition of $J(S)$ and $J(S_{\pm},\xi)$ it follows that:

$$\xi \cdot \mathbf{J}(S) = \bar{J}(S,\xi) = J(S_+,\xi) - J(S_-,\xi) \quad . \tag{23}$$

We are now ready to state the generalized cosine law for radiant intensity.

Let $N(x)$ be a uniform radiance distribution over the hemisphere $\Xi(\xi'(x))$ at each point x of the boundary S of a region Y in an optical medium, where $\xi'(x)$ is the unit outward normal to S at x. Then the vector (surface) radiant intensity $\mathbf{J}(S)$ as defined in (22) has the property that:

$$\boxed{\xi \cdot \mathbf{J}(S) = |\mathbf{J}(S)| \cos \mathscr{v}} \tag{24}$$

where "$|\mathbf{J}(S)|$" denotes the magnitude of $\mathbf{J}(S)$ and "$\mathscr{v}$" denotes the angle between ξ and the direction of $\mathbf{J}(S)$. Furthermore:

$$\boxed{|\mathbf{J}(S)| = \max_{\xi} \bar{J}(S,\xi) \quad .} \tag{25}$$

The parallel of (24) and (25) with the irradiance case in (8) and (9) of Sec. 2.8 is exact. In particular, from (23) we can now write:

$$\boxed{\bar{J}(S,\xi) = \bar{J}(S,\mathbf{m})\mathbf{m}\cdot\xi \quad ,} \tag{26}$$

where $\mathbf{m}$ is the direction of $\bar{J}(S)$. This is the radiant intensity counterpart to (16) of Sec. 2.8. The special case (18) of Lambert's law now follows upon applying to (26) the conditions stated for (18). In particular Y now degenerates into a plane, we let $N(x) = 0$ on S_-, and $N(x)$ be constant on S_+. It should be noted in passing that (24) holds for regions including non-point sources. The duality between J and H now becomes clear upon comparison of, say (26) above with (16) of Sec. 2.8: a point x in the irradiance context is replaced by a surface S in the intensity context; the set $\Xi(\xi)$ in the irradiance context is replaced by the point ξ in the intensity context.

2.10 Polarized Radiance

In this section we shall develop an operational definition of polarized radiance. The development shall take as a point of departure the notion of empirical radiance introduced in Sec. 2.5. The details of the development shall be kept to a minimum, as we will not in this work make extensive use of the concept of polarized radiance. For a somewhat more detailed theoretical discussion of polarized radiance suitable for geophysical applications, the reader is referred to Chapter XII of Ref. [251].

Before going into the technical details of how to measure polarized radiance, a few comments may be made on the reason for wanting to measure polarized radiance in natural optical media. The first and most important reason is that the systematic documentation of the state of polarization of submarine (and atmospheric) light fields increases our store of basic optical knowledge of the world in which we live. For those of a more practical turn of mind, it may suffice to add that knowledge of the kind and amount of polarization extant in a natural light field could yield efficient means of increasing visibility in both the atmosphere and the sea. For, the contrast of objects seen against a sky or underwater background is occasionally increased when viewed through a material which can transmit polarized light in various amounts depending on how the material is held and oriented. If we possess systematic tabulations of polarized light fields and some workable theoretical models of such fields, these empirical observations can be more deeply explored and applied. Finally, there is the question, still not fully resolved--especially for the hydrologic optics branch of geophysical optics--of whether and to what extent polarized light is used by creatures in navigating, in foraging, and in their biological growth cycles.

In order to help resolve such questions and in order to add to our knowledge of the light fields in natural hydrosols in a systematic manner, we must develop a precise but workable means of measuring and theorizing about polarized light fields. A small but definite beginning in this direction will be attempted in this section and subsequently in Chapter 4 wherein the equation of transfer for the polarized radiance vector is used to derive a theoretical model of polarized light fields in natural optical media.

One final comment is in order. It will be recalled that our approach to hydrologic optics is through the tenets of radiative transfer theory and, as a consequence, we are committed to study the natural light fields on a phenomenological level. In particular, as pointed out in Sec. 2.0, we have agreed to adopt those instruments of investigation which make quantitatively precise, all of the optical phenomena visible to the human eye. One may then--in view of this observation--argue that in extending the capabilities of our instruments to detect and measure polarized light fields we are transcending the bounds originally set down by us when we embarked on the development of the concepts of radiometry. It may be observed, however, that whenever it is deemed necessary to extend the radiative transfer phenomena of concern to hydrologic optics in particular, or geophysical optics in general, the extension will be made solely on its merits to add to the descriptive power of these branches of radiative transfer theory.* In the present discussion, the extension of the radiometric concept of radiance to the polarized level not only fulfills this general criterion, but interestingly enough, still keeps the collection of radiometric concepts within that small, select circle of concepts which are directly observable by the unaided eye. For indeed, the polarization of the light of the sky or a submarine light field is directly observable to the unaided (but practiced) human eye. The physiological basis for this capability of direct observation is the dichroic nature of either the material comprising the yellow spot of the retina or perhaps that of certain of the optic nerve fibers themselves. (Dichroic materials are also found in natural deposits, e.g., in the form of tourmaline crystals, and were already used in the early devices for detecting polarized light.) It is the small but adequate amount of dichroic material in the retina which thus permits the unaided eye to detect and the brain to record the presence of linearly polarized light in a natural light field. This innate ability of the eye to detect polarized light was reported by Haidinger in 1846, and the elusive but yet visually observable pattern seen by the eye is known as *Haidinger's brush*. An informative description of how to facilitate the detection of Haidinger's brush in skylight is given by Minnaert in Ref. [182].

---

*The case for an extension of the classical scalar theory to the polarized level ultimately involves no less than the consistency of the classical scalar theory in the context of polarized light fields. See (17) of Sec. 13.11.

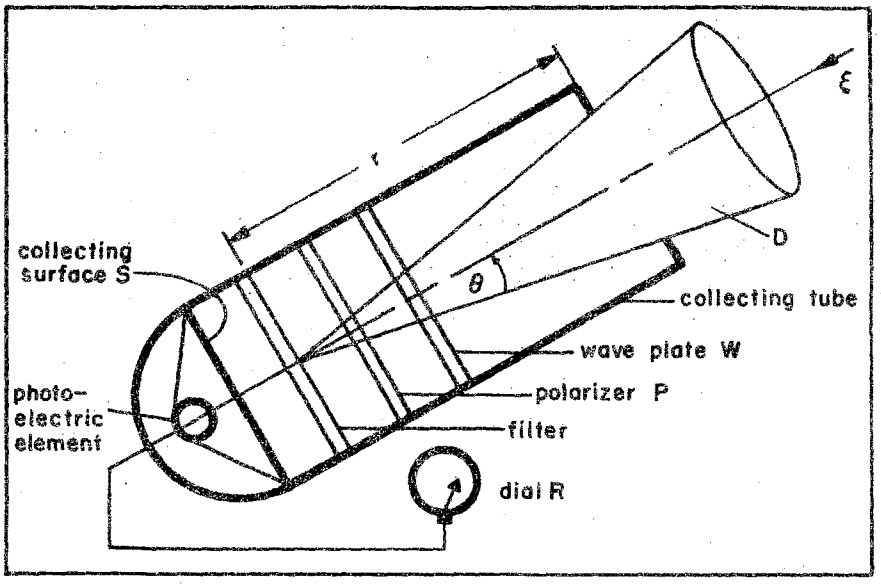


FIG. 2.25 Schematic details of a radiance meter fitted with a polarizer P and a variable wave plate W, for measuring polarized radiance distribution.

Operational Definition of Polarized Radiance

The operational definition of polarized radiance we shall adopt has been chosen for its inherent simplicity and its amenability to be linked with the classical Stokes vector for polarized light. Thus the polarized radiance vector will, on the one hand, be tied directly to observable qualities of natural light fields and, on the other, be rigorously representable by means of concepts extant in the electromagnetic picture of light.

We begin with a radiance meter, as described in Sec. 2.5, and adjoin to the meter, at the base of the tube, a polarizer P and a variable wave plate W. The order in which the entering light encounters these devices is important and is depicted in Fig. 2.25: the light is to encounter the variable wave plate first, and the polarizer second; then it passes on through the filter to the photoelement below. This relative placement of W and P is the essential point to observe here; where the filter is relative to W and P is, however, immaterial as far as ideal detectability of polarized flux is concerned.

The polarizer P, which is made from a dichroic crystal or a sheet of polaroid, is mounted so that it is rotatable about the axis of the cylindrical tube of the radiance meter. The orientation of the optic axis of the polarizer is important in what follows; therefore it is essential that some

means be provided for the clear marking of the position of the optic axis relative to the radiance meter's tube, or some other fixed part of the radiance meter. Further, if absolute radiance measurements are desired, the transmittance of P over the spectrum is required. The ideal transmittance of P is 1/2 for unpolarized light.

The wave plate can be made of some negatively doubly refracting material such as calcite, and is assembled (at least for the introductory discussion below) so that a wide range of optical path lengths is available at the twist of a knob. For example, a Babinet compensator type of arrangement may be employed. Later on, when the radiance meter is readied for field use, the wave plate may be replaced by an attachment fashioned from a single sheet of some good grade of circularly polarizing material of known transmittance over the spectrum. The ideal transmittance of W is 1, and that of circularly polarizing material, 1/2. The fact that a circular polarizing material can be used in lieu of a variable wave plate will become clear after the observable radiance vector has been defined.

The next step in the present operational definition of polarized radiance is to take the radiance meter, set for frequency ν, to a point x in the environment, and direct it so that flux enters the tube along the direction ξ at time t. Then one systematically varies the angle ψ of the polarizer's optic axis, starting from the vertical plane, with a given fixed retardation $\varepsilon \geq 0$ of the wave plate. (See Fig. 2.26) In fact one varies ψ from 0 radians (so that the optic axis is in the vartical plane) and increases ψ clockwise (when seen looking into the tube from the front of the tube, i.e., looking along the direction of travel of the photons) to π radians (so that the optic axis is again in the vertical plane). As this is done, one should note how the recorded flux varies with ψ and that the variation is of period π. In particular, for $\varepsilon = 0$ and a general light field, the variation of radiance turns out to be representable in the form:

$$\frac{1}{2}\left[I + Q\cos 2\psi + U \sin 2\psi\right] \quad ,$$

where we have written:

"I" for $2\bar{N}$

"Q" for $2\Delta N \cos 2\psi_o$

"U" for $2\Delta N \sin 2\psi_o$

where, in turn, we have written:

"$\bar{N}$" for $(N_{max} + N_{min})/2$

"ΔN" for $(N_{max} - N_{min})/2$,

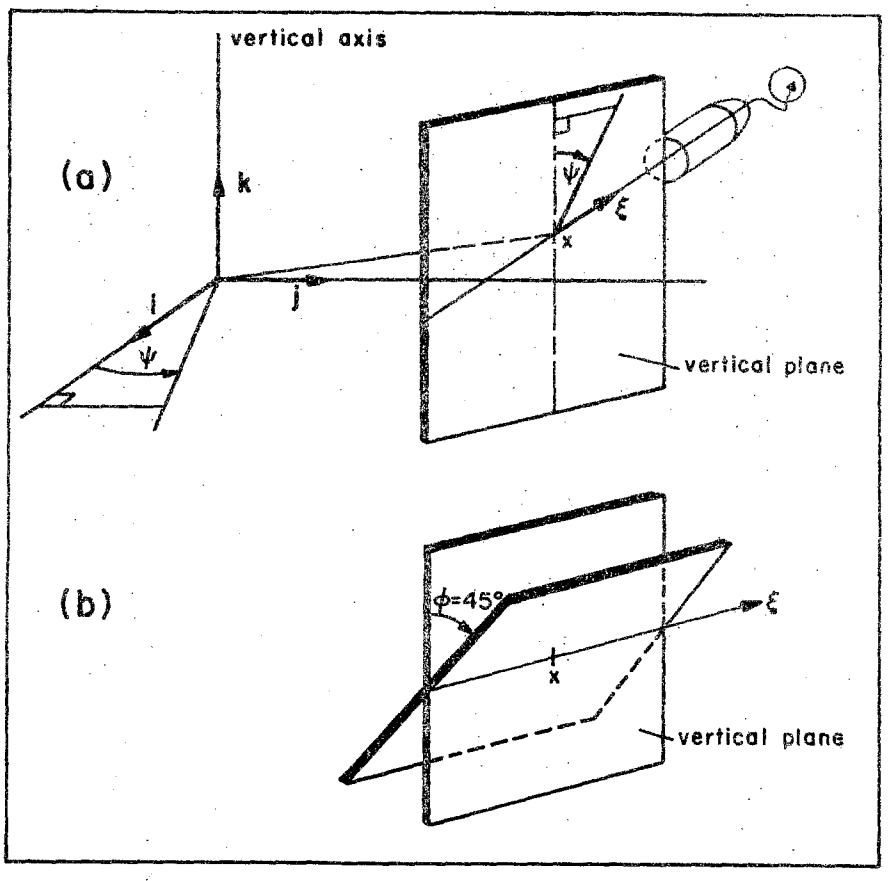


FIG. 2.26 How to measure and record the standard observable radiance vector.

and where "N_{max}" and "N_{min}" denote, respectively, the maximum and minimum radiance readings when ψ is varied from $\psi = 0$ to $\psi = \pi$. ψ_o is the angle of occurrence of the maximum reading N_{max}.

Further experimentation, with now a general ε-setting on W and with ψ varying over the interval $[0,\pi]$, shows the full form of the radiance variation to be:

$$\frac{1}{2}\left[I + Q\cos 2\psi + (U\cos\varepsilon - V\sin\varepsilon)\sin 2\psi\right] \qquad (1)$$

where V is determinable by a simple trigonometric analysis of the recorded curves obtained by fixing $\psi \neq 0$ and varying ε. (See, e.g., Ref. [193]). If this reading is obtained at point x, for the direction ξ at time t, for frequency ν, and with a P-setting ψ, and a W-setting ε, then we will agree to denote

it by "$N(x,\xi,t,\nu,\psi,\varepsilon)$", or, if x,ξ,t, and ν are understood, we will denote it simply by "$N(\psi,\varepsilon)$", for short. The development of the empirical basis of this quantity (using the "S", "D", "F" notation for radiant flux) is parallel to the unpolarized radiance case of Sec. 2.5, and therefore need not be repeated here. Expression (1) constitutes the desired operational definition of the polarized radiance $N(x,\xi,t,\nu,\psi,\varepsilon)$.

The Standard Stokes and Standard Observable Vectors

The operational definition (1) for polarized radiance and the experimental considerations leading to it draw out the remarkable fact that the most general polarized radiance field can be characterized by four functions, I,Q,U and V whose values are determined once a selection of ψ,ε along with x,ξ,t and ν are given. In view of the potentially infinite variety of specific forms that polarized radiance fields can assume, this is indeed a remarkably simple characterization and representation of the entire class of possible fields. This theoretical characterization of the polarized light field by I,Q,U, and V was first systematically studied by Stokes in 1852. We shall write:

$$\text{"S"} \quad \text{for} \quad (I,Q,U,V)$$

and call S the *standard Stokes vector*. S is a function which assigns to each choice of ψ,ε, along with x,ξ,t, and ν, an ordered quadruple of radiance numbers obtained as described above.

There is an alternate method of quantitatively documenting a polarized light field. Instead of obtaining I,Q,U, and V as described above, one may obtain four direct readings $N(\psi,\varepsilon)$ for the following four special pairs of settings ψ and ε. We write:

$$\text{"}{}_1N\text{"} \quad \text{for} \quad N(0,0)$$

$$\text{"}{}_2N\text{"} \quad \text{for} \quad N(\pi/2,0)$$

$$\text{"}{}_3N\text{"} \quad \text{for} \quad N(\pi/4,0)$$

$$\text{"}{}_4N\text{"} \quad \text{for} \quad N(\pi/4,\pi/2)$$

We then go on to form an ordered quadruple from these numbers; we write:

$$\text{"N"} \quad \text{for} \quad ({}_1N,{}_2N,{}_3N,{}_4N) \ ,$$

and call **N** the *standard observable vector*. **N** is a function which assigns to each choice of x,ξ,t, and ν the four numbers shown. Observe how **N** requires use of W only for the setting ${}_4N$. Readers familiar with the concepts of polarized light will see that each $N(\pi/4, \pi/2)$ can be obtained by means of a single reading using a piece of circularly polarizing material.

We shall see that S and N are equivalent descriptions of polarized light fields in the sense that knowledge of either description allows the deduction of the other. Since N is operationally the simpler of the two radiance vectors, we shall henceforth work with N. But before going on to the exlusive use of N, we shall establish some important and useful connections between S and N. Then, since S is computable directly from electromagnetic theory, we will have available a direct tie between N and electromagnetic concepts, to be used when needed.

As an illustration of how N characterizes the commonly occurring polarized radiances, consider the following list of special cases where "N" denotes the relative magnitudes of the components $_iN$ of N.

Verbal Description	Observable Radiance Vector
vertically linearly polarized radiance	$\frac{1}{2}(2N,0,N,N)$
horizontally linearly polarized radiance	$\frac{1}{2}(0,2N,N,N)$
linearly polarized radiance at +45°	$\frac{1}{2}(N,N,2N,N)$
linearly polarized radiance at -45°	$\frac{1}{2}(N,N,0,N)$
right circularly polarized radiance	$\frac{1}{2}(N,N,N,0)$
left circularly polarized radiance	$\frac{1}{2}(N,N,N,2N)$
unpolarized radiance	$\frac{1}{2}(N,N,N,N)$

To tie in these conventions with electromagnetic conventions, recall first that the optic axes of P and W lie initially in the standard preferred orientation, i.e., they lie in a vertical plane. Now the E-vector in vertically polarized light by convention lies in a vertical plane as it crosses planes perpendicular to the direction of travel. (Recall that the E-field is a transverse field.) In +45° linearly polarized light the E-vector lies in a plane tilted +45° from the vertical plane containing the direction of travel of the ray associated with the electric field E. The + direction of the +45° is measured clockwise as one looks along the direction of travel of the ray. Finally, right circularly polarized light is by convention that light associated with an E-vector whose tip describes a clockwise circular motion on a stationary plane perpendicular to its direction of travel as seen on the incident side of the plane. The most general light field can be resolved into its linear and elliptical components. This is the *Polarization Composition theorem* of Stokes.

Analytic Link Between S and N

The connection between the two vectors **S** and **N** is easily established by means of (1). On the basis of (1) and the definition of ${}_iN$, i = 1,2,3,4, we have:

$$N = \frac{1}{2}\,[I + Q]$$

$$N = \frac{1}{2}\,[I - Q]$$

$$N = \frac{1}{2}\,[I + U]$$

$$N = \frac{1}{2}\,[I - V]$$

From this set of equations we may construct the matrix $\mathscr{P}$ which transforms **S** into **N**. Thus let us write:

$$\text{"}\mathscr{P}\text{"} \quad \text{for} \quad \left(\frac{1}{2}\right) \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Then:

$$\mathbf{N} = \mathbf{S}\mathscr{P} \tag{2}$$

$$\mathbf{S} = \mathbf{N}\mathscr{P}^{-1} \tag{3}$$

where:

$$\mathscr{P}^{-1} = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

As an example of the use of (3), we obtain the following representation of vertically polarized radiance in terms of Stokes vectors:

$$\frac{1}{2}(2N,0,N,N) \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} = (N,N,0,0) \quad .$$

The reader will find it instructive to use (3) to obtain a list of Stokes vector representations of the seven special observable vectors given above.

Standard and Local Reference Frames

Up to this point in the exposition of the polarized radiance vector all operational activity has implicitly taken place in a terrestrial coordinate system. In particular the ψ-setting of the polarized radiance meter was such that if $\psi = 0$, then the optic axis of the polarizer P of the meter is in a vertical plane. (See Fig. 2.26 (a).) Such a frame of reference for polarized radiance measurements we call a *standard reference frame*. We now introduce a second reference frame--the *local reference frame*--whose main virtue and reason for being is that at each point in an optical medium it permits a simple means of experimental determination of the polarized volume scattering function. Furthermore, the introduction of the local reference frame considerably facilitates the formulation and handling of the various forms of the transfer equation in the polarized context.

To establish a local reference frame at a point x in an optical medium, two directions ξ' and ξ must be given (in, say, a terrestrial coordinate system) such that ξ' and ξ uniquely determine a plane. In other words, the only requirement on ξ' and ξ is that they not be collinear. We shall call ξ' and ξ, respectively, the *incident* and *scattered* directions, and the plane they determine together with a point x, the *plane of scattering*. See Fig. 2.27.

Once a plane of scattering is determined by x, ξ' and ξ, the incident polarized radiance is measured as follows: place the radiance meter at x so as to allow the flux in the direction ξ' to enter the meter's collecting tube. With the ψ-setting at 0 with respect to the meter's tube, rotate the entire radiance meter around ξ' so that the optic axis of the polarizer P lies in that plane A' through ξ' and perpendicular to the scattering plane. With the radiance meter so oriented, perform the four operations leading to ${}_iN$, $i = 1,2,3,4$. Designate the *local observable vector* by "$N_{\phi'}$" where "ϕ'" denotes the angle through which the vertical plane through ξ' must be rotated *clockwise* around ξ'--when looking along the direction of ξ'--so as to become coincident with the plane A'. Thus ϕ' varies from 0 to π; similarly with ϕ for the radiance determination in the scattered direction ξ. In general it can be shown (Sec. 111, Ref. [251]) that the standard observable vector N is related to a local observable vector N_ϕ by means of the equation:

$$N_\phi = N\mathcal{L}(\phi) \tag{4}$$

where we have written:

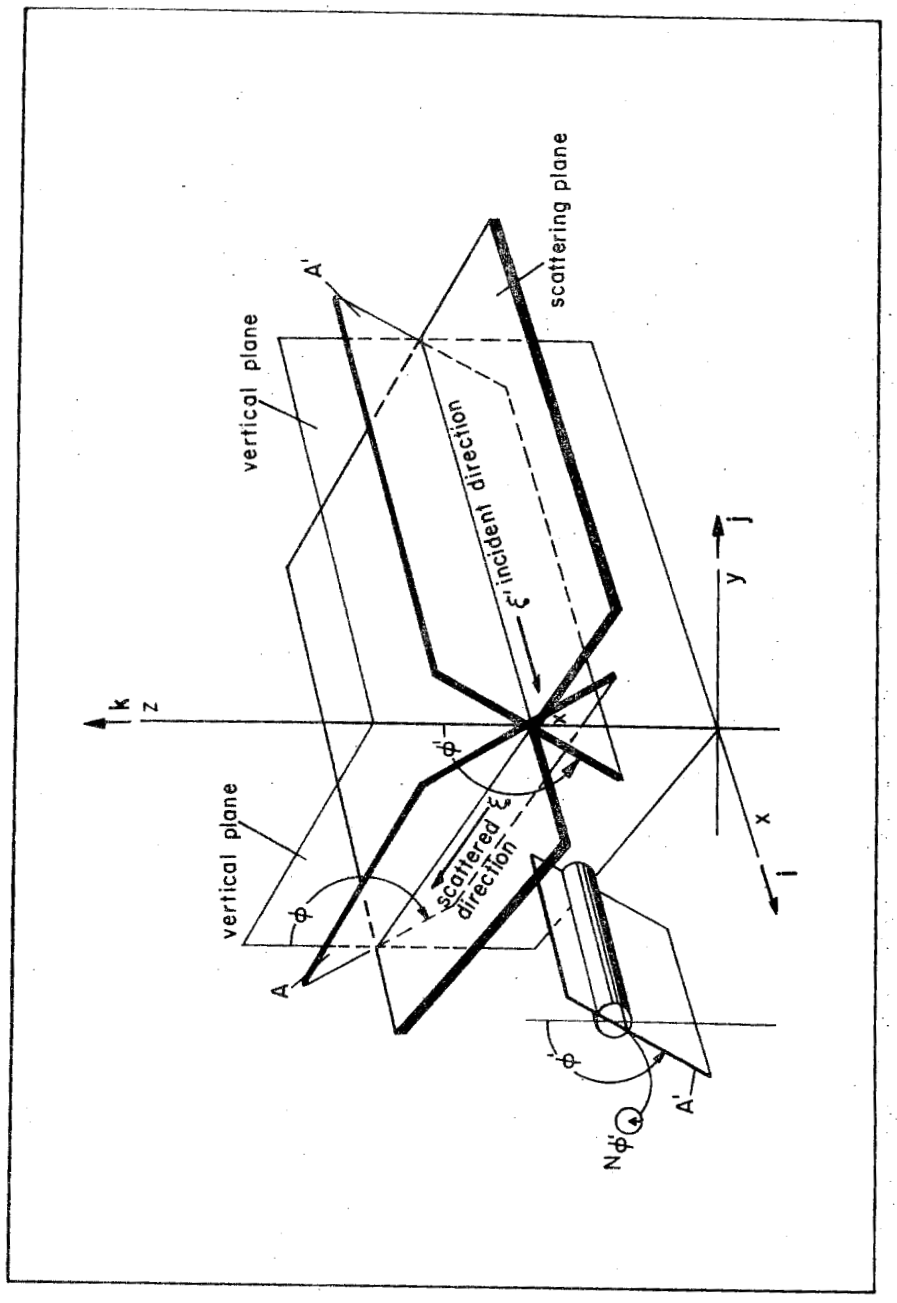


FIG. 2.27 The plane of scattering, and associated angle conventions.

"$\mathcal{L}(\phi)$" for

$$\begin{bmatrix} \cos^2\phi - \frac{1}{2}\sin 2\phi & \sin^2\phi + \frac{1}{2}\sin 2\phi & \sin^2\phi - \frac{1}{2}\sin 2\phi & 0 \\ \sin^2\phi - \frac{1}{2}\sin 2\phi & \cos^2\phi + \frac{1}{2}\sin 2\phi & \sin^2\phi + \frac{1}{2}\sin 2\phi & 0 \\ \sin 2\phi & -\sin 2\phi & \cos 2\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

It is useful to observe that $\mathcal{L}(\phi)$ has the properties:

$$\mathcal{L}(\phi_1+\phi_2) = \mathcal{L}(\phi_1)\,\mathcal{L}(\phi_2) \tag{5}$$

$$\mathcal{L}^{-1}(\phi) = \mathcal{L}(-\phi) \tag{6}$$

Thus, in particular, the inverse $\mathcal{L}^{-1}(\phi)$ of $\mathcal{L}(\phi)$ is obtained simply by replacing ϕ by $-\phi$ in $\mathcal{L}(\phi)$.

The following example will illustrate the use of (4). Fig. 2.26 (b) depicts a beam of +45° linearly polarized radiant flux proceeding along direction ξ at point x. Hence $\mathbf{N} = 1/2(\mathbf{N},\mathbf{N},2\mathbf{N},\mathbf{N})$. The reference frame is now switched from the standard reference frame at x to a local reference frame at x defined by a rotation ϕ of magnitude +45° around ξ. In order to find the components of the given beam of polarized flux in this new frame, we first note that:

$$\mathcal{L}(45°) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Hence:

$$\mathbf{N}_{45°} = \tfrac{1}{2}(\mathbf{N},\mathbf{N},2\mathbf{N},\mathbf{N}) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \tfrac{1}{2}(2\mathbf{N},0,\mathbf{N},\mathbf{N}) \quad ,$$

that is, $\mathbf{N}_{45^\circ}$ is vertically polarized radiance in the new frame of reference, as was to be expected.

The primary advantage of introducing local reference frames and their corresponding local observable radiance vectors lies in the fact that the volume scattering matrix obtained in most natural optical media can be given a simple standard form whenever such frames are used. The general relation (4) permits the resultant scattering matrix to be fitted into discussions using the standard observable radiances.

Radiant Flux Content of Polarized Radiance

Having extended the concept of radiance to the polarized context, the question now arises as to the necessary connection that exists between the readings of a radiance meter with and without polarization attachments. Specifically, let $\mathbf{N}$ be $({}_1N, {}_2N, {}_3N, {}_4N)$, the observable radiance vector (in either standard or local form) at a point and a given direction at that point. Further, let N be the simultaneous radiance reading of the meter at the same point and same direction with the polarizer P and wave plate W removed. What is the connection between $\mathbf{N}$ and N? This question, interestingly, cannot be answered within the theoretical framework of radiative transfer *per se*; of course it can be answered on the empirical level quite easily. However, to establish the desired theoretical connection one must appeal to some relatively finer-grained picture of light phenomena, such as electromagnetic theory. On such a more fundamental level both N and the components of $\mathbf{N}$ are representable in terms of the principal construct on that level: the electromagnetic wave. The desired connection can be established by suitably relating these electromagnetic representations of N and the ${}_iN$. On that level the desired connection is readily forthcoming (see, e.g., Ref. [43]) and is of the form:

$$N = {}_1N + {}_2N \quad . \tag{7}$$

This relation is interpreted as described above and under the assumption that P has ideal transmittance 1/2 for unpolarized flux and W has ideal transmittance 1.

Of course these ideals are not attained in practice. However, with (7) as a starting point, the associated practical version is readily established. The customary operational definition of the transmittance of the polarizer assembly is as follows. We write:

$$\text{"T"} \quad \text{for} \quad \max_{\psi}\left(N(\psi,0)/N\right) \tag{8}$$

Further, $N(\psi,0)$ is, as the notation implies, the radiance reading with the W setting ε equal to 0 (or W removed entirely) and with P in place and with optic axis rotated an amount ψ, in the usual way. For example, it follows from (1), in the case of linearly polarized radiance that $N(\psi,0)$ varies

ideally as $\cos^2\psi$. Hence $N(\psi,0)$ reaches its ideal maximum of $N(0,0) = N$ at $\psi = 0$. In practice, however, $N(0,0) < N$, and so $T < 1$. It is now easy to establish that the practical counterpart to (7) is:

$$N = ({}_1N + {}_2N)/T \tag{9}$$

where ${}_1N$ and ${}_2N$ are measured with the same polarizer P as that used to obtain T in (8) and with the identical disposition of W as that used for (8), i.e., either W is in the tube and $\varepsilon = 0$, or W is removed entirely.

To summarize, if a radiance tube is fitted with attachments to allow the determination of the observable radiance vector $({}_1N, {}_2N, {}_3N, {}_4N)$, then the associated reading N of the meter without these accouterments is generally related to the vector components by means of (9), with T defined as in (8). In this way the radiant flux content N of **N** is established.

On the basis of relation (9) or its suitable generalizations, it is possible to use tabulated polarized radiance data to compute all the usual unpolarized radiance, scalar irradiance and vector irradiance quantities, etc. formulated in the preceding sections, simply by replacing "N" everywhere in those formulas by "$({}_1N+{}_2N)/T$". In this sense then we understand polarized radiance data to be more general than unpolarized radiance data, for it includes the latter as a special case.

2.11 Examples Illustrating the Radiometric Concepts

In this section, we conclude our discussion of geometrical radiometry and, before going on to the discussion of photometry, we consider some examples which may serve to illustrate in some depth the various radiometric concepts and the relations among them. The contents of this section are intended to serve a multiple purpose. First of all we take the opportunity of collecting together some worked examples in geometrical radiometry which illustrate the theory developed above; secondly, various special topics of only limited interest to hydrologic optics *per se* are considered on the basis of their intrinsic radiometric merits; and finally the section serves as a repository for certain special radiometric results needed as a matter of course in the later developments of this work.

Example 1: Radiance of the Sun and Moon

We illustrate the use of the empirical radiance definition (1) of Sec. 2.5 by using it to compute the empirical field radiances of the sun and moon. Now in (1) of Sec. 2.5, S is the area collecting the flux P(S,D) funneling down the set D of directions from either the sun or the moon. Hence S may be chosen at will and we fix it in this example as a

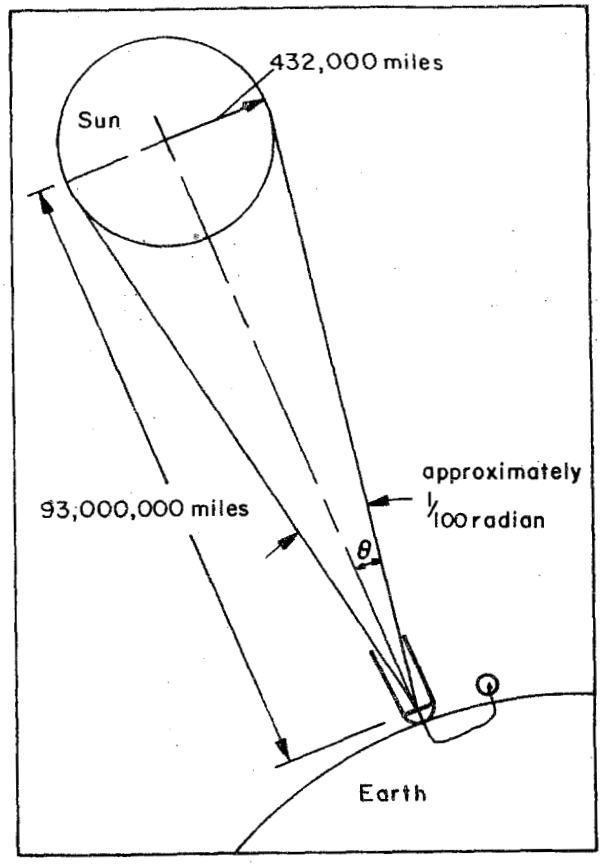


FIG. 2.28 Approximate angular subtense of the sun at the earth.

square meter of plane surface just outside the atmosphere and whose normal when extended goes through the center of the sun or moon. For the purpose of computing N(S,D) we choose D to be the solid conical set of directions from any point on the collecting surface to and within the limb of the sun or moon. See Fig. 2.28. We consider first the case of the sun.

The sun is a nearly spherical body with diameter nearly 864,000 miles and at a distance of about 93,000,000 miles from the earth. It follows that the half-angle subtense θ of the sun at the earth's surface is very nearly:

$$\theta = 4.32 \times 10^5 / 9.3 \times 10^7$$

$$= 4.65 \times 10^{-3} \text{ radians}$$

Hence, by (12) of Sec. 2.5, the solid angle subtense $\Omega(D)$ of the sun is:

$$\Omega(D) = \pi \times (4.65)^2 \times 10^{-6}$$

$$= 6.78 \times 10^{-5} \text{ steradians}$$

Now Table 1 of Sec. 2.4 gives an order of magnitude estimate of 10^3 watts/m^2 of the irradiance $P(S,D)/A(S)$ produced by the sun's radiation over the whole spectrum and which is incident through D on a surface S normal to the sun's rays. A more accurate estimate of this irradiance produced outside the atmosphere is 1396 watts/m^2; see Ref. [128]. A still more meaningful alternate estimate of $H(S,D)$ can be made for the interval of wavelengths in the visible spectrum (approximately 400 to 700 millimicrons). In this case $H(S,D)$-estimates vary from 542 watts/m^2 to 555 watts/m^2 (p. 31, Ref. [185]), see also Ref. [128]). Using the first estimate we obtain*:

$$N(S,D) = H(S,D)/\Omega(D)$$

$$= 5.42 \times 10^2/6.78 \times 10^{-5}$$

$$= 8 \times 10^2 \text{ watts/m}^2 \times \text{steradian} \quad (1)$$

This radiance is the overall average radiance of the sun's disk as seen just outside the atmosphere and over the wavelengths of the visible spectrum. (Hence the set F of frequencies of Sec. 2.3 now consists of all frequencies from approximately 4×10^{14} to 7×10^{14}/sec.)

A good rule of thumb for remembering the angular subtense of the sun is that its entire disk subtends an angle of about 1/100 of a radian. The more exact estimate is given above. In other words the sun subtends about the same angle as a disk of a centimeter diameter at a meter's distance.

We turn now to the case of the moon. The geometric and radiometric principles are the same as in the case of the sun. And again, the crucial point of the calculation rests in the estimate of $H(S,D)$. For this case we assume that the irradiance $H(S,D)$ of the sun is on the order of 7×10^5 times that of the full moon over the visible spectrum. (See, e.g., Fig. 1.12 and Table 2 of Sec. 2.12.) In other words we assume that for the case of the moon, $H(S,D) = 7.75 \times 10^{-9}$ watts/m^2. Estimates of this ratio vary considerably. The one just chosen is an order of magnitude estimate only for the purposes of the present example.

The moon is a nearly spherical body with diameter nearly 2100 miles and at a distance of about 240,000 miles. It follows that the half-angle subtense θ of the moon at the earth's surface is very nearly:

$$\theta = (1.05 \times 10^3)/(2.4 \times 10^5)$$

$$= 4.38 \times 10^{-3} \text{ radians.}$$

---

*At sea level under a clean dry atmosphere, $H(S,D)$ on the order of 472 watts/m^2. See also Table 2, Sec. 1.2.

Hence, by (12) of Sec. 2.5, the solid angle subtense $\Omega(D)$ of the moon is:

$$\Omega(D) = \pi(4.38)^2 \times 10^{-6}$$

$$= 6.00 \times 10^{-5} \text{ steradians} .$$

Using the adopted estimate we obtain:

$$N(S,D) = H(S,D)/\Omega(D)$$

$$= 7.75 \times 10^{-4}/6.00 \times 10^{-5}$$

$$= 13 \text{ watts/m}^2 \times \text{steradian} \qquad (2)$$

This may be used as an overall average radiance of the full moon's disk as seen at sea level on a clear night and over the wavelengths of the visible spectrum. An extensive literature exists with reference to lunar photometry and radiometry. See, e.g., [8].

In conclusion we note that the rule of thumb adopted for the angular size of the sun as seen from earth evidently also holds for the moon. For more detailed radiometric information on the radiant energy output of the sun, the reader may consult, e.g., Sec. 1.1 and Refs. [185] and [128]. Detailed discussion is made of the estimates of the solar irradiances in the latter references.

Example 2: Radiant Intensity of the Sun and Moon

The present example illustrates the use of the concept of radiant intensity as defined in (1) of Sec. 2.9.

We begin by computing the radiant intensity of the hemisphere S of the sun visible from the earth. Let ξ be the unit vector pointing from the center of the sun to the center of the earth. Then the radiant intensity $J(S,\xi)$ of S in the direction ξ is given by (14) of Sec. 2.9, where $N(x,\xi)$ is the surface radiance of the sun in the direction ξ at a point x on S. In Example 1 we estimated the field radiance of S for radiant flux in the wavelength interval from 400 to 700 millimicrons. Now, by the radiance invariance law (2) of Sec. 2.6, the estimate of Example 1 may be taken as the surface radiance of the sun over S, the radiance $N(x,\xi)$ being sensibly independent of x on S. Then if "N" denotes this fixed surface radiance, (14) of Sec. 2.9 yields:

$$J(S,\xi) = N\int_S \xi\cdot\xi'(x)\, dA(x)$$

$$= NA(S')$$

where A(S') is the area of the projection S' of S on a plane perpendicular to ξ. The area A(S') is readily determinable. From the data in Example 1, we have:

$$A(S') = \pi(4.32)^2 \times 10^{10} (\text{miles})^2$$

$$= \pi(4.32)^2 \times (1.6)^2 \times 10^6 \times 10^{10} \ (\text{meters})^2$$

$$= 1.5 \times 10^{18} \ (\text{meters})^2$$

Using this estimate of A(S') and the estimate of N(S,D) for the sun given in Example 1, we have:

$$J(S,\xi) = 8 \times 10^6 \times 1.5 \times 10^{18}$$

$$= 1.2 \times 10^{25} \text{ watts/steradian} \qquad (3)$$

as the radiant intensity of a hemisphere of the sun facing the earth and over the visible spectrum.

The sun is radiometrically a point source (Sec. 2.9) with respect to points on the earth and may thus be imagined to be compressed to its center x. Furthermore, we may evidently assume that $J(S,\xi)$ is independent of ξ. Hence (17) of Sec. 2.9 is applicable, and we can estimate the total radiant flux output of the sun over the visible spectrum to be:

$$P(x) = 4\pi J(S,\xi)$$

$$= 1.5 \times 10^{26} \text{ watts.}$$

Turning now to the case of the moon, we have a slightly more interesting geometrical situation arising from the possible phases of the moon. Fig. 2.29 depicts this situation. If "S'" now denotes the projection of a lunar hemisphere on a plane normal to the direction ξ, then we have by means of (14) of Sec. 2.9:

$$J(S,\xi) = \frac{NA(S')}{2} (1 + \cos \theta)$$

where N is the surface radiance of the lighted hemisphere of the moon, as estimated*, e.g., in Example 1, and θ is the phase angle of the moon as described in Fig. 2.29. Thus at full moon, $\theta = 0$ and $J(S,\xi)$ is in particular NA(S). To estimate this product we first compute:

$$A(S') = \pi(1.05)^2 \times 10^6 \ (\text{miles})^2$$

$$= \pi(1.05)^2 \times (1.6)^2 \times 10^5 \times 10^6 \ (\text{meters})^2$$

$$= 8.9 \times 10^{12} \ (\text{meters})^2 \quad .$$

---

*The precise analysis of the gradation of the radiance distribution over the sunlit hemisphere of the moon is a delicate problem. The estimate here is deliberately kept simple in order to first emphasize the radiometric geometry essentials. A source reference on radiometry of the moon and planets is [8].

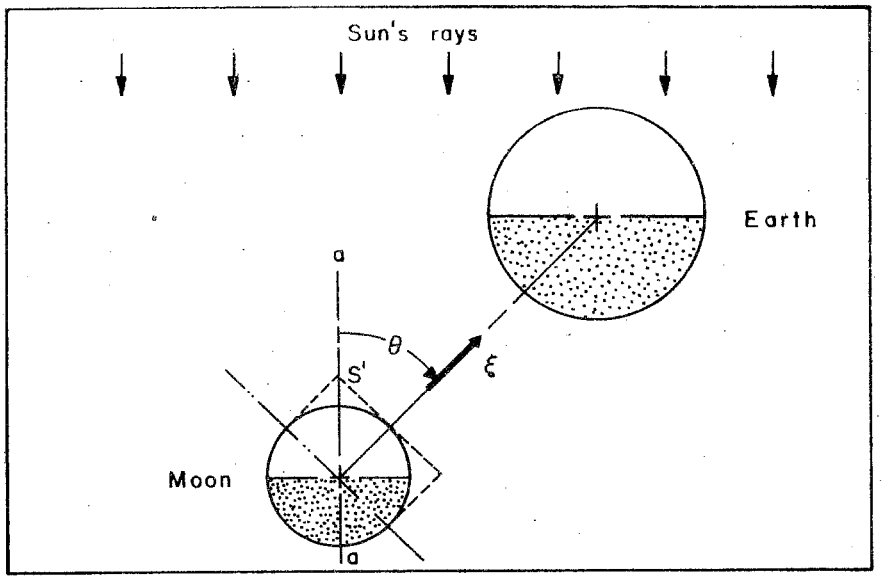


FIG. 2.29 Simple phase diagram for the earth-moon system

Using 13 watts/(m^2 × steradian) for N (justified by means of the radiance invariance law) we have:

$$NA(S') = 1.3 \times 10 \times 8.9 \times 10^{12} = 1.2 \times 10^{14} \text{ watts/steradian}$$

as the radiant intensity of the surface of the full moon over the visible spectrum. Hence for any phase θ, the corresponding radiant intensity of the lighted surface S of the moon in direction ξ (Fig. 2.29) is:

$$J(S,\xi) = 0.6 \times 10^{14}(1 + \cos\theta) \text{ watts/steradian} \qquad (4)$$

We conclude this example by computing the total radiant flux content of the reflected radiant flux from the moon, over the visible spectrum. Using the radiant intensity estimate just made, and assuming N to be independent of direction S, and the moon to be a point source at its center x as seen from the earth, we then integrate J(x,ξ) over all directions to obtain the requisite radiant flux, according to (15) of Sec. 2.9. Thus if "x" denotes the center of the moon and D' is now Ξ, Equation (15) of Sec. 2.9 becomes:

$$P(x) = \int_{\Xi} J(x,\xi)\, d\Omega(\xi)$$

$$= \frac{NA(S')}{2} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} (1+\cos\theta) \sin\theta \, d\theta \, d\phi = 2\pi NA(S)$$

To perform this integration, a "lunar based" polar coordinate system was used with θ measured as shown in Fig. 2.29 and ϕ measured from 0 to 2π around the axis a -a in a plane perpendicular to the page. We might have expected this relation on intuitive grounds: the radiant flux output of an entire sphere lighted uniformly all over, should be just twice that given off by one hemisphere. Thus P(x) is 2π times NA(S) instead of 4π times NA(S'). Hence:

$$P(x) = 2\pi \times 1.2 \times 10^{14}$$

$$= 7.5 \times 10^{14} \text{ watts}$$

is the radiant flux output of the moon over the visible spectrum.

Example 3: Radiant Flux Incident on Portions of the Earth

In this example, Equations (7) of Sec. 2.4 and (8) of Sec. 2.5 will be illustrated. Now, from Example 1, we find that at each point x just outside the atmosphere of the earth we have $H(x,D,t,F) = 542$ watts/m^2 funneling down a narrow cone D from the disc of the sun and with wavelengths over the visible spectrum F. Suppose S is some portion of the earth's surface accessible to the sun's rays, as in Fig. 2.30. To compute $\Phi(S,D,t,F)$, we establish a polar coordinate system as depicted in the figure. We first deduce that:

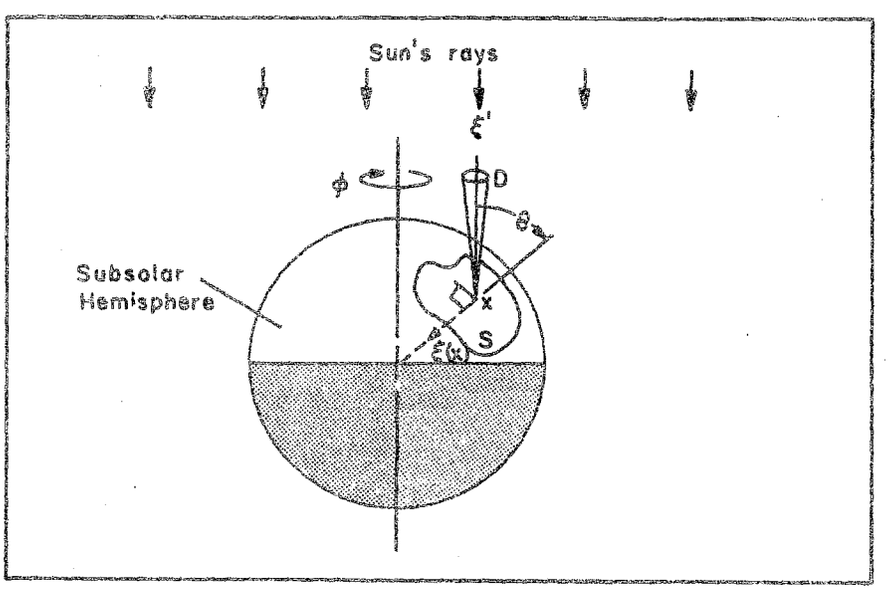


FIG. 2.30 Geometry for solar irradiation calculations

$$H(x,D,t,F) = \int_F H(x,D,t,\nu)\, dl(\nu) \qquad (5)$$

which follows from (5) of Sec. 2.4 by means of a theorem of elementary calculus. Then by (7) of Sec. 2.4 we have:

$$\Phi(S,D,t,F) = \int_S H(x,D,t,F)\, dA(x)$$

Next, from (8) of Sec. 2.5:

$$H(x,D,t,F) = \int_D N(x,\xi',t,F)\xi\cdot\xi'\, d\Omega(\xi')$$

where we now explicitly use the fact that wavelengths are over the visible spectrum F. Since D is small and the sun's field radiance is uniform of magnitude N over D we can estimate H(x,D,t,F) fairly accurately by means of the equality:

$$H(x,D,t,F) = N\xi'\cdot\xi(x)\Omega(D) \quad ,$$

where N and $\Omega(D)$ were estimated for the sun in Example 1. Furthermore, $\xi(x)$ is the unit inward normal to the earth's surface at x, and ξ' is the direction from the center of the sun to the center of the earth. Using this representation of H(x,D,t,F) in the preceding integral for $\Phi(S,D,t,F)$, we arrive at the expression:

$$\Phi(S,D,t,F) = N\Omega(D) \int_S \xi'\cdot\xi(x)\, dA(x)$$

$$= NA(S')\Omega(D) \qquad (6)$$

where A(S') is the area of the projection S' of S on a plane normal to the direction ξ' of the sun's rays.

As a specific example, we use N and $\Omega(D)$ as in Example 1, and let S be the sub solar hemisphere of the earth. Then:

$$A(S') = \pi(4)^2 \times 10^6 \text{ (miles)}^2$$

$$= \pi(4)^2 \times (1.6)^2 \times 10^6 \times 10^6$$

$$= 1.3 \times 10^{14} \text{ (meters)}^2$$

Hence:

$$\Phi(S,D,t,F) = 8 \times 10^6 \times 1.3 \times 10^{14} \times 6.78 \times 10^{-5}$$

$$= 7 \times 10^{16} \text{ watts} \qquad (7)$$

over the visible spectrum. The corresponding radiant flux $\Phi(S_1,D,t,F)$ incident on any proper portion S_1 of the entire subsolar hemisphere is simply obtained by finding $A(S_1')/A(S')$, where now S_1' is the projection of S_1 on a plane normal to the sun's rays, and then multiplying 7×10^{16} by this fraction; or alternatively, 542 watts/m^2 by $A(S_1')$. Of course these estimates are somewhat crude, and serve only to illustrate the correct mathematical use of the geometric radiometry formulas deduced above. The present estimate of $\Phi(S,D,t,F)$ omits, e.g., the effect of the atmosphere which at each point subtly attenuates and augments the solar influx by permitting absorption, scattering, and interreflections with the earth below.

Example 4: Irradiance Distance-Law for Spheres

In this and several of the examples below we shall explore some interesting consequences of the irradiance integral (8) of Sec. 2.5.

We begin the investigations by considering a spherical surface S of radius a with uniform radiance distribution of magnitude N at each point. Suppose that S is viewed at a point x a distance r from the center y of S. The lines of sight lie in a vacuum and the background radiance of S is zero. See Fig. 2.31 (a). We ask: what is the irradiance $H(x,\xi)$ at point x? Here ξ is the direction from y to x.

Equation (8) of Sec. 2.5 is readily applied to the present situation. For the present case we may use the radiance invariance law to say that $N(x,\xi') = N$ for every ξ' in the conical set D of directions subtended by S at x. Hence (8) of Sec. 2.5 becomes:

$$\begin{aligned} H(x,\xi) &= N\int_D \xi\cdot\xi'\, d\Omega(\xi') \\ &= N\int_0^{2\pi}\int_0^{\theta}\cos\theta'\sin\theta'\, d\theta'\, d\phi' \\ &= 2\pi N\int_0^{\theta}\cos\theta'\sin\theta'\, d\theta' \\ &= \pi N\sin^2\theta \\ &= \pi N(a/r)^2 \quad . \end{aligned}$$

If we write, *ad hoc*,

$$\text{"}H_r\text{"} \quad \text{for} \quad H(x,\xi) \quad ,$$

then we have found that:

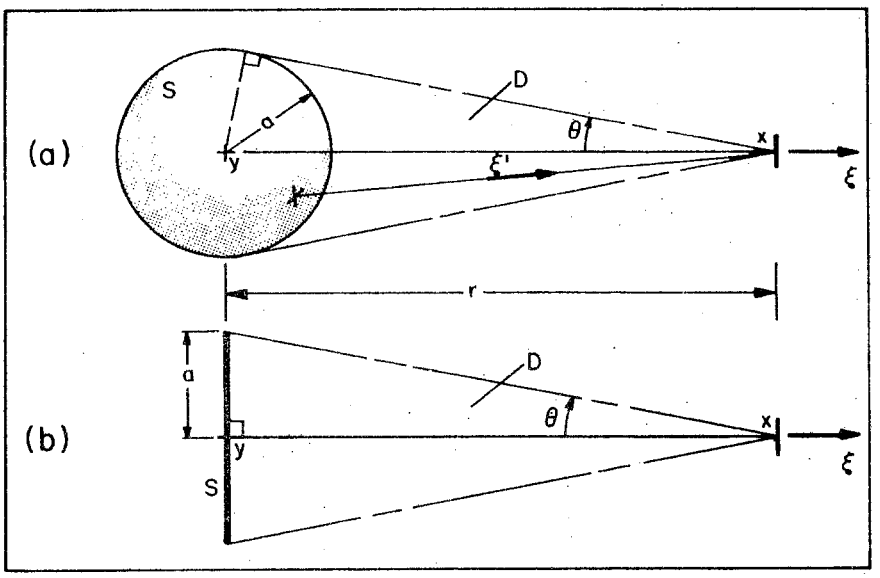


FIG. 2.31 Deriving the Irradiance Distance-Law for spheres and disks

$$\boxed{H_r = AN/r^2} \quad , \tag{8}$$

where we have written:

$$\text{"A"} \quad \text{for} \quad \pi a^2 \quad ,$$

i.e., A is the area of a great circle of S; alternatively A is the area of projection of S on a plane perpendicular to ξ. From (14) of Sec. 2.9 applied to the present case, we may write:

$$\boxed{H_r = J/r^2} \tag{9}$$

where we have written:

$$\text{"J"} \quad \text{for} \quad AN \quad .$$

It is to be particularly noted that H_r varies *precisely* as the inverse square of the distance r, where $a \leq r$. If $r = a$, then:

$$H_a = \pi N \quad . \tag{10}$$

Example 5: Irradiance Distance-Law for Circular Disks; Criterion for a Point Source

Equation (8) of Sec. 2.5 will now be used to derive the law governing the irradiance produced by a circular disk S of uniform radiance. See Fig. 2.31 (b). In that figure is depicted a circular disk of radius a and of uniform surface radiance N at each point. The disk is viewed at point x on the perpendicular through the center y of S at a distance r from the center. The set D of the lines of sight from x to S lies in a vacuum and the background radiance of S is zero. What is the radiance $H(x,\xi)$ at point x? Here ξ is the direction from y to x.

Equation (8) of Sec. 2.5 can be applied to the present situation, as in the case of Example 4. Thus, (8) of Sec. 2.5 becomes:

$$H(x,\xi) = N\int_D \xi\cdot\xi'\, d\Omega(\xi')$$

$$= N\int_0^{2\pi}\int_0^{\theta} \cos\theta' \sin\theta'\, d\theta'\, d\phi'$$

$$= 2\pi N\int_0^{\theta} \cos\theta' \sin\theta'\, d\theta'$$

$$= \pi N \sin^2\theta$$

$$= \pi N\, a^2/(a^2+r^2) \quad .$$

If we write, *ad hoc*:

$$\text{"}H_r'\text{"} \quad \text{for} \quad H(x,\xi)$$

and further, we write:

$$\text{"}A\text{"} \quad \text{for} \quad \pi a^2$$

and:

$$\text{"}J\text{"} \quad \text{for} \quad AN \quad ,$$

then we have found that:

$$H_r' = AN/(a^2+r^2) \tag{11}$$

or:

$$H_r' = J/(a^2+r^2) \tag{12}$$

From this we find first of all that H_r', unlike H_r of Example 4, does not vary precisely as the inverse square of r, where $r \geq 0$. However, in the special case of $r = 0$, we have:

$$H_o' = \pi N \quad . \tag{13}$$

Further, in the other extreme, i.e., when r is very much larger than a, H_r' varies very nearly as the inverse square of r.

By examining more closely the difference between H_r and H_r', we arrive at the basis for the definition of a point source given in Sec. 2.9. Suppose then we compare H_r and H_r' which are, respectively, the irradiances produced by a sphere of radius a and a circular disk of radius a both of uniform radiance N. Toward this end we form the difference:

$$H_r - H_r' = J\left[\frac{1}{r^2} - \frac{1}{(a^2+r^2)}\right]$$

and then form the relative difference:

$$(H_r - H_r')/H_r' = \frac{a^2+r^2}{r^2} - 1 = a^2/r^2$$

This relative-difference expression is the basis for the following statements: *The relative difference between the irradiance* H_r *and* H_r' *is less than 1% whenever* $r > 10a$. *More generally: the irradiance produced by a finite object of uniform radiance decreases as the inverse square of the distance from that object, within an error of 1 percent, whenever the distance from the object is more than 10 times greater than the object's largest transverse linear dimension.* This alternate statement follows readily from the preceding analysis. Some further study is made in Example 6 of related questions. Observe that the associated solid angle of the circular cone of half angle 1/10 radian is very nearly $\pi(1/10)^2 = \pi/100 \approx 1/30$ steradian, in which lies the origin of the solid angle number used in the point source criterion of Sec. 2.9.

Example 6: Irradiance Distance-Law for General Surfaces

We devote this example to the elucidation of the common denominator of Examples 4 and 5; the net result being the formula for the irradiance distance-law for a general surface S of uniform radiance N viewed, as in Fig. 2.32, from an external vantage point x along a set of paths defined by a collection D of directions, each path of which lies in a vacuum.

The derivation of the required $H(x,\xi)$ begins, as in Examples 4 and 5, with (8) of Sec. 2.5, but now proceeds as follows:

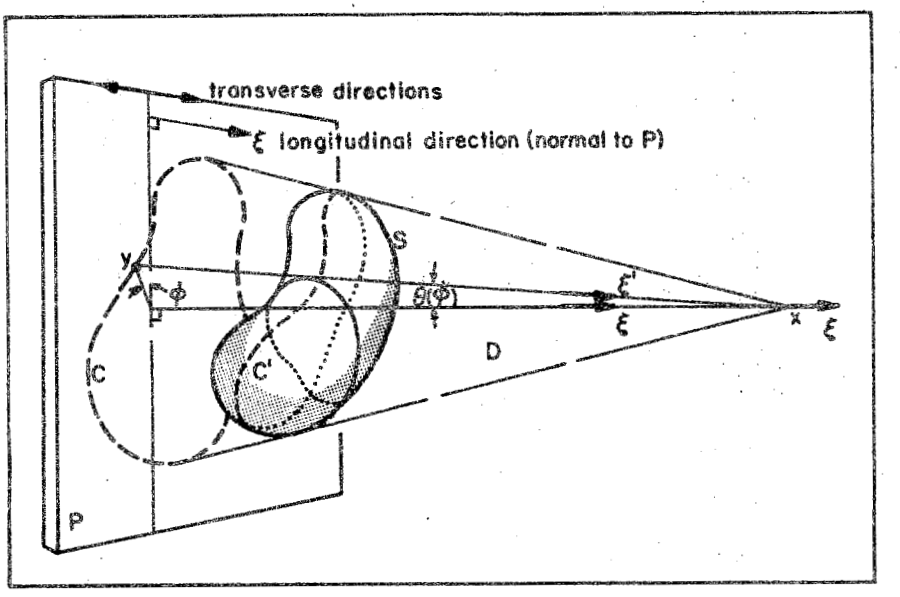


FIG. 2.32 Deriving the Irradiance Distance-Law for general surfaces

$$
\begin{aligned}
H(x,\xi) &= N\int_D \xi\cdot\xi'\, d\Omega(\xi') \\
&= N\int_0^{2\pi}\int_0^{\theta(\phi)} \cos\theta' \sin\theta'\, d\theta'\, d\phi \\
&= \frac{N}{2}\int_0^{2\pi}\int_0^{\theta(\phi)} d(\sin\theta')\, d\phi \\
&= \frac{N}{2}\int_0^{2\pi} \sin^2\theta(\phi)\, d\phi
\end{aligned}
$$

Let us write, *ad hoc*:

$$\text{"H"} \quad \text{for} \quad H(x,\xi) \quad .$$

With this, we have attained the required result:

$$\boxed{H = \frac{N}{2}\int_0^{2\pi} \sin^2\theta(\phi)\, d\phi \quad .} \tag{14}$$

This formula for H reduces to the expressions for H_r and H_r' when the function $\theta(\cdot)$ is suitably prescribed for all ϕ from 0 to 2π. In particular θ is a constant function in the preceding two cases. More importantly, the reader should observe the remarkable fact that the irradiance H *depends only on the integral over the outline* C *of* S, as may be seen by studying the central projection of S onto the background plane P (of Fig. 2.32) which is perpendicular to ξ. Hence it is literally immaterial to H what the longitudinal structure of S is as regards the computation of H at a fixed point x, as long as S has the given outline C on P, and also has uniform radiance N. Of course the shape of S is important when it is decided to let x *vary*, and indeed the distance-law for $H(x,\xi)$ depends critically on the longitudinal shape of S and in this context takes its most general form displayed in the above equation for H.

An alternate form of the distance-law for irradiance is obtained when we write:

$$\text{"}\Omega\text{"} \quad \text{for} \quad \int_D \xi' \, d\Omega(\xi')$$

Hence:

$$\boxed{H = N\xi\cdot\Omega} \quad . \tag{15}$$

When the size Ω of D is small--e.g., when S is a point source at x, then we have, very nearly:

$$\Omega = \xi\Omega$$

and in this special case (15) yields:

$$H = N\Omega \; . \tag{16}$$

If A is the projected area of S then in this case we have very nearly:

$$\Omega = A/r^2 \quad ,$$

where r is the distance from x to S. In this way we return to the inverse square law for H in the limit of large r (or small A).

Still one more form for H, i.e., $H(x,\xi)$, is obtainable using the concept of vector irradiance introduced in Sec.2.8. Thus we have

$$\boxed{H = \xi\cdot\mathbf{H}} \tag{17}$$

where in the present case we have written:

$$\text{"}\mathbb{H}\text{"} \quad \text{for} \quad \int_D N\xi' \, d\Omega(\xi')$$

As a corollary we have:

$$\boxed{H = N\Omega} \tag{18}$$

An important and useful special case of (14) occurs when $\theta(\phi)$ is independent of ϕ. This happens when the surface S is a surface of revolution about the direction ξ. (See Fig. 2.32). In such a case (14) becomes:

$$\boxed{H = \pi N \sin^2 \theta} \quad .$$

In particular, if S is an infinite plane, then at all distances r from S, S subtends a half angle $\theta = \pi/2$. Hence $H = \pi N$ for all r in such a case.

Example 7: Irradiance via Line Integrals

The present example is designed to let us investigate in greater depth the irradiance integral (14) of Example 6 which showed that the irradiance produced at a fixed point x by an arbitrary surface of uniform radiance depended only on the *angular* outline of S as seen at the point x. Our goal in this example will be to cast equation (14) into explicit line integral form over the curve C which defines the outline of S.

Figure 2.33 (a) is a reconstruction of Figure 2.32 with surface S omitted. What is left is the geometric essence of the irradiance calculations done in Example 6. Specifically, we have retained the central projection of S on plane P through point x. The boundary C of this projection of S on P is a closed curve characterized by means of the function $\theta(\cdot)$ which assigns to each ϕ, $0 \leq \phi \leq 2\pi$ an angle $\theta(\phi)$, which determines point y on C as shown in Fig. 2.33. We denote by "O" the foot of the perpendicular dropped on P from x. Further, "$r(\phi)$" will denote the distance from the fixed point x to the variable point y on C.

With these preliminaries established, we can write (14) in the form:

$$H = \frac{N}{2}\int_0^{2\pi} \frac{\sin\theta(\phi)}{r(\phi)} \cdot r(\phi) \sin\theta(\phi) \, d\phi \quad .$$

The integral was rewritten this way to make use of the fact that:

$$r(\phi) \sin\theta(\phi) \, d\phi$$

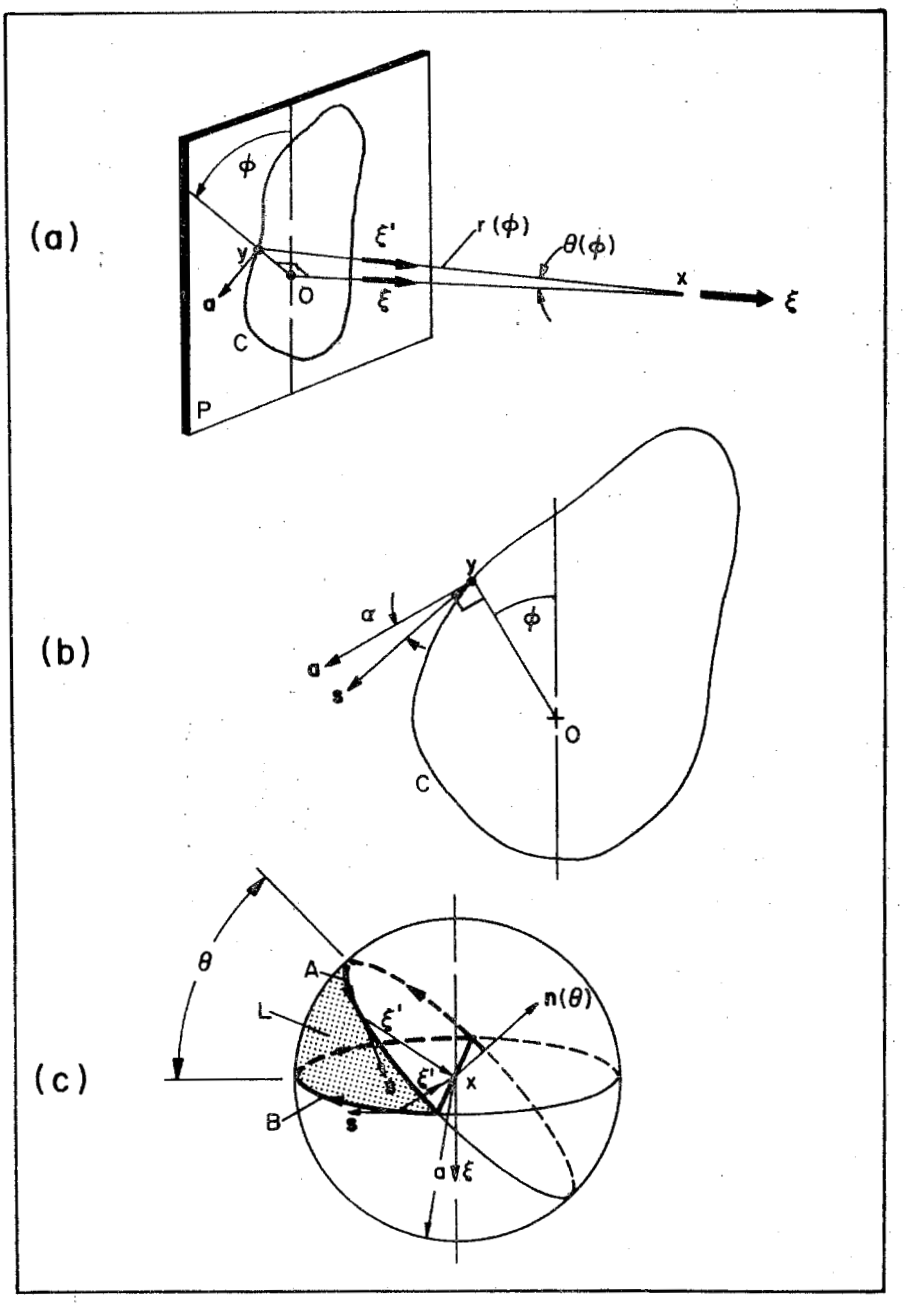


FIG. 2.33 Setting for calculating irradiances via line integrals.

is an element of length da along the direction of the unit vector **a** in P, normal to Oy, and at point y on C, as shown in Fig. 2.33 (b), which is a plan view of P. The element of length da is related to an element of arc length ds along C at y by a projection, and by definition of ds:

$$da = \cos\alpha \, ds$$

where α is the angle between **a** and the unit tangent vector **s** to C at y; hence:

$$r(\phi)\sin\theta(\phi)\, d\phi = da = \mathbf{a}\cdot\mathbf{s}\, ds \quad .$$

The preceding integral then may be written:

$$H = \frac{N}{2}\int_C \frac{\sin\theta(\phi)}{r(\phi)}\, \mathbf{a}\cdot\mathbf{s}\, ds \quad .$$

Next we observe, by means of Fig. 2.33, that:

$$\xi' \times \xi = a \sin\theta(\phi) \quad ,$$

where ξ is the unit outward normal to P (outward relative to ξ', i.e., such that $\xi'\cdot\xi > 0$). Hence:

$$H = \frac{N}{2}\int_C \frac{(\xi'\times\xi)\cdot\mathbf{s}\, ds}{r(\phi)}$$

The triple box product of vectors in the integrand may be rearranged so that we obtain for C (or C'!) in Fig. 2.32:

$$\boxed{H = \frac{1}{2} N\xi\cdot\int_C \frac{\mathbf{s}\times\xi'\, ds}{r(\phi)}} \tag{19}$$

Comparing this with (15) we deduce that:

$$\Omega = \frac{1}{2}\int_C \frac{\mathbf{s}\times\xi'\, ds}{r(\phi)} \tag{20}$$

Equation (20) displays a line integral representation of Ω, and (19) displays the desired line integral representation of H.

As an illustration of (20), let C' be the boundary of a spherical lune L of angular opening θ, on a sphere of radius a, as shown in Fig. 2.33 (c). Thus L is now a specific instance of the general surface S of Fig. 2.32, and C may actually be taken as *any* outline of S (as, e.g., C' in Fig. 2.32). Note the present placement of point x and the direction ξ. The contribution to Ω over the upper arc A of C is clearly:

$$\frac{1}{2}\int_A \frac{\mathbf{s}\times\xi'}{r(\phi)}\,ds = \frac{\pi}{2}\,\mathbf{n}(\theta)$$

where $\mathbf{n}(\theta)$ is the unit normal to the plane containing arc A, and directed such that:

$$\mathbf{n}(\theta) = \mathbf{s}\times\xi' \quad ,$$

when A is traversed as shown in the figure. Further, $r(\phi) = a$ for all ϕ. The contribution to Ω over the arc B of C is clearly:

$$\frac{1}{2}\int_B \frac{\mathbf{s}\times\xi'}{r(\phi)}\,ds = \frac{\pi}{2}\,\xi$$

The integrals over A and B were evaluated immediately by noting that over A, $\mathbf{s}\times\xi'$ is a fixed unit vector, namely $\mathbf{n}(\theta)$; and over B, $\mathbf{s}\times\xi'$ is the unit vector ξ, the unit inward normal to the plane of arc B. The arc lengths of A and B are each $a\pi$. Hence for the present case:

$$\Omega = \frac{\pi}{2}\,(\xi + \mathbf{n}(\theta)) \quad .$$

Observe that if $\theta = 0$, then, $\mathbf{n}(\theta) = -\xi$, and $\Omega = 0$. If $\theta = \pi$, then $\mathbf{n}(\pi) = \xi$, and $\Omega = \pi\xi$. If L is of uniform radiance N, then, by (15) or (19):

$$H = N\,\xi\cdot\Omega = \frac{N\pi}{2}\,(1 + \xi\cdot\mathbf{n}(\theta)) \quad .$$

Example 8: Solid Angle Subtense of Surfaces

The integral form of the solid angle subtense $\Omega(D)$ of a set D of directions, as given in (10) of Sec. 2.5, will now be recast into a form which arises when the solid angle subtense of specific surfaces (either real or hypothetical) are under consideration. Thus, consider the surface S depicted in Fig. 2.34 (a) where S is shown viewed from an external vantage point x. Let "D(S,x)" denote the set of all directions from points of S to x. Our present goal is to derive the expression for $\Omega(D(S,x))$ (or "$\Omega(S,x)$" for short) in the form of a surface integral over S.

We begin by letting "D" in (10) of Sec. 2.5 be replaced by "D(S,x)". The result is:

$$\Omega(S,x) = \int_{D(S,x)} \sin\theta\,d\theta\,d\phi$$

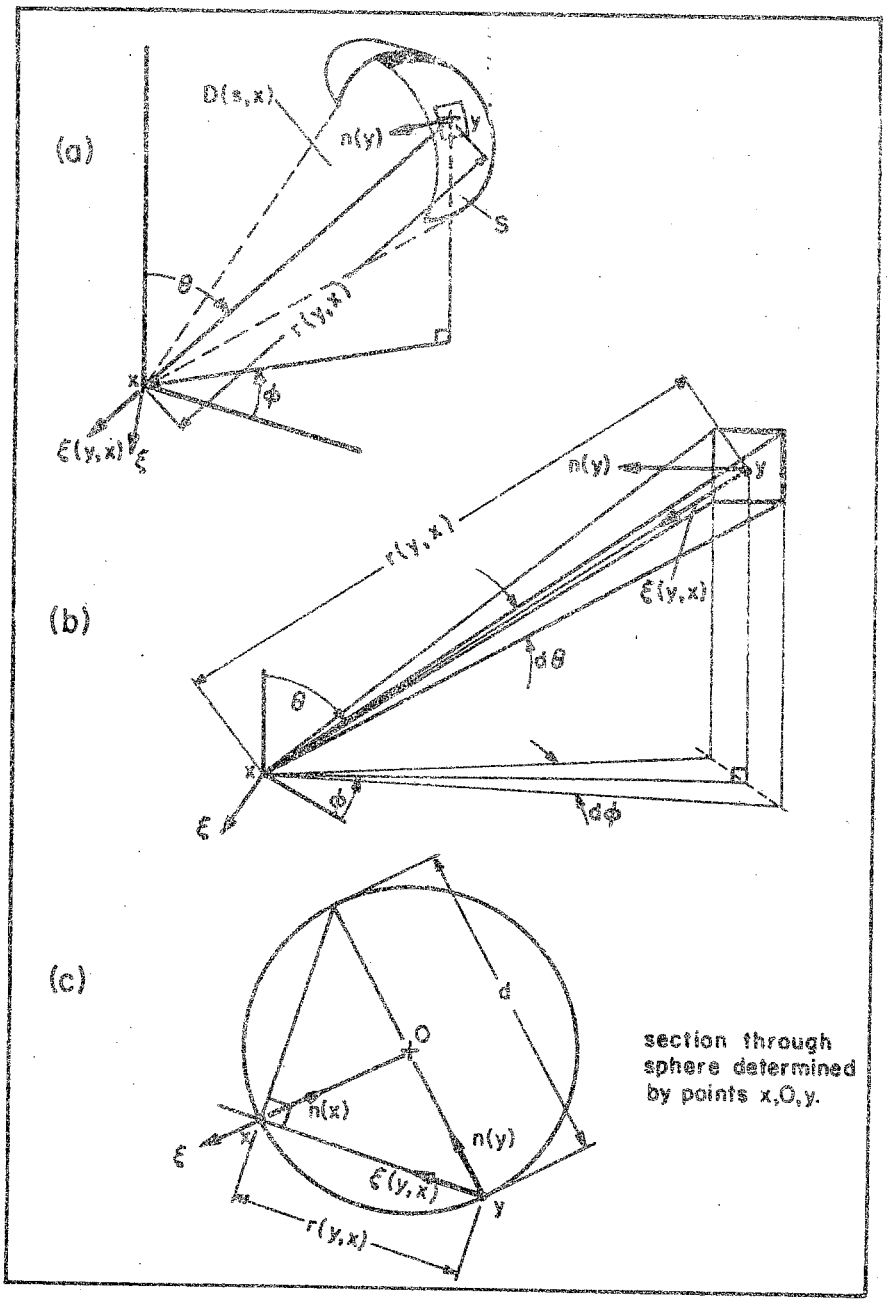


FIG. 2.34 Calculating the solid angle subtense of "tangible" surfaces.

The conventions for measuring θ and ϕ are summarized in Fig. 2.34 (a). In particular the details of the integration over a part of S about a point y on S are depicted in part (b) of Fig. 2.34. Points y and x determine a direction $\xi(y,x)$, as shown. It may be seen from part (b) of Fig. 2.34, that the relation between a small patch of S of area A(y) about y is related to its projection's area A' on a plane perpendicular to $\xi(y,x)$ by the formula:

$$A'(y) = A(y)\mathbf{n}(y)\cdot\xi(y,x) \quad (\geq 0, \text{ by choice of } S)$$

where $\mathbf{n}(y)$ is the unit outward normal to S at y. Hence the solid angle subtense of the patch of S about y is:

$$\frac{A'(y)}{r^2(y,x)} = \frac{\mathbf{n}(y)\cdot\xi(y,x)A(y)}{r^2(y,x)} \quad .$$

The entire solid angle subtense of S at x is obtained by adding up all these solid angle subtenses of the component patches comprising S:

$$\Omega(S,x) = \int_S \frac{\mathbf{n}(y)\cdot\xi(y,x)}{r^2(y,x)}\, dA(y) \tag{21}$$

It is of interest to observe that the set function $\Omega(\cdot,x)$ is non-negative valued, S-additive and S-continuous (compare these properties with those of the radiometric concepts in Sec. 2.3). Thus for every x and pair S_1, S_2 surfaces with disjoint sets $D(x,S_1)$ and $D(S_2,x)$ we have:

$$\Omega(S_1,x) + \Omega(S_2,x) = \Omega(S_1 \cup S_2, x)$$

which is the S-*additivity property*; further:

$$\text{If } A(S) = 0, \quad \text{then } \Omega(S,x) = 0 \quad .$$

In other words, the latter statement, the S-*continuity property* for $\Omega(\cdot,x)$, asserts that $\Omega(S,x) > 0$ only if $A(S) > 0$.* It follows from these additivity and continuity properties of Ω and the calculus that the ratio $\Omega(S,x)/A(S)$ has a limit as $S \to \{y\}$, where y is some point of S. Indeed:

$$\lim_{S\to\{y\}} \Omega(S,x)/A(S) = \frac{\mathbf{n}(y)\cdot\xi(y,x)}{r^2(y,x)} \quad .$$

*The converse clearly does not hold; thus, give a counterexample for: If $\Omega(S,x) = 0$, then $A(S) = 0$.

Writing "$d\Omega(y,x)/dA(y)$" for the limit operation above, we can then state that:

$$\frac{d\Omega(y,x)}{dA(y)} = \frac{n(y)\cdot\xi(y,x)}{r^2(y,x)} \tag{22}$$

Equation (22) yields, for a given point x, the value of the general area derivative of the solid angle function $\Omega(\cdot,x)$ at point y of an arbitrary surface, where $n(y)$ is the unit outward normal to the surface at y, $r(y,x)$ is the distance from x to y and $\xi(y,x)$ is the unit vector from y to x and such that the dot product of $\xi(y,x)$ and $n(y)$ is non-negative (this fixes the sense of "outward" for $n(y)$). As an interesting exercise the reader should show that if x *and* y in (22) are on a spherical surface S of diameter d, (see Fig. 2.34 (c)) then:

$$\frac{d\Omega(y,x)}{dA(y)} = \frac{1}{[n(y)\cdot\xi(y,x)]d^2}$$

The representation of $\Omega(S,x)$ in (21) is of particular value when the surface S is relatively concrete and has a specific analytic description, (parts of spheres, walls, and relatively tangible surfaces in general), whereas (10) of Sec. 2.5 is of greatest use when no surface S is specifiable and when instead a set D of directions *per se* is to be assigned a solid angle value. We close this example with the observation that all of Euclid's Optics [36] can be placed on a solid modern mathematical foundation using (21) and its various logical corollaries. (The translation of the first theorem in Euclid's Optics is given as a motto at the beginning of Volume I of this work. The theorem thus has several levels of meaning.)

Example 9: Irradiance via Surface Integrals

We return now to the integral for irradiance given in (8) of Sec. 2.5 and cast it into that form which is most useful when one must take into specific account the surface radiance of some surface S producing an irradiance $H(x,\xi)$ at some point x outside of S. The geometric setting for the present example is depicted by Fig. 2.34 (a), where at each point y of the surface S, there is given a surface radiance distribution $N(y,\cdot)$. We assume that all directions in $D(S,x)$ lie in a vacuum, that $D(S,x)$ is less than a hemisphere, and that the irradiance contributions to $H(x,\xi)$ come only from S so that $N(x,\xi')$ in (8) of Sec. 2.5 is zero for ξ' outside of $D(S,x)$. Hence (8) of Sec. 2.5 may be written:

$$H(x,\xi) = \int_{D(S,x)} N(x,\xi')\xi\cdot\xi' \, d\Omega(\xi')$$

In the present study the dummy variable "ξ'" is chosen to be the name of the variable direction $\xi(y,x)$ used in Example 8.

Then, in view of the radiance invariance law (2) of Sec. 2.6:

$$N(y,\xi') = N(x,\xi') \quad ,$$

it is clear that $H(x,\xi)$ is also represented by the integral:

$$H(x,\xi) = \int_S N(y,\xi')\xi\cdot\xi' \, d\Omega(y,x) \quad .$$

This may be written in the form:

$$H(x,\xi) = \int_S N(y,\xi')\xi\cdot\xi' \, \frac{d\Omega(y,x)}{dA(y)} \, dA(y)$$

which, by (22) is reducible to:

$$\boxed{H(x,\xi) = \int_S N(y,\xi(y,x)) \, \frac{\xi\cdot\xi(y,x)\mathbf{n}(y)\cdot\xi(y,x)}{r^2(y,x)} \, dA(y)} \qquad (23)$$

Equation (23) is the desired surface integral representation of $H(x,\xi)$. Suppose we write:

$$\text{"}\mathbf{H}(S,x)\text{"} \quad \text{for} \quad \int_S N(y,\xi(y,x)) \, \frac{\xi(y,x)\mathbf{n}(y)\cdot\xi(y,x)}{r^2(y,x)} \, dA(y) \quad .$$

This is the surface radiance counterpart to the field radiance definition of $\mathbf{H}(x)$ in (2) of Sec. 2.8. Then (23) becomes:

$$H(x,\xi) = \xi\cdot\mathbf{H}(S,x) \quad . \qquad (24)$$

Equation (24) suggests that the condition imposed at the outset, namely that $D(S,x)$ be less than a hemisphere, can evidently be relaxed. In that event (24) is generalizable to:

$$\bar{H}(x,\xi) = \xi\cdot\mathbf{H}(S,x) \quad , \qquad (25)$$

the proof of which is left to the reader.

If we assume that the point x is *inside* a closed surface S, then (23) still holds but with $\mathbf{n}(y)$ now being interpreted, if desired, as an *inward* unit normal from y to x. In that case, $\mathbf{H}(S,x)$ of (24) formally reduces to $\mathbf{H}(x)$ in (2) of Sec. 2.8. These observations suggest that the true field radiance counterpart to (25) is:

$$\bar{H}(x,\xi) = \xi\cdot\mathbf{H}(x,D) \qquad (26)$$

where we have written:

$$\text{"}\mathbb{H}(x,D)\text{"} \quad \text{for} \quad \int_D \xi' H(x,\xi')\, d\omega(\xi') \tag{27}$$

The connection between $\mathbb{H}(x,D)$ and $H(x)$ of Sec. 2.8 is clearly that:

$$H(x) = \mathbb{H}(x,\Xi) \quad . \tag{28}$$

An interesting special case of (23) arises when S is part of the inside of a spherical surface. In Fig. 2.34 (a) imagine x and y to be on the same spherical surface of diameter d, and now ξ in (23) is to be the unit outward normal to S at x, i.e., $\xi = -n(x)$. Under such conditions (see Fig. 2.34 (c)) it follows that:

$$\xi \cdot \xi(y,x) = n(y) \cdot \xi(y,x)$$

and

$$r(y,x) = d\omega(y) \cdot \xi(y,x) \quad .$$

Hence if $N(y,\xi(y,x)) = N$, over S, (23) yields:

$$H(x,\xi) = \frac{NA(S)}{d^2}$$

for every x on the sphere, and arbitrary subset S of the sphere.

Example 10: Radiant Flux Calculations

The irradiance integral (23) may be applied to the following radiometric setting, depicted in Fig. 2.35, which arises frequently in practice. A surface Y has a prescribed surface radiance $N(y,\cdot)$ at each point y. Surface X, which is disjoint from Y, receives an amount P(Y,X) of radiant flux from Y. It is required to express the amount P(Y,X) in terms of $N(y,\cdot)$ and the areas of X and Y, assuming the space between X and Y is a vacuum. Now from (23) we have for each x and ξ an expression for the irradiance $H(x,\xi)$, so that we can immediately compute P(Y,X) in terms of $H(x,n(x))$, using (6) of Sec. 2.4 (in which D is now $\Xi(n(x))$):

$$P(Y,X) = \int_X H(x,n(x))\, dA(x) \quad ,$$

where n(x) is the unit inward normal to X at x. Hence

$$P(Y,X) = \int_X \int_Y N(y,\xi(y;x)) \frac{n(x)\cdot\xi(y,x)\,\xi(y,x)\cdot n(y)\, dA(y)\, dA(x)}{r^2(y,x)}$$

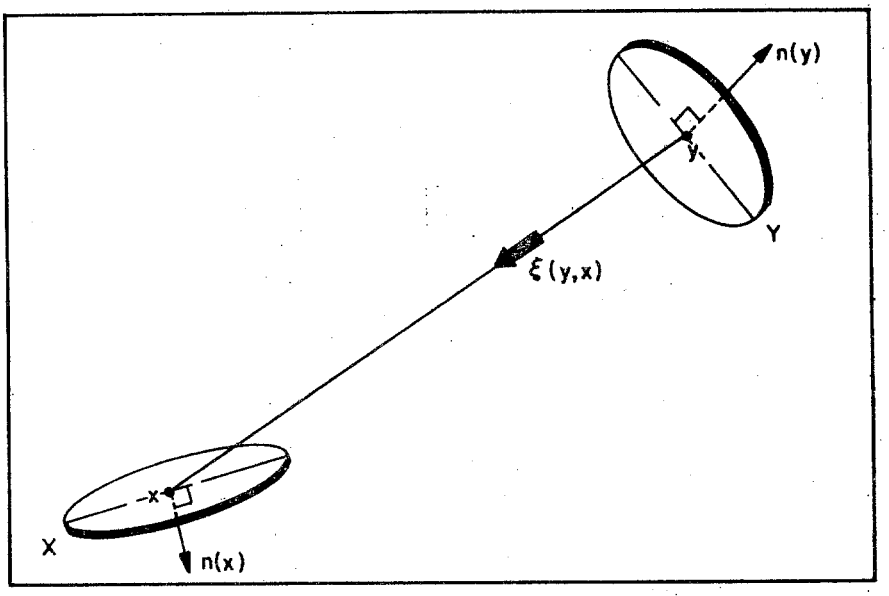


FIG. 2.35 A radiant flux calculation for two disjoint surfaces.

If we write

$$\text{"K(y,x)"} \quad \text{for} \quad \frac{\mathbf{n}(x)\cdot\xi(y,x)\,\xi(y,x)\cdot\mathbf{n}(y)}{r^2(y,x)}$$

Then, more succinctly,

$$P(Y,X) = \int_X \int_Y N(y,\xi(y,x))K(y,x)\, dA(y)\, dA(x) \tag{29}$$

Observe that $K(\cdot,\cdot)$ is a symmetric function, i.e., for every x and y,

$$K(x,y) = K(y,x) \quad .$$

If the areas A(X) and A(Y) of X and Y are small--say when each is a point source with respect to any point on the other--then (29) yields the useful approximate relation:

$$P(Y,X) = N(Y,X)\ K(y,x)\ A(Y)\ A(X) \tag{30}$$

where x and y are some fixed points of X and Y, respectively and "N(Y,X)" denotes the surface radiance of Y in the direction

$\xi(y,x)$ of X. Writing, *ad hoc*:

$$H(Y,X) \quad \text{for} \quad P(Y,X)/A(X)$$

and

$$\text{"}J(Y,X)\text{"} \quad \text{for} \quad N(Y,X)A(Y) \quad ,$$

Equation (30) becomes:

$$H(Y,X) = J(Y,X)\ \xi(y,x) \quad , \tag{31}$$

which is a highly compact formulation of several well-known radiometric laws, with built-in cosine laws for both irradiance and radiant intensity, and furthermore, with built-in inverse square law for irradiance and direct square law for radiant intensity.

Example 11: Intensity Area-Law for General Surfaces

This example serves to illustrate some further facets in the duality between irradiance and radiant intensity developed in various earlier sections throughout this chapter. In particular we now direct attention to the intensity counterparts of the relations (15)-(18) in Example 6 of this section.

Thus, starting with (22) of Sec. 2.9 as a conceptual base, let us write:

$$\text{"}\mathbf{A}\text{"} \quad \text{for} \quad \int_S \xi'(x)\ dA(x) \quad .$$

Then if a surface S has a constant uniform surface radiance N over the part $S(\xi)$ defined by a direction ξ (Sec. 2.9) then

$$\boxed{J = N\xi\cdot\mathbf{A}} \tag{32}$$

which is the exact intensity-counterpart to (15), and where we have written:

$$\text{"}J\text{"} \quad \text{for} \quad J(S,\xi) \quad .$$

When the shape of S is nearly planar, e.g., when $\xi'(x)$ varies within a solid angle 1/30 steradians over S, then we have, very nearly:

$$\mathbf{A} = \xi A$$

where A is the area of S. In this special case (32) yields the present counterpart to (16):

$$J = NA \quad . \tag{33}$$

If S is a point source with respect to a point x, a distance r from S, then the apparent area of another *point* source S'

similar in shape to S is related to that of S by the equation:

$$A(S') = Cl^2$$

where C is a constant and l is some given linear dimension of S. In this way we return to the direct square law for J in the limit of large r (or small A). That is, using this estimate of A(S') in (33), we obtain the present counterpart to the inverse square estimate of irradiance; and as it stands by itself, the preceding equation is the dual to the relation $\Omega = A/r^2$ used to estimate solid angle subtenses of point sources.

Still one more form for J, i.e., $J(S,\xi)$, is obtainable using the concept of vector area introduced in Sec. 2.9. Thus we have:

$$\boxed{J = \xi \cdot \mathbf{J}} \qquad (34)$$

which is the dual to (17), and where in the present case we have written:

$$\text{"}\mathbf{J}\text{"} \quad \text{for} \quad N\int_S \xi'(x)\, dA(x) \quad .$$

As a corollary we have:

$$\boxed{J = N\mathbf{A}} \quad , \qquad (35)$$

which is the dual to (18).

Example 12: On the Possibility of Inverse nth Power Irradiance Laws

The cumulative evidence of the preceding examples, beginning with Example 4, shows the predominant role played in radiometry by the inverse square law for irradiance. The law is evident in various guises in formulas (8) and (9) for spherical surfaces, in the point-source criterion of Example 5, in the discussion of Example 6, in (21), (22), (29), and finally, its dual (the direct square law for radiant intensity) is evident in the discussions of Example 11. All of this evidence appears to lead inexorably to the generalization that the distance fall-off of irradiance produced by flux from all real surfaces of uniform radiance must *eventually* (i.e., for large enough distance r) assume the inverse-square behavior with r. This guess is essentially correct. However, the result of Example 5 shows that for intermediate distances r, the irradiance decrease with r need not be exactly an inverse square type of decrease. A question of some interest now arises as to necessary conditions that may govern the *rate* of such decrease. For example, can a surface S be found such

that the irradiance H_r over some interval of distances r from the surface falls off exactly as C/r^3, (where C is a constant)? Or, perhaps S can be found such that H_r behaves exactly as C/r over some interval of r. In general, can a surface S be found such that over some interval of r, H_r is of the form C/r^n, where n is any number? In this example we devote some attention to these questions as they are of intrinsic interest and are of the kind which aid in forming a good intuition about the laws of geometrical radiometry.

Before attempting a systematic search for surfaces of the kind discussed above, let us consider some special cases which may point the way to an appropriate methodical approach. Figure 2.36 (a) depicts a sphere S of radius a and uniform radiance N. The results of Example 4 let us conclude that the irradiance H_r at distance r from the center of S is given by:

$$H_r = N(a^2/r^2)$$

Fig. 2.36 (c) depicts an infinite plane S of uniform radiance N. The results of Example 6 indicate that:

$$H_r = \pi N$$

for every r. Further, Fig. 2.36 (b) depicts an infinite conical surface S of half angle θ_o. Once again the results of Example 6 indicate that:

$$H_r = \pi N \sin^2 \theta_o$$

for every r, i.e., H_r in this case is independent of r but less than πN by the fixed factor $\sin^2 \theta_o$. Finally, Fig. 2.36 (d) depicts a general bounded closed smooth surface S of uniform radiance N which encloses a positive volume of space. Since S encloses a positive volume of space, there is a point x in S about which a sphere S_1 can be drawn such that S_1 lies wholly in S. Since S is bounded, there is a sphere S_2 with center x, such that S lies wholly in S_2. The spheres need not be tangent to S. It follows that H_r, the irradiance produced by S at any point outside S_2 a distance r from x, must obey the following equalities:

$$N(a_1^2/r^2) \leq H_r \leq N(a_2^2/r^2) \quad ,$$

where a_1 and a_2 are the radii of S_1 and S_2 respectively. This set of inequalities leads us to the following assertion: if S is *any bounded surface enclosing positive volume and with uniform surface radiance* N, *then associated with* S *is an irradiance function* H_r *whose graph is bounded by two inverse square curves.* Thus for sufficiently large r, H_r is arbitrarily closely described by an inverse square relation in r.

In view of the evidence just reviewed, the first main observation toward resolving the question before us may be made: *If* S *is a surface with uniform radiance* N *and the associated* H_r *is of the form* C/r^n *with* $n \neq 2$ *for every* $r \geq a$, *where* a *is some nonnegative number, then* S *is either (a) not*

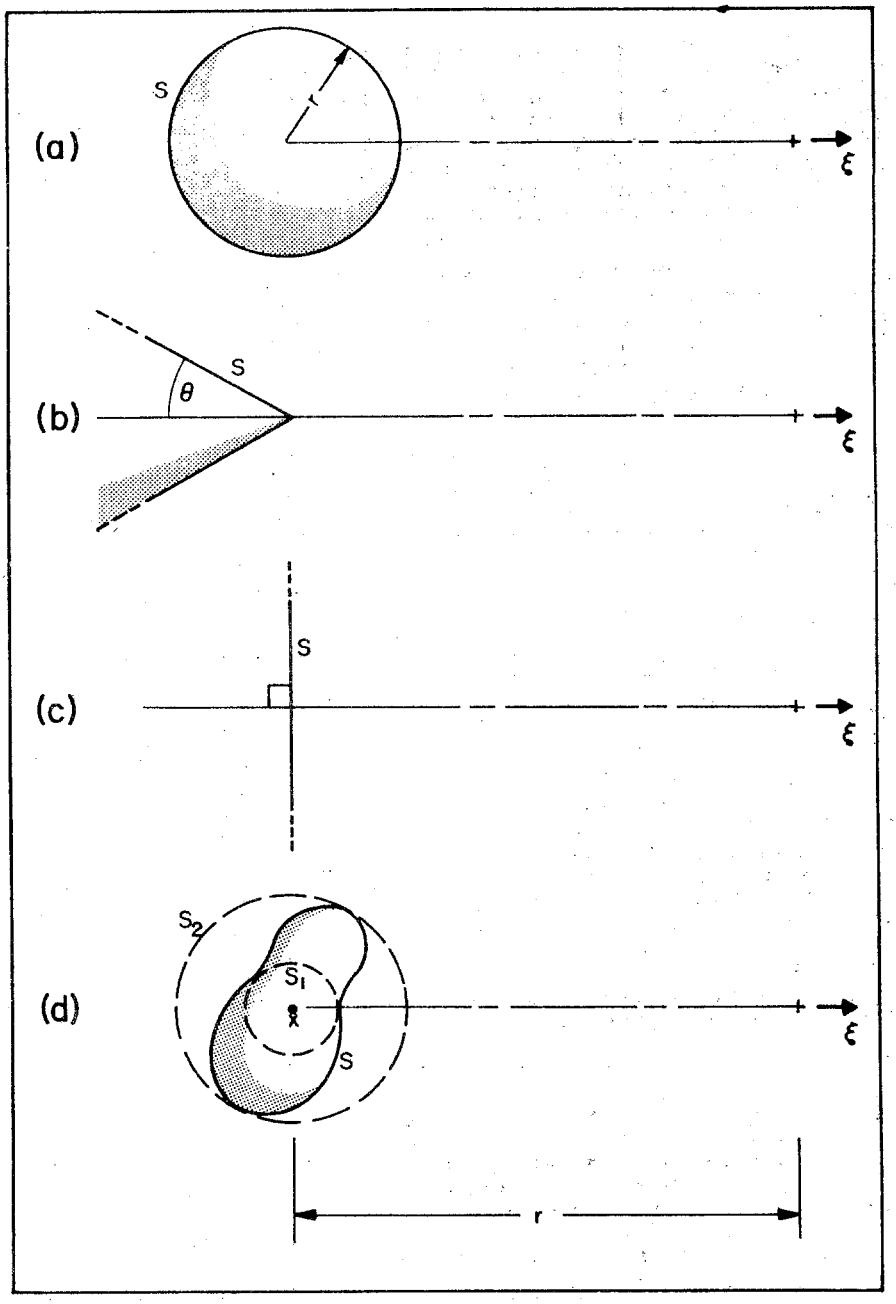


FIG. 2.36 Attempting to generalize the inverse square law for irradiance.

bounded or (b) does not enclose a positive volume.

Let us now attempt to find the actual shape of a surface S with the property that its associated irradiance varies precisely as C/r^n. Two conditions can be fixed at the outset in order to keep our initial search within reasonable confines. First we assume that $n \geq 0$. Secondly, we search only for surfaces S which, like those in (a)-(c) of Fig. 2.36, are surfaces of revolution about the direction ξ. It follows from (14) that the associated H_T is given by:

$$H_T = \pi N \sin^2 \theta(r) \quad ,$$

where $\theta(r)$ is the angle that the tangent to S from the point of observation, y, makes with the direction $-\xi$. (See Fig. 2.37.) Thus by specifying y and knowing S, we can in principle compute $\theta(r)$. Now use the hypothesized property of S to set up the following two equations:

$$\frac{C}{r^n} = H_T = \pi N \sin^2 \theta(r) \quad .$$

We can evaluate the constant C by observing that: if S is a smooth surface and $r = a$, then $\theta(r) = \pi/2$. This follows intuitively, e.g., from the observation that the surface has the appearance of an infinite plane for an observer at $r = a$. Hence for smooth surfaces:

$$\frac{C}{a^n} = \pi N$$

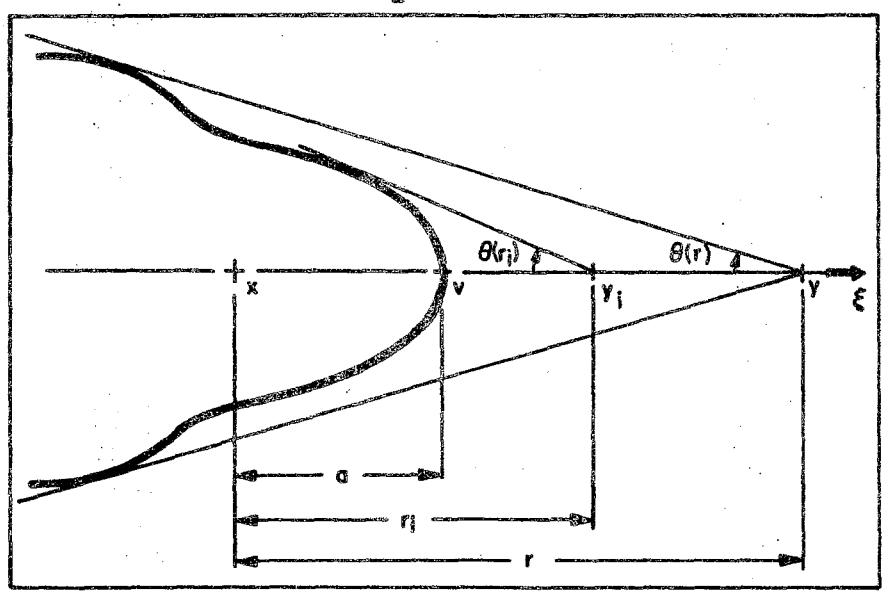


FIG. 2.37 Finding the shape of the luminous surface which has an inverse nth power irradiance law.

so that:

$$\sin \theta(r) = \left(\frac{a}{r}\right)^{n/2} \tag{36}$$

is the necessary connection between r and $\theta(r)$, where a is some fixed length associated with S. In the case of a spherical surface S and for the case of $n = 2$, a is simply the radius of S. In general if the surface S divides all of space into two separate parts (as, e.g., a plane or a paraboloid of revolution) then we agree that a is the distance from the *vertex* v of S to some *inside* point x, the *center* of S, where y is, by agreement, *outside* of S.

On the basis of relation (36) a graphical construction procedure can be evolved for the requisite surface S. First choose n, with $n > 0$ and choose a, with $a > 0$. Then select a set of distances $r_1, r_2, \ldots, r_k$, such that $a \leq r_1 < r_2 < \ldots < r_k$. Equation (36) may now be used to compute the associated angles $\theta(r_1), \theta(r_2), \ldots, \theta(r_k)$. These angles are used in the following manner. At each point y_i which lies a distance r_i from the center x along the axis of revolution of S, draw two straight lines making angles $\pm\theta(r_i)$ with the direction $-\xi$, (see Fig. 2.37). If the r_i have been spaced sufficiently closely together, then one may visually detect the *envelope* of the lines just drawn, i.e., the *curve* which is tangent to each straight line of the family just constructed. This envelope is the cross section of the desired surface S; i.e., by spinning this envelope around the direction ξ, the requisite S is formed.

Some experimentation with the preceding construction procedure yields information about how the surfaces S vary in shape as a function of the power n. Thus let the parameter a be fixed. Then for every n in the range $0 < n < 2$, we find that the associated surface S(n) is unbounded. The closer n is to 0, the more of a conical shape is exhibited by S(n) about its vertex. The limiting curve S(0) is a degenerate infinite cone $\theta_0 = \pi/2$, of the kind depicted in (b) of Fig. 2.36. The closer n is to 2, the more spherical is the shape of S(n) in the neighborhood of the vertex. The limiting curve S(2) is a sphere of radius a. The constructions of the surfaces S(n) for $n > 2$ at first present rather puzzling anomalies. By choosing the range of the values $r_1, \ldots, r_k$ sufficiently small and having the r_i closely packed together, one can construct the surfaces S(3), S(4), ..., within small regions around their vertices. In each case where $n > 2$, there is a critical distance r_0 from the center x beyond which the envelope construction degenerates. The larger n is, the smaller is the corresponding critical distance r_0.

These graphical experiments in constructing the surface S(n) for which the inverse nth power law for irradiance holds, especially in the case of $n > 2$, indicate the need for a relatively precise analytical approach to the problem of determining S(n). We shall now briefly direct some attention to such an approach.

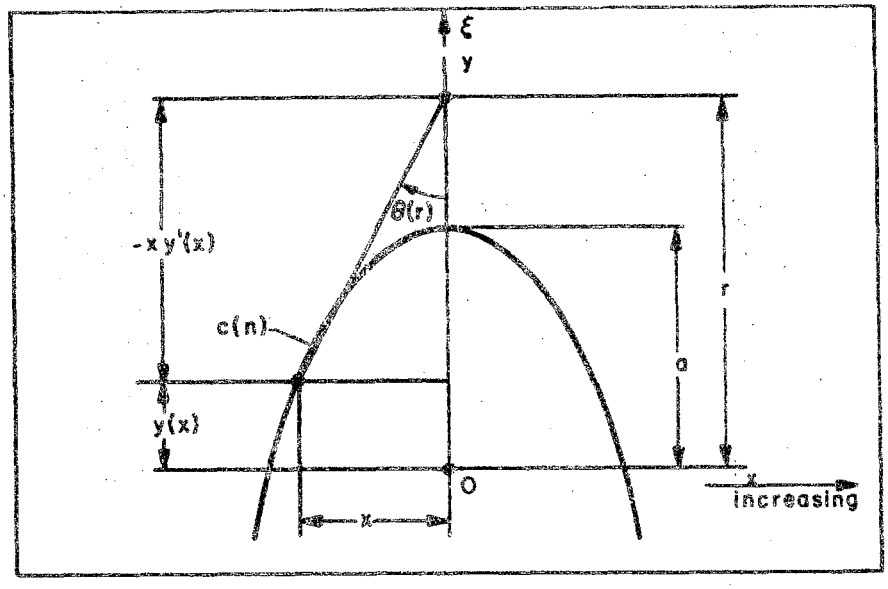


FIG. 2.38 Imbedding Fig. 2.37 in a cartesian frame.

Figure 2.38 shows an xy coordinate frame in which the cross section of surface S(n) is depicted by curve C(n). The irradiance meter is imagined to be at a point on the y-axis a distance r from the origin O of the frame. The unit inward normal ξ to the collecting surface of the meter is directed along the positive direction of the y-axis. The origin of the frame serves as the *center* of C(n), and the *vertex* of C(n) is a point on the y-axis a distance a from the origin. The curve C(n) is represented by some function y(·). Our primary goal is to obtain the differential equation for the function y(·) of the curve C(n). The starting point is equation (36) in the form:

$$\sin^2 \theta(r) = \left(\frac{a}{r}\right)^n$$

which, as we have seen, combines the inverse nth power requirement on the irradiance H_r with the general formula for H_r. We now systematically replace r, and $\theta(r)$, using the description of C(n) by y(·). Let us denote the derivative of y(x) with respect to x by "y'(x)". First of all r is clearly the algebraic sum of two terms: y(x) and -xy'(x), i.e., $r = y(x) - xy'(x)$, as a glance at Fig. 2.38 would show. Secondly, it is also clear from the figure that:

$$\tan \theta(r) = x/xy'(x)$$

$$= 1/y'(x)$$

From this:

$$\sin\theta(r) = \frac{1}{\sqrt{1 + (y'(r))^2}} \quad .$$

Hence:

$$(y'(x))^2 + 1 = [y(x) - xy'(x)]^n \, (1/a^n) \quad . \tag{37}$$

We have essentially reached our goal. Equation (37) is the differential equation for C(n). With an eye toward expediting the solution of (37) we shall rearrange it into the form:

$$\boxed{y(x) = xy'(x) + a\left[1+(y'(x))^2\right]^{1/n}} \tag{38}$$

Equation (38), as it stands, has the Gestalt of a Clairaut equation, an equation which is readily solvable in parametric form:

$$\boxed{\begin{aligned} x(t) &= -\frac{2at}{n}\left[1+t^2\right]^{\frac{1}{n}-1} \qquad (39)\\ y(t) &= -\frac{2at^2}{n}\left[1+t^2\right]^{\frac{1}{n}-1} + a\left[1+t^2\right]^{\frac{1}{n}} \qquad (40) \end{aligned}}$$

This equation also has a singular solution of the form

$$\boxed{y(x) = t_o x + a\left[1+t_o^2\right]^{1/n}} \tag{40a}$$

which represents straight lines of slope t_o. These singular solutions evidently can yield the degenerate conical case $\theta_o = \pi/2$ depicted in (c) of Fig. 2.36.

Setting n = 2 in (39) and (40), and eliminating the parameter t, we obtain:

$$x^2(t) + y^2(t) = a^2 \quad .$$

Hence C(2) is a circle with center at the origin (0,0), and of radius a, as expected. Setting n = 1 in (39) and (40), and eliminating the explicit appearance of the parameter t, we obtain

$$y(t) = -x^2(t)/4a + a$$

In this case S(1) is a paraboloid of revolution with focus at the origin (0,0), axis of symmetry along the y-axis, vertex at (0,a), and intercepting the x-axis at $x = \pm 2a$. The surface S(1) is typical, as far as size (unboundedness) and general orientation is concerned, of all S(n) with $0 < n < 2$.

For a curve-tracing analysis of a general S(n) we find from (39) and (40) that:

$$y(t)/x(t) = \left(1-\frac{n}{2}\right) t - \frac{n}{2t}$$

which is helpful in gauging the location of points on the curve. For example, if $n = 2$, then:

$$y(t)/x(t) = -\frac{1}{t}$$

for all t, $0 < t < \infty$. This shows that, as $t \to \infty$, a quadrant of the circle S(2) is traced out and $y(t)/x(t) \to 0$. This tracing is depicted in (b) of Fig. 2.39. Further, if $n = 1$, then:

$$y(t)/x(t) = \frac{t}{2} - \frac{1}{2t} \quad .$$

In this case, as t varies in the range $0 < t < \infty$, one branch of S(1) is traced out, and $y(t)/x(t) \to +\infty$. The tracing of this branch of the parabola S(1) is shown in (a) of Fig. 2.39. Finally, for the general case $n > 2$, we see that $\lim_{t\to\infty} y(t)/x(t) = +-\infty$, indicating that y(t) becomes much larger than x(t) as $t \to \infty$. This in itself is not too informative, but when coupled with the observation that for large t,

$$x(t) \quad \text{behaves like} \quad -\frac{2a}{n}\, t^{\frac{2-n}{n}}$$

and

$$y(t) \quad \text{behaves like} \quad a\left(1-\frac{2}{n}\right) t^{\frac{2}{n}} \quad ,$$

we gain further insight into the behavior of the curves. From these observations we cull the following information: as $t \to \infty$,

$$\begin{aligned} \text{for} \quad 0 < n < 2: &\quad x(t) \to -\infty, \quad y(t) \to -\infty \\ n = 2: &\quad x(t) \to -a, \quad y(t) \to 0 \\ n > 2: &\quad x(t) \to 0, \quad y(t) \to +\infty \quad . \end{aligned}$$

The behavior for the case $n > 2$ continues to present puzzling features. Thus, when $n > 2$, $x(t) \to 0$ for large t, indicating that $|x(t)|$ attains a maximum for some t. Examining (39) for this possibility, we see that for C(n),

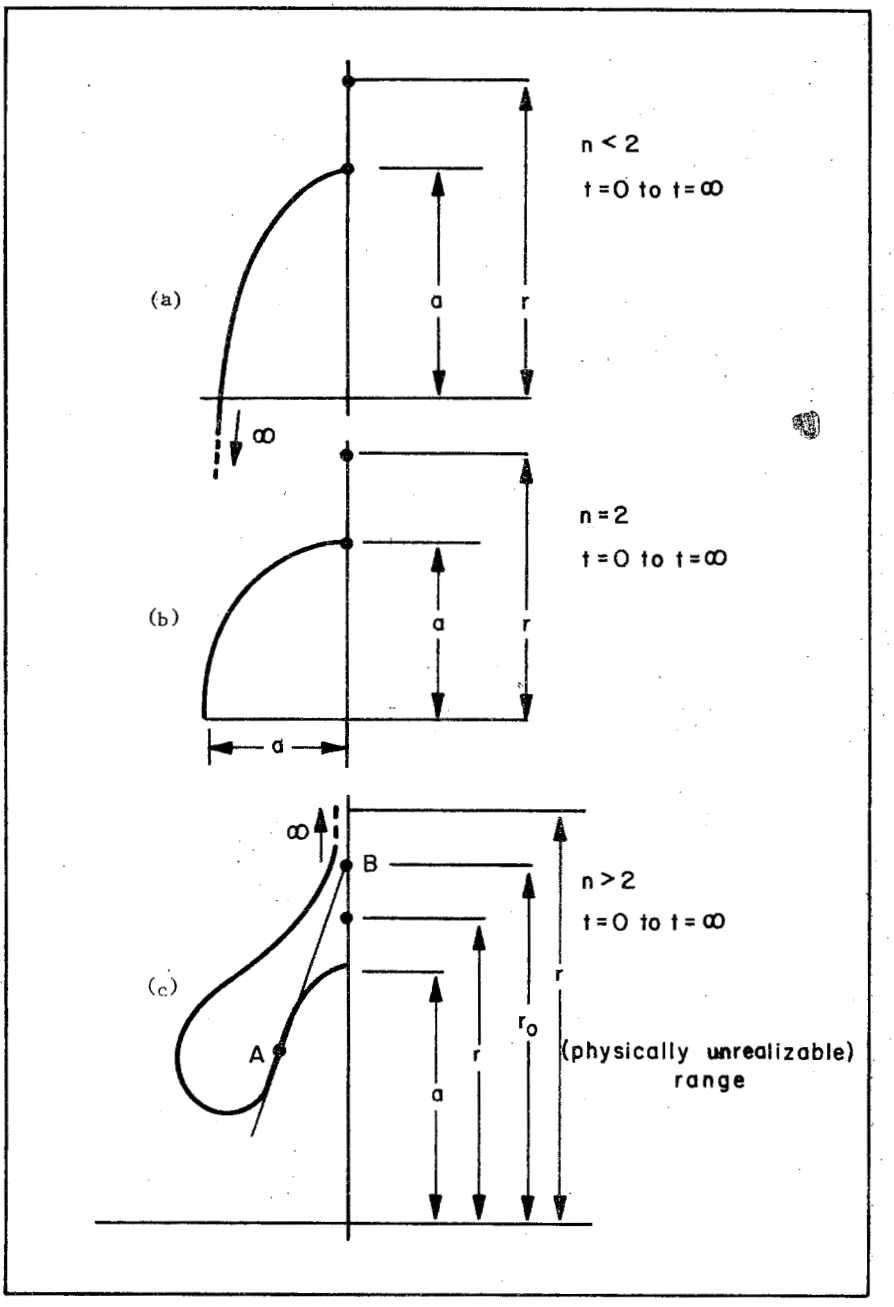


FIG. 2.39 Some cross sections of surfaces which produce irradiance fall-off like $1/r^n$ (see text).

$$\max\ |x(t)| \quad \text{occurs at} \quad t = \sqrt{\frac{n}{n-2}}$$

From this relation, it is at once clear that we obtain observable maxima for $n > 2$, and no observable (i.e., real) maxima for $0 < n < 2$. For example, we expect for $n = 3$, a maximum value of $|x(t)|$ at $t = \sqrt{3/(3-2)} = \sqrt{3}$. However, we expect no real maximum for $n = 1$, since $t = \sqrt{1/(1-2)} = \sqrt{-1}$. Finally, we inquire if, for the case $n > 2$, $S(n)$ ever intersects the x-axis as do its lower-n counterparts. Thus in (40) we set $y(t) = 0$, and find that:

$$y(t) = 0 \quad \text{when} \quad t = \sqrt{\frac{n}{2-n}} \quad .$$

Hence if $n > 2$, then $y(t) \neq 0$ for every real t. This means that $C(n)$ does not meet the x-axis for $n > 2$.

Summarizing the behavior of $C(n)$ for $n > 2$: as t varies from 0 to $+\infty$, a branch of $C(n)$ is traced out which is of the general form shown in (c) of Fig. 2.39. This figure explains the source of our difficulties in the geometric constructions based on equation (36). The construction is able to generate a branch of $S(n)$ from the vertex to the first point of inflection at point A. Beyond A, the branch of $C(n)$ and $S(n)$ itself has no conventional physical interpretation. Some interesting unconventional interpretations can be made; however, we leave it to the reader's initiative to interpret the meaning of $C(n)$ beyond point A. The tangent to $C(n)$ at A meets the y-axis at a point B, which determines the critical range r_0 for which the physically realizable inverse nth power law holds. Part (c) of Fig. 2.39 is drawn from computed data for the case $n = 3$. The associated set of parametric coordinates are given in the table below, which was computed by Mrs. Judith Marshall.

Table 1

Computed Values for Part (c) of Figure 2.39
(The case $n = 3$)

t	x(t)	y(t)	t	x(t)	y(t)
0	-0	1	2.0	-0.456	0.797
0.1	-0.066	0.997	2.5	-0.444	0.826
0.2	-0.130	0.987	3.0	-0.430	0.802
0.3	-0.194	0.971	3.5	-0.417	0.906
0.4	-0.242	0.954	4.0	-0.403	0.958
0.5	-0.287	0.933	10.	-0.307	1.580
1.0	-0.420	0.840	20.	-0.246	2.43
1.5	-0.455	0.798	30.	-0.213	3.25
1.6	-0.458	0.793	40.	-0.194	3.91
1.7	-0.460	0.790	50.	-0.180	4.60

Several general concluding comments can now be made on the problem of the inverse nth power law for irradiance. First, there is the constantly recurring use of or reference to the integer 2 throughout the most general of the preceding discussions. Observe how critically 2 enters into the following tabular classification of our main results:

Table 2

n	The surface S(n) inducing $H_r = C/r^n$	Radius of curvature at vertex	Range of validity
0	Curve with half-angle $\theta_o = \pi/2$ (Fig 2.36 (c))	0 (= 0/2a)	$r \geq a$
1	Paraboloid of revolution (Fig. 2.38)	1/2a	$r \geq a$
2	Sphere (Fig. 2.38)	1/a (= 2/2a)	$r \geq a$
3	3rd order luxoid (Fig. 2.39 (c))	3/2a	$r_o \geq r \geq a$
n	nth order luxoid (Fig. 2.39 (c))	n/2a	$r_o \geq r \geq a$

In this classification we encounter classical euclidean surfaces for all n, $0 < n \leq 2$, but a definite break occurs at $n = 2$, as has been repeatedly evident in the curve-tracing discussion above. All this is apparently closely related to the fact that we live in a three dimensional world, or at any rate, the radiometric laws above are represented most naturally in euclidean frames of dimension 3. The three dimensional geometric frame has been used implicitly throughout our discussions. We are thus led to conjecture that radiometry in a two dimensional world would have a ubiquitous inverse first power "irradiance" law and radiometry on a line would have its inverse zero power irradiance law. It is interesting to speculate on the theory and utility of k-dimensional geometric radiometry in which very likely the "irradiance" in such a geometry will obey an inverse k-1 power law, and to contemplate the potentially rich multiplicity of irradiances, scalar irradiances, and radiant intensities and their manifold interconnections latent in such geometries. Here the dualities brought out between irradiance and radiant intensity in the preceding examples are likely to attain their deepest and broadest meanings. These observations are commensurate with the conclusions, already brought out in the studies in Sec. 99 of Ref. [251], that radiometry and radiative transfer are

meaningfully formulable in arbitrarily general spaces. These interesting matters are left to the reader for further consideration.

Example 13: Irradiance from Elliptical Radiance Distributions

We shall now illustrate the use of equation (17) of Sec. 2.5 by computing the irradiance distribution associated with an important type of theoretical radiance distribution, namely the elliptical radiance distribution. The elliptical radiance distribution arises in the study of light patterns at great depths in oceans, lakes, and other natural hydrosols. It is a convenient theoretical standard against which to compare the light patterns that actually subsist in nature. The physical basis for the elliptical radiance distribution will be considered in Chapter 4. For the present we shall be concerned only with its geometric properties. In particular we shall investigate its structure with respect to direction, to variation in eccentricity, and also compute its associated vector and scalar irradiances.

We begin by setting the geometric stage of the computation. Figure 2.40 (a) depicts a terrestrial coordinate frame (Sec. 2.4) and the plane of a radiant flux collector oriented as usual by its unit inward normal ξ. Let the associated direction angles of ξ be θ and ϕ. Thus if $\theta = \pi$ and $\phi = 0$, say, then $\xi = -\hat{k}$ and the collector receives flux from the upper hemisphere, i.e., receives flux flowing in the directions of Ξ_+. In general, when ξ is the unit inward normal to the collecting surface, the incident radiant flux is along the directions of the hemisphere $\Xi(\xi)$, as defined in Sec. 2.4.

Next we define an *elliptical radiance distribution* at x, of eccentricity ε, $0 \leq \varepsilon \leq 1$, and magnitude N, to be a radiance distribution of the form:

$$N(x,\xi') = N/(1+\varepsilon\xi'\cdot\hat{k}) \tag{41}$$

An alternate mode of representation of $N(x,\xi')$ is by means of polar and azimuthal angles. Thus (41) may be written

$$N(x,\theta',\phi') = N/(1+\varepsilon \cos\theta') \quad . \tag{42}$$

The upper diagrams of Fig. 2.41 show four plots of elliptical radiance distributions $N(x,\cdot)$ of eccentricity $\varepsilon = .25, .50, .75, .95$. For small values of ε near 0 the associated distribution is predominantly spherical. For values of ε near 1, the associated distribution is long and narrow. When $\theta = 0$, the flow is upward and smallest; when $\theta = \pi$, the flow is downward and greatest, thus simulating, at least qualitatively, the real flows in nature. (We are using surface rather than field radiance.) The "size" of the distribution is governed by N, being the radiance in the horizontal directions, i.e., ξ's with angles of the form $(\pi/2,\phi)$. The ratio of downward (zenith) to horizontal radiances in the present model is given by:

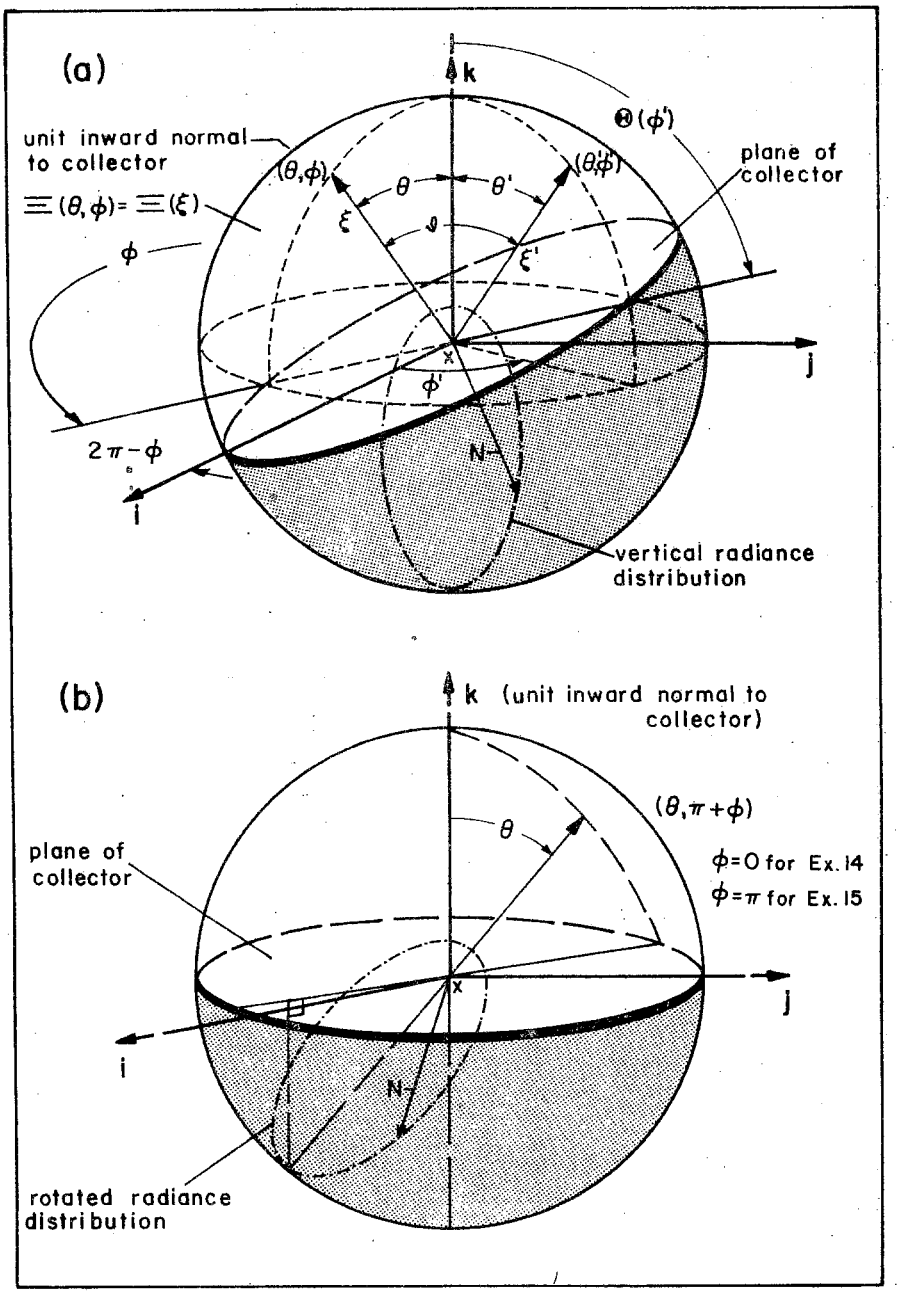


FIG. 2.40 Some calculation details for irradiance from elliptical radiance distributions.

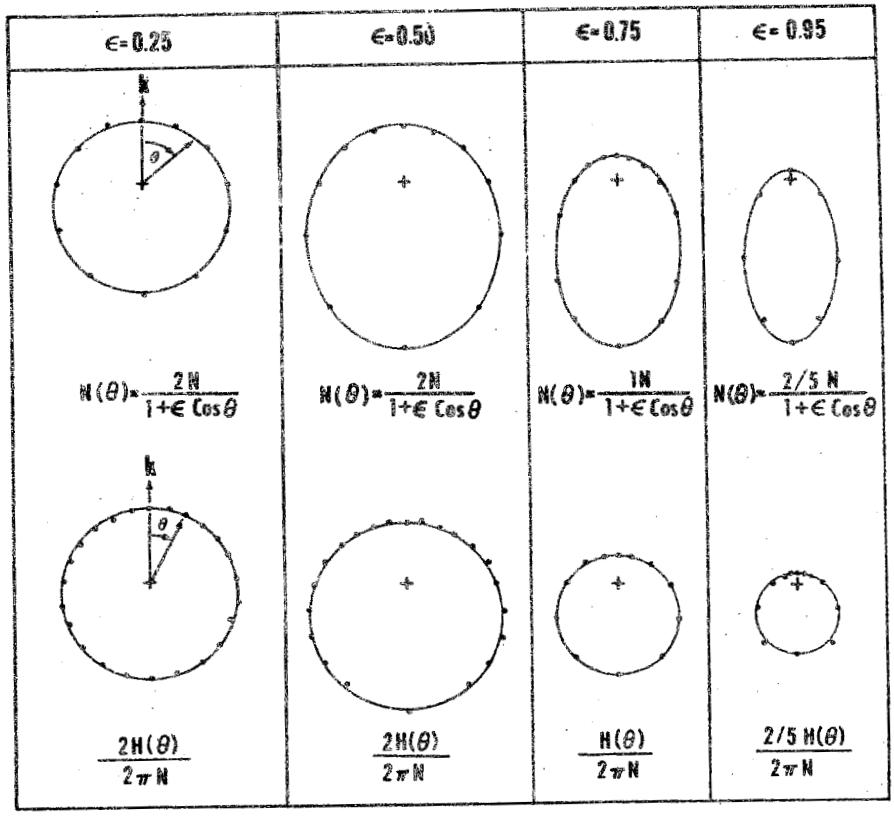


FIG. 2.41 Some representative irradiance distributions $H(\theta)$ associated with elliptical radiance distributions. The points are calculated from (48). The solid curves are circles, showing a possible simplification of (48) for engineering calculation purposes.

$$N(x,\pi,\phi')/N(x,\pi/2,\phi') = 1/(1-\epsilon)$$

Thus, the nearer ϵ is to 1, the greater is this ratio. The ratio of zenith (downward) to nadir (upward) radiance is:

$$N(x,\pi,\phi')/N(x,0,\phi') = \frac{1+\epsilon}{1-\epsilon}$$

Turning now to the computation of $H(x,\theta,\phi)$, we start with (17) of Sec. 2.5:

$$H(x,\theta,\phi) = \int_{\Xi(\theta,\phi)} N(x,\theta',\phi') \cos z\ell \sin\theta' \, d\theta' \, d\phi' \qquad (43)$$

where $\Xi(\theta,\phi)$ is the hemisphere $\Xi(\xi)$, as shown in part (a) of Fig. 2.40. This range of integration in (43) may be given explicitly:

$$H(x,\theta,\phi) = \int_{\phi'=0}^{2\pi} \int_{\theta'=0}^{\Theta(\phi')} N(x,\theta',\phi') \cos\varphi \sin\theta' \, d\theta' \, d\phi' \tag{44}$$

where $\Theta(\phi')$ is the angle between $\mathbf{k}$ and ξ', i.e., the variable direction of integration in the collector's plane which has azimuth ϕ'. $\Theta(\phi')$ may be determined from the functional relation:

$$\Theta(\phi') = \text{arc cot } \{-\tan\theta \cos(\phi-\phi')\} \quad . \tag{45}$$

Thus, e.g., if $\theta = 0$, $\Theta(\phi') = \text{arc cot } \{0\} = \pi/2$ for every ϕ', $0 \leq \phi \leq 2\pi$. Eq. (43) can be put into a more convenient form by using the fact that, as far as the quantity $H(x,\theta,\phi)$ is concerned, it is immaterial whether, on the one hand, the collector is tipped in the frame of reference of the radiance distribution as in (a) of Fig. 2.40; or on the other hand, the collector is held still and the radiance distribution is appropriately tipped in the frame of the collector as in (b) of Fig. 2.40. The computational merit of the arrangement in (b) is superior to that in part (a) of Fig. 2.40, and we shall adopt it in the present illustration. The salient change resulting in this switch of points of view is in the functional form of $N(x,\theta',\phi')$. Indeed, a glance at (a) and (b) of Fig. 2.40 shows that the "vertical" radiance distribution in part (a) has undergone a rigid rotation to the "tipped" form in part (b), and rotated about the vertical axis so that $\mathbf{k}$ goes into the unit vector whose angles are $(\theta,\pi+\phi)$.

The details of the transformation of $N(x,\theta',\phi')$ into its new form $N'(x,\theta',\phi')$ constitute a simple exercise in analytic geometry and are left for the reader to formulate (recall (18) of Sec. 2.5). The resultant form is:

$$N'(x,\theta',\phi') = \frac{N}{1 + \varepsilon(-\sin\theta \sin\theta' \cos\phi' + \cos\theta \cos\theta')} \tag{46}$$

in which we have set $\phi = 0$ since the desired irradiance $H(x,\theta,\phi)$ is independent of ϕ for the present radiance distribution, which is assumed symmetrical about the vertical. We can partially check (46) by letting $\theta' = \theta$ and $\phi' = \pi$. The resultant radiance is:

$$N'(x,\theta,\pi) = N/(1+\varepsilon)$$

which is precisely the magnitude of $N(x,0,\phi)$, as was to be expected. Using (46), it is clear from (b) of Fig. 2.40 that the desired irradiance $H(x,\theta,\phi)$ is given by:

$$H(x,\theta,\phi) = \int_{\phi=0}^{2\pi} \int_{\theta'=0}^{\pi/2} N'(x,\theta',\phi') \cos\theta' \sin\theta' \, d\theta' \, d\phi' \tag{47}$$

The integration details of (47) are straightforward and are therefore omitted. Writing, *ad hoc*:

$$\text{"}H(\theta)\text{"} \quad \text{for} \quad H(x,\theta,\phi)$$

we then evaluate (47) to obtain:

$$\boxed{H(\theta) = \frac{2\pi N}{\varepsilon^2}\left\{1+\varepsilon\cos\theta-\sqrt{1-\varepsilon^2\sin^2\theta}-(\cos\theta)\ln\left[\frac{(1+\varepsilon)(1+\cos\theta)}{\sqrt{1-\varepsilon^2\sin^2\theta}+\cos\theta}\right]\right\}} \tag{48}$$

which is the desired functional form of the irradiance distribution under an elliptical radiance distribution of eccentricity ε and magnitude N. The reader should now show how to use (48), without the need of further computation, for the case where the axis of the elliptical distribution is originally tilted at θ_0 from the vertical, and the angle between this axis and the unit inward normal to the collecting surface is $\mathscr{N}_0$.

Let us study some of the properties of $H(\theta)$. First of all, by setting $\theta = \pi$, we obtain the downward irradiance induced by the radiance distribution in (41) on the collecting surface:

$$H(\pi) = -\frac{2\pi N}{\varepsilon^2}\left[\varepsilon + \ln(1-\varepsilon)\right] \quad . \tag{49}$$

The upward irradiance is obtained by setting $\theta = 0$:

$$H(0) = \frac{2\pi N}{\varepsilon^2}\left[\varepsilon - \ln(1+\varepsilon)\right] \quad . \tag{50}$$

The net downward irradiance is therefore:

$$\begin{aligned} \bar{H}(\pi) &= H(\pi) - H(0) \\ &= \frac{2\pi N}{\varepsilon^2}\left[-2\varepsilon + \ln\left(\frac{1+\varepsilon}{1-\varepsilon}\right)\right] \end{aligned} \tag{51}$$

which is positive for all ε, $0 < \varepsilon < 1$. $\bar{H}(\pi)$ is the magnitude of the vector irradiance $\mathbf{H}(x)$ associated with (41). The direction of $\mathbf{H}(x)$ is evidently $-\mathbf{k}$. It may be verified directly from (48) that:

$$\bar{H}(\theta) = \bar{H}(\pi) \cos \psi \tag{52}$$

where we have written

$$\text{"}\bar{H}(\theta)\text{"} \quad \text{for} \quad H(\theta) - H(\psi)$$

and

$$\text{"}\psi\text{"} \quad \text{for} \quad \pi - \theta$$

Equation (52) is a specific example of (16) of Sec. 2.8.

Observe next that $H(\theta) \to \pi N$ as $\varepsilon \to 0$. This is to be expected since the elliptical radiance distribution becomes spherical as $\varepsilon \to 0$, and as we now know, a spherical radiance distribution of magnitude N, induces an irradiance πN. This fact about the limit of $H(\theta)$ as $\varepsilon \to 0$ may be seen relatively readily for a special case by letting $\varepsilon \to 0$ in (49). Indeed, expanding $\ln(1-\varepsilon)$ in a power series in ε, we have, for very small ε, as an approximation:

$$H(\pi) = 2\pi N \left(\frac{1}{2} + \frac{\varepsilon}{3} + \frac{\varepsilon^2}{4} \right) ,$$

$$H(0) = 2\pi N \left(\frac{1}{2} - \frac{\varepsilon}{3} + \frac{\varepsilon^2}{4} \right) ,$$

whence:

$$\bar{H}(\pi) = 4\pi\varepsilon N/3$$

The scalar irradiance induced by the elliptical radiance distribution (41) is also of interest. Using the representation (41) in (3) of Sec. 2.7 we have:

$$\boxed{h(\varepsilon) = \frac{2\pi N}{\varepsilon} \ln \left(\frac{1+\varepsilon}{1-\varepsilon} \right)} \tag{53}$$

where we have written, *ad hoc*:

$$\text{"}h(\varepsilon)\text{"} \quad \text{for} \quad h(x,t) ,$$

Note that:

$$\lim_{\varepsilon \to 0} h(\varepsilon) = 4\pi N .$$

For small ε, we have, very nearly:

$$h(\varepsilon) = 4\pi N \left[1 + \frac{\varepsilon^2}{3} \right] .$$

Comparing (53) and (51) we see that:

$$\bar{H}(\pi) = \frac{1}{\varepsilon}\,[h(\varepsilon)-h(0)] \quad . \tag{54}$$

Thus, the magnitude of the vector irradiance associated with an elliptical radiance distribution of eccentricity ε and magnitude N is $1/\varepsilon$ times the difference between the scalar irradiance $h(\varepsilon)$ associated with the distribution and the scalar irradiance associated with a *uniform* radiance distribution of magnitude N.

Finally, we consider the hemispherical scalar irradiances associated with (41) (see (7) of Sec. 2.7), writing *ad hoc*:

"$h(\varepsilon,-)$" for $h(x,-k,t)$

"$h(\varepsilon,+)$" for $h(x,k,t)$

we have for an elliptical radiance distribution:

$$h(\varepsilon,-) = -\frac{2\pi N}{\varepsilon}\ln(1-\varepsilon) \tag{55}$$

$$h(\varepsilon,+) = \frac{2\pi N}{\varepsilon}\ln(1+\varepsilon) \tag{56}$$

Adding these two and comparing the sum with (53) we have:

$$h(\varepsilon,+) + h(\varepsilon,-) = h(\varepsilon) \quad ,$$

which illustrates (9) of Sec. 2.7. In the two-flow theory of light fields, to be studied in Chapters 8 and 9, the following ratios are of interest in that theory (see also (30) of Sec. 10.7):

$$h(\varepsilon,-)/H(\pi) = \varepsilon\ln(1-\varepsilon)/(\varepsilon+\ln(1-\varepsilon)) \tag{57}$$

$$h(\varepsilon,+)/H(0) = \varepsilon\ln(1+\varepsilon)/(\varepsilon-\ln(1+\varepsilon)) \tag{58}$$

These ratios constitute convenient measures of the "collimatedness" of the elliptical radiance distribution. Thus for the case $\varepsilon = 0$, the distribution is spherical and the very antithesis of collimatedness. In this case:

$$\lim_{\varepsilon\to 0} h(\varepsilon,-)/H(\pi) = 2 \quad .$$

A similar limit, namely 2, holds for $h(\varepsilon,+)/H(0)$. In the other extreme, i.e., when ε is near 1, the elliptical distribution of downward flux is relatively collimated. In this case:

$$\lim_{\varepsilon\to 1} h(\varepsilon,-)/H(\pi) = 1 \quad .$$

On the other hand, and somewhat unexpectedly, the upward radiance approaches a certain shape for which:

$$\lim_{\epsilon \to 1} h(\epsilon,+)/H(0) = \ln 2/(1-\ln 2) \quad .$$

Equation (48) is plotted for the four values of ϵ given in Fig. 2.41. The plots of $2H(\theta)/2\pi N$ are shown in the lower line of that figure.

Example 14: Irradiance from Polynomial Radiance Distributions

The present example is assigned the task of developing a generalization of the elliptical radiance distribution considered in Example 13, and of developing a formula for the associated irradiance distribution. The main lesson of this example is one not of importance to radiometry *per se*. Rather, it is designed to bring to the reader's attention the fact that many of the techniques of classical polynomial and power series theory are available to help obtain analytical representations of the radiance distributions measured in nature and on which, in turn, one can base practical methods of computing the associated irradiance distributions.

Suppose then, that an empirical radiance distribution $N(x,\cdot)$ can be represented for each θ and ϕ by the following polynomial in $\cos\theta$:

$$N(x,\theta,\phi) = \sum_{j=0}^{n} a_j P_j(\cos\theta) \qquad (59)$$

where $P_j(\cos\theta)$ is the Legendre polynomial (in $\cos\theta$) of the first kind and of integral order j. The number n may be finite or infinite, as required. Here we are assuming that $N(x,\cdot)$ is a radiance distribution symmetrical about the vertical but of a form which has a quite general θ-dependence. As in the case of the elliptical radiance distribution in Example 13, we can use the fact that the irradiance produced by $N(x,\theta,\phi)$ in (59) depends only on the angle $\mathscr{l}$ between its axis of symmetry and the inward normal to the collecting surface. Therefore we can use the results of this example, without further effort, to compute irradiance on any collecting surface when the angle $\mathscr{l}$ between the axis of the symmetrical distribution and the unit inward normal to the collecting surface is known. Hence the assumption of the form (59) constitutes no loss of generality in this sense.

We first observe that the coefficients a_j are readily determinable from the tabulated data $N(x,\theta,\phi)$. Indeed, using the orthogonality properties of $P_j(\cos\theta)$, we have from (59) (and cf. (3) of Sec. 6.3):

$$a_k = \frac{(2k+1)}{2} \frac{1}{2\pi} \int_{\Xi} N(x,\theta',\phi') P_k(\cos\theta') \sin\theta' \, d\theta' \, d\phi'$$

i.e.,

$$a_k = \frac{(2k+1)}{2} \int_0^{\pi} N(x,\theta',\phi') P_k(\cos\theta') \sin\theta' \, d\theta' . \quad (60)$$

Hence, if $N(x,\cdot)$ is known, performing the operation on $N(x,\cdot)$ as defined in (60), yields a_k for every $k = 1,\ldots,n$. By writing:

$$\text{"}\mu\text{"} \quad \text{for} \quad \cos\theta'$$

and

$$\text{"}N(\mu)\text{"} \quad \text{for} \quad N(x,\theta',\phi') \quad ,$$

Equation (60) takes the relatively compact form:

$$\boxed{a_k = \frac{(2k+1)}{2} \int_{-1}^{+1} N(\mu) P_k(\mu) \, d\mu} \quad (61)$$

Equation (61) can be evaluated by any of several avaliable numerical quadratures, given the radiance data $N(\mu)$. Having obtained the a_k in this way, we now can go on to consider the computation details of $H(x,\cdot)$ associated with $N(x,\cdot)$. With (17) of Sec. 2.5 as a starting point and using (59) we can write:

$$H(x,\theta,\phi) = \iint_{\Xi(\theta,\phi)} \left(\sum_{j=0}^{n} a_j P_j(\cos\theta') \right) \cos\nu \sin\theta' \, d\theta' \, d\phi' \quad (62)$$

where ν is the angle between the directions (θ',ϕ') and (θ,ϕ). The preceding integral can be transformed into an alternate form by adopting the technique used in Example 13. (See Fig. 12.40 (b).) Thus Equation (62) can be written:

$$H(x,\theta,\phi) = \int_{\phi'=0}^{2\pi} \int_{\theta'=0}^{\pi/2} \left(\sum_{j=0}^{n} a_j P_j(\cos\nu') \right) \cos\theta' \sin\theta' \, d\theta' \, d\phi' \quad (63)$$

and which may be viewed as the present counterpart of (47), wherein:

$$\cos\nu' = \sin\theta \sin\theta' \cos\phi' + \cos\theta \cos\theta' \quad .$$

Equation (63) now stands in a form which is readily evaluable. Toward this end, observe that the sum of Legendre polynomial terms can be written in the form:

$$\sum_{j=0}^{n} a_j P_j(\cos \mathscr{V}') = \sum_{j=0}^{n} b_j \cos^j \mathscr{V}' \qquad (64)$$

where the numbers b_j, $j = 0, \ldots, n$, are obtained by expanding each $P_j(\cos \mathscr{V}')$ in the left side of (64) in powers of $\cos \mathscr{V}'$ and collecting together like powers of $\cos \mathscr{V}'$. Each "b_j" therefore denotes the coefficient of $\cos^j \mathscr{V}'$ so obtained. Hence from knowledge of the a_j, the numbers b_j are readily computed. Tables of Legendre Polynomials are available for the determinations of the b_j.

Next write, *ad hoc*:

"x" for $\sin\theta \sin\theta'$

"y" for $\cos\theta \cos\theta'$.

Then for every j, 1,..., n

$$(\cos \mathscr{V}')^j = (x \cos\phi' + y)^j$$

$$= \sum_{i=0}^{j} {}^jC_i \, x^{j-i} \, y^i \cos^{j-i}\phi' \quad ,$$

where "${}^jC_i$" as usual denotes the combinatorial coefficient of the ith term in the jth power of a binomial. Using this expansion in (63), with the help of (64), we can rearrange (63) to read:

$$H(x,\theta,\phi) =$$

$$= \sum_{j=0}^{n} b_j \int_{\phi'=0}^{2\pi} \int_{\theta'=0}^{\pi/2} \sum_{i=0}^{j} {}^jC_i x^{j-i} y^i \cos^{j-i}\phi' \cos\theta' \sin\theta' \, d\theta' \, d\phi'$$

$$= \sum_{j=0}^{n} b_j \sum_{i=0}^{j} {}^iC_j \int_{\theta'=0}^{\pi/2} x^{j-i} y^i \cos\theta' \sin\theta' \left[\int_{\phi'=0}^{2\pi} \cos^{j-i}\phi' \, d\phi' \right] d\theta' \, . \qquad (65)$$

Observe that:

$$\int_{\phi'=0}^{2\pi} \cos^{j-i}\phi' \, d\phi' = 2\sqrt{\pi} \, \frac{\Gamma\left(\frac{j-i+1}{2}\right)}{\Gamma\left(\frac{j-i+2}{2}\right)} \, \Delta_{ij} \quad ,$$

where

$$\Delta_{ij} = \begin{cases} 0 \text{ if } j-i \text{ is odd} \\ 1 \text{ if } j-i \text{ is even} \end{cases} \quad .$$

Let "I_{ij}" momentarily denote the value of this integral of $\cos^{j-i}\phi'$. Then (65) becomes:

$$H(x,\theta,\phi) =$$

$$= \sum_{j=0}^{n} b_j \sum_{i=0}^{j} {}^{i}C_j I_{ij} \int_{0}^{\pi/2} \sin^{j-i}\theta \, \sin^{j-i+1}\theta' \, \cos^{i}\theta \, \cos^{i+1}\theta' \, d\theta'$$

$$= \sum_{j=0}^{n} b_j \sum_{i=0}^{j} {}^{i}C_j I_{ij} \, \sin^{j-i}\theta \, \cos^{i}\theta \int_{0}^{\pi/2} \sin^{j-i+1}\theta' \, \cos^{i+1}\theta' \, d\theta' \qquad (66)$$

Observe that:

$$\int_{0}^{\pi/2} \sin^{j-i+1}\theta' \, \cos^{i+1}\theta' \, d\theta' = \frac{1}{2} \, \frac{\Gamma\left(\frac{j-i+2}{2}\right) \, \Gamma\left(\frac{i+2}{2}\right)}{\Gamma\left(\frac{j+4}{2}\right)} \quad ,$$

where "$\Gamma(z)$" once again denotes the value of the gamma function Γ at z. Let us write "J_{ij}" for the product of I_{ij} and the latter integral. Hence:

$$J_{ij} = \sqrt{\pi} \cdot \frac{\Gamma\left(\frac{j-i+1}{2}\right) \, \Gamma\left(\frac{i+2}{2}\right)}{\Gamma\left(\frac{j+4}{2}\right)} \, \Delta_{ij} \quad .$$

If we write:

$$\text{"}C_j(\theta)\text{"} \quad \text{for} \quad \sum_{i=0}^{j} {}^{i}C_j \, J_{ij} \, \sin^{j-i}\theta \, \cos^{i}\theta$$

and

$$\text{"}H(\theta)\text{"} \quad \text{for} \quad H(x,\theta,\phi) \quad ,$$

then (66) becomes:

$$H(\theta) = \sum_{j=0}^{n} b_j C_j(\theta) \quad . \qquad (67)$$

This is the desired formula for the irradiance distribution associated with the radiance distribution in (59). Observe that the numbers ${}^{i}C_j\, J_{ij}$ are evaluable once for all, so that to use (67) with particular radiance distribution $N(x,\cdot)$ it is required only to evaluate the a_k by means of (61) to the desired degree of accuracy, and to obtain the b_k using (64). It is left as an exercise for the reader to evaluate ${}^{i}C_j J_{ij}$ and to obtain explicit formulas for the b_k in terms of the a_k for the first few values of $k = 1, 2, \ldots, n$, and to make a list of them so that the use of (67) is reduced to simply finding the a_k for each new application.

The reader may verify that the scalar irradiance h associated with a radiance distribution $N(x,\cdot)$ of the form in (59) is given by:

$$h = 2\pi \sum_{j=0}^{n} b_j \left(1 - (-1)^{j+1}\right)\Big/ j + 1 \quad . \qquad (68)$$

We close this example by observing two special cases of the polynomial distributions. First we note that the set of polynomial radiance distributions discussed above contains as a special case the elliptical radiance distribution of Example 13. To see this, in (64) choose the a_j subject to the condition that:

$$b_j = (-\varepsilon)^j N$$

for every integer $j \geq 0$, where $0 \leq \varepsilon < 1$, and where N is a non-negative number. Then:

$$N(x,\theta,\phi) = N \sum_{j=0}^{\infty} (-\varepsilon \cos\theta)^i$$

$$= N/(1 + \varepsilon \cos\theta) \quad ,$$

which is the form of (42). Secondly, an interesting radiance distribution associated with heavily overcast skies is a special case of (59). This is the "Moon-Spencer Sky" representation of radiance distributions and takes the following form. For every θ, $\pi/2 \leq \theta \leq \pi$:

$$N(x,\theta,\phi) = N(x,\pi/2,\phi)(1-2\cos\theta) \ .$$

The empirical details of determining this distribution may be found in Ref. [186].

Example 15: On the Formal Equivalence of Radiance and Irradiance Distributions

The present sequence of illustrations of the radiometric concepts is concluded with a discussion of the theoretical possibility of reversing the usual path between radiance and irradiance distributions. *We shall show that, given an irradiance distribution* $H(x,\cdot)$ *it is possible, in principle, to deduce the associated radiance distribution* $N(x,\cdot)$. This course of action is the reverse of that taken in the various Examples above, and in the discussion of Sec. 2.5. The theoretical and experimental significance of this reversal of the usual computation procedure was touched on briefly in Sec. 2.5 wherein also a practical scheme for obtaining $N(x,\cdot)$ from $H(x,\cdot)$ was suggested. The main purpose of this example is to show that this reverse path is possible not only on a numerical level, but also on an exact function-theoretic level. This is tantamount to showing that (8) of Sec. 2.5, when viewed as an integral equation with unknown $N(x,\cdot)$, has a unique solution in terms of the irradiance distribution $H(x,\cdot)$. We shall discuss this point of view in detail, as it affords an opportunity to illustrate how the use of advanced vector space concepts can facilitate the solutions of certain radiometric problems.

We can phrase the present problem in precise terms as follows: *Given:* the irradiance distribution $H(x,\cdot)$ at a point x in an optical medium. *Required:* the associated radiance distribution $N(x,\cdot)$. Now, for every direction ξ and point x we have, by (8) of Sec. 2.5:

$$H(x,\xi) = \int_{\Xi(\xi)} N(x,\xi')\xi\cdot\xi' \, d\Omega(\xi') \ . \tag{69}$$

Let us write:

$$\text{"}C(x)\text{"} \quad \text{for} \quad \frac{1}{2\pi}\int_{\Xi(\xi)} [\ ]\xi\cdot\xi' \, d\Omega(\xi') \ .$$

We call $C(x)$ the *cosine operator*, for obvious reasons. Then (69) can be written as:

$$\frac{H(x,\cdot)}{2\pi} = N(x,\cdot)\ C(x) \quad , \tag{70}$$

where "$N(x,\cdot)C(x)$" means: "operate on the radiance distribution $N(x,\cdot)$ by substituting $N(x,\cdot)$ into the square bracket of the integral operator $C(x)$." For example, if at point x, $N(x,\cdot)$ is a uniform radiance distribution with magnitude N, then for every ξ:

$$\begin{aligned} N(x,\cdot)C(x) &= \frac{1}{2\pi} \int_{\Xi(\xi)} N\xi\cdot\xi'\ d\Omega(\xi') \\ &= \frac{N}{2\pi} \int_{\Xi(\xi)} \xi\cdot\xi'\ d\Omega(\xi') \\ &= \frac{N}{2} \quad . \end{aligned} \tag{71}$$

Up to this point in the present example our deliberations have been relatively elementary and were without exception motivated by physical intuition. But now when we ask: "Can we determine $N(x,\cdot)$ knowing $H(x,\cdot)$ and $C(x)$?", we leave the domain of physical intuition and are asking a purely mathematical question. Perhaps even in this general radiometric setting some reader may see a physical reason for an affirmative answer to the query. For instance, by starting with the simpler setting in (21) of Sec. 2.5 and by letting the number of equations of the type considered there increase indefinitely and by being assured at each step along such a course that $N(x,\cdot)$ is determinable from $H(x,\cdot)$, perhaps by following such a line of thought one can be convinced of the general determinability of $N(x,\cdot)$ from $H(x,\cdot)$. Indeed, it is most desirable that some assurance be generated in such a manner. But at the present moment we are confronted by a mathematical question and in view of its important relevance to applications we prefer to settle it using now the rules of mathematics.

To begin to answer the preceding question we generate a mathematical setting in which the question suggests some further action toward the present goal. The appropriate setting is obtained by considering the set $\mathcal{N}(x)$ of all radiance distributions at point x. Next we observe the interesting fact that the sum of any two such radiance distributions is again in the set $\mathcal{N}(x)$. For example, if $N(x,\cdot)$ and $N'(x,\cdot)$ are in $\mathcal{N}(x)$, then the function:

$$N(x,\cdot) + N'(x,\cdot)$$

is in $\mathcal{N}(x)$ and by definition assigns to each direction ξ at x the sum $N(x,\xi) + N'(x,\xi)$ of the two radiances $N(x,\xi)$ and

$N'(x,\xi)$. The sum of two radiances is again a radiance*. That is, if each of $N(x,\cdot)$ and $N'(x,\cdot)$ is physically admissible then a lighting arrangement could be conceived so that $N(x,\cdot) + N'(x,\cdot)$ was realizable. Observe, however, that we need not introduce the preceding observation as an additional justification for the assertion about $\mathcal{N}(x)$ containing the sum of $N(x,\cdot)$ and $N'(x,\cdot)$ whenever it contains each. That assertion is simply the result of the present definition of $\mathcal{N}(x)$, and of the general definition of radiance. Next we observe that if $N(x,\cdot)$ is in $\mathcal{N}(x)$, so is $cN(x,\cdot)$, where c is a non negative real number. The physical plausibility of this assertion is obvious. As a result of these observations we see that $\mathcal{N}(x)$ is part of the vector space $V(x)$ of *all* functions $f(x,\cdot)$ at x over the domain Ξ, and with dimensions of radiance. In fact $\mathcal{N}(x)$ forms what is referred to in mathematical terminology as a *non negative cone* of $V(x)$ which, by definition, is closed under formation of sums, and multiplication by non negative real numbers. (Take all the unit vectors in a subset Ξ_0 of Ξ and form the set of all products $c\xi$, with $c \geq 0$, and ξ in Ξ_0. Describe the geometrical appearance of this set.)

Now what is the purpose of all this collecting together of huge families $\mathcal{N}(x)$ of radiance distributions? Simply this: by collecting together the members of $\mathcal{N}(x)$ in the fashion just exhibited, the operator $C(x)$ defined above takes on the crucial role of a *linear transformation* from $V(x)$ to $V(x)$ and in this setting our original question, "Can we obtain $N(x,\cdot)$ from $H(x,\cdot)$?," takes a deeper and mathematically meaningful cast.

Before rephrasing the question in the vector space terminology it may be well to include, simply for completeness, a comment about what it means for $C(x)$ to be a linear transformation. It means this: If $f(x,\cdot)$ and $g(x,\cdot)$ are any two functions in $V(x)$ and a and b are any two real numbers, then

$$[af(x,\cdot) + bg(x,\cdot)]\, C(x) = a[f(x,\cdot)C(x)] + b[g(x,\cdot)C(x)]$$

where $f(x,\cdot)C(x)$ and $g(x,\cdot)C(x)$ are again members of $V(x)$, being images of $f(x,\cdot)$ and $g(x,\cdot)$ under $C(x)$. Observe that $C(x)$ acting on a function with dimensions of radiance yields up once again a function with dimensions of radiance.

We can now ask our question about $H(x,\cdot)$ and $N(x,\cdot)$ as follows: "Is the linear transformation $C(x)$ from $V(x)$ to $V(x)$ a one-to-one transformation when restricted to the part $\mathcal{N}(x)$ of $V(x)$?" By $C(x)$ being "one-to-one", is meant that $C(x)$ sends exactly one radiance function into each modified irradiance function of the form: $H(x,\cdot)/2\pi$ defined in (70). Then, having given an irradiance distribution $H(x,\cdot)$, we are thereby assured that there is one and only one radiance function that it comes from (i.e., is associated with). Hence, whenever

---

*This may be taken as intuitively obvious at this point of the exposition. Formally, it is a consequence of the interaction principle of Chapter 3.

$C(x)$ is one-to-one we are encouraged to find that unique radiance distribution $N(x,\cdot)$ which, by (70), yields up $H(x,\cdot)$.

It turns out that the linear transformation $C(x)$ is indeed one-to-one when restricted to $\mathcal{N}(x)$ and we find its inverse $C^{-1}(x)$ as follows. We begin by observing that:

$$\int_{\Xi} \left(H(x,\xi)/2\pi\right)\, d\Omega(\xi) = \int_{\Xi} \left(N(x,\cdot)C(x)\right)\, d\Omega(\xi)$$

$$= \frac{1}{2\pi} \int_{\Xi} \left[\int_{\Xi(\xi)} N(x,\xi')\xi'\cdot\xi \, d\Omega(\xi') \right] d\Omega(\xi)$$

$$= \frac{1}{2\pi} \int_{\Xi} N(x,\xi') \left[\int_{\Xi(\xi')} \xi'\cdot\xi \, d\Omega(\xi) \right] d\Omega(\xi')$$

$$= \frac{1}{2} \int_{\Xi} N(x,\xi')\, d\Omega(\xi')$$

$$= \frac{1}{2}\, h(x) \quad .$$

For our present purpose, let us write:

$$\text{"}|f(x,\cdot)|\text{"} \quad \text{for} \quad \int_{\Xi} |f(x,\xi)|\, d\Omega(\xi) \quad ,$$

where $f(x,\cdot)$ is in $V(x)$. In particular, if $f(x,\cdot)$ is in $\mathcal{N}(x)$, then $|f(x,\cdot)|$ is the scalar irradiance associated with a general radiance distribution $f(x,\cdot)$ at point x. The main thing the preceding calculation has shown is that:

$$|N(x,\cdot)C(x)| = \frac{1}{2}\, |N(x,\cdot)| \quad . \tag{72}$$

The significance of this equality for the present discussion is crucial, and we pause to make this significance clear. The significance becomes clear when it is pointed out that the scalar irradiance $h(x)$ acts as the "length" of the vector $N(x,\cdot)$ in $V(x)$. Indeed, the bars around "$f(x,\cdot)$" in the definition above are there to point up the easily verified fact that $|f(x,\cdot)|$ is analogous to the absolute value of a number or vector; and it may be shown that all the essential properties of length that we carry with us from euclidean space hold also for the numbers $|N(x,\cdot)|$. We call $|N(x,\cdot)|$ (i.e., $h(x)$ in this case) the *radiometric norm* of $N(x,\cdot)$, to point up this similarity between $N(x,\cdot)$ and the usual concept of norm or length of a vector.

The significance of (72) can now be stated: the linear transformation C(x) has the property that it maps a radiance function into one which has exactly half the norm (i.e., "length") of the original radiance distribution. In short, C(x) has the *norm contracting* property with contracting factor 1/2. The mathematical consequence of this fact is immediate: we now can use the well known norm-contracting theorem of vector space theory, as stated, e.g., in Ref. [251] for the radiative transfer context, to assert that the inverse $C^{-1}(x)$ of C(x) exists, and that, indeed:

$$C^{-1}(x) = \sum_{j=0}^{\infty} (I - C(x))^j \quad , \tag{73}$$

where I is the identity transformation, i.e.,

$$f(x,\cdot)I = f(x,\cdot)$$

for every $f(x,\cdot)$ in V(x). This identity transformation can be written as an integral operator. Thus if we write:

$$\text{"I"} \quad \text{for} \quad \int_{\Xi(\xi)} [\,]\delta(\xi-\xi')\, d\Omega(\xi')$$

it may be verified that I is the identity operator on V(x) whenever δ is the Dirac delta function (on the space with Ω as measure). Then if we go on to write:

$$\text{"D(x)"} \quad \text{for} \quad \frac{1}{2\pi}\int_{\Xi(\xi)} [\,]\{2\pi\delta(\xi-\xi') - \xi\cdot\xi'\}\, d\Omega(\xi')$$

we have the equivalent form for (73), where

$$D(x) = \{I-C(x)\}$$

and

$$\boxed{C^{-1}(x) = \sum_{j=0}^{\infty} D^j(x)} \quad . \tag{74}$$

and

$$\boxed{N(x,\cdot) = \frac{H(x,\cdot)}{2\pi}\, C^{-1}(x)} \tag{75}$$

where, in turn, we have defined $D^j(x)$ recursively by writing:

$$\text{"}D^{0}(x)\text{"} \quad \text{for} \quad I \quad ,$$

and for every positive integer j:

$$\text{"}D^{j}(x)\text{"} \quad \text{for} \quad D^{j-1}(x)D(x) \quad ,$$

and where, finally, "$D^{j-1}(x)D(x)$" denotes the customary integral operation on $D^{j-1}(x)$ as an integrand in $D(x)$. Equation (74) yields the requisite inverse of $C(x)$, and the solution of our present problem is summarized in (75).

Observe that to use the norm-contracting theorem in Ref. [251] we actually need the fact that $I-C(x)$ is a norm-contracting operator. The reader may now easily verify that:

$$\left|N(x,\cdot)\left(I-C(x)\right)\right| = \frac{1}{2}\left|N(x,\cdot)\right| \quad ,$$

so that $I-C(x)$ is, indeed, norm-contracting with contracting factor 1/2, and the norm-contracting theorem statement yields (73) and hence (74).

Aside from the relatively advanced mathematical objects involved in (74) (namely, Dirac delta functions, and two-dimensional iterated integration) the algebraic essence of (74) is identical to that of the formula used by every high school student summing a geometric series of the form:

$$(1-x) + (1-x)^2 + (1-x)^3 + \ldots$$

whose value is clearly $1/x$ and where x is any number with absolute value less than 1. Now, instead of squaring (1-x), i.e., multiplying (1-x) by itself, we are required to operate with $I-C(x)$ on itself. Thus, e.g.,

$$N(x,\cdot)\left[I-C(x)\right]^2 =$$

$$= \frac{1}{2\pi}\int_{\Xi(\xi'')} \left[\frac{1}{2\pi}\int_{\Xi(\xi')} N(x,\xi)\left(2\pi\delta(\xi'-\xi)-\xi'\cdot\xi\right)d\Omega(\xi)\right]\left(2\pi\delta(\xi''-\xi')-\xi''\cdot\xi'\right)\,d\Omega(\xi')$$

To obtain the form for $\left(I-C(x)\right)^2$ itself, simply remove "$N(x,\cdot)$" and "$N(x,\xi)$" where they occur in the preceding equality. Thus, as in the case of computing the "fraction" $1/x$ by using solely multiplication, addition and subtraction repeatedly, so too can we compute "$1/C(x)$", i.e., $C^{-1}(x)$ using solely integration, multiplication, addition and subtraction, repeatedly. The norm-contracting theorem states that by continuing sufficiently far, $C^{-1}(x)$ can be arbitrarily closely approximated.

The error engendered by stopping the computation of $C^{-1}(x)$ in (74) at the kth term may be readily computed. Thus, write,

$$\text{"}N^{(k)}(x,\cdot)\text{"} \quad \text{for} \quad \frac{1}{2\pi}\sum_{j=0}^{k} H(x,\cdot)D^j(x)$$

so that $N^{(k)}(x,\cdot)$ serves as the kth order approximation to the desired distribution $N(x,\cdot)$. Then the radiometric norm of the difference between $N(x,\cdot)$ and $N^{(k)}(x,\cdot)$ is:

$$\begin{aligned}
\left|N(x,\cdot)-N^{(k)}(x,\cdot)\right| &= \left|\sum_{j=k+1}^{\infty} \frac{H(x,\cdot)}{2\pi} D^j(x)\right| \\
&= \sum_{j=k+1}^{\infty} \left|\frac{H(x,\cdot)}{2\pi} D^j(x)\right| \\
&= \sum_{j=k+1}^{\infty} \frac{1}{2^j} \left|\frac{H(x,\cdot)}{2\pi}\right| \\
&= \frac{1}{2^k} \left|\frac{H(x,\cdot)}{2\pi}\right|
\end{aligned}$$

The reader may use as specific cases in (75) the formulas for $H(x,\theta,\phi)$ in (48) and (67) of Examples 13 and 14 in order to recover the associated radiance distributions of those examples. These will afford non trivial examples of (75).

We close this discussion with some general assertions to which one is naturally led after contemplating the lesson of the present example. The assertions concern the possibility of still further equivalences between radiance and other radiometric concepts which are natural generalizations of the concept of irradiance. Recall that irradiance was defined empirically by specifying a small plane surface S onto each point of which radiant flux could be incident within the set $\Xi(\xi)$, where ξ is the unit inward normal to S. If now we replace $\Xi(\xi)$ by *any* fixed conical set $D(\xi)$ of directions of positive solid angle content specified in some way with respect to ξ, then the generalized irradiance distribution $H(x,D(\cdot))$, as defined in (4) of Sec. 2.4, is equivalent to $N(x,\cdot)$ in the same sense that $H(x,\cdot)$ and $N(x,\cdot)$ were shown to be equivalent in the present example. This is the first assertion to which we are led. Its proof is left to the reader.

The lesson of the present example can be carried still further than the point reached in the preceding paragraph. Let "$S(x,\xi)$" denote a collecting surface S which is a convex surface of revolution of fixed shape and size whose location and orientation in an optical medium X are uniquely specified

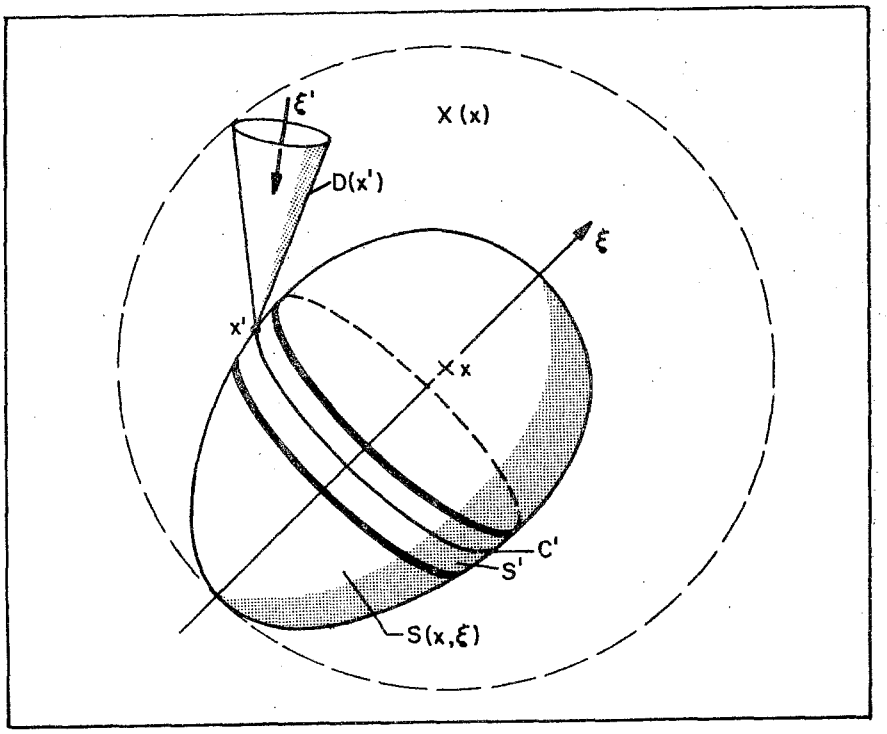


FIG. 2.42 A diagram of a radiometrically adequate collector. How many of them are there? (See text)

by locating a fixed point x on (or within) S and giving the direction ξ of the sensed axis of revolution of S. A typical surface of the type $S(x,\xi)$ is pictured in Fig. 2.42. Let S' be a *proper band* of latitude circles on S, i.e., such that S' has positive area and such that to the points x' of each latitude circle C' of S' there is assigned a right circular cone D(x') of directions whose axis direction ξ' lies in the plane of x' and ξ and makes a given angle with ξ, and of common positive solid angle opening $\Omega(D(x'))$. We shall require that the values $\xi\cdot\xi'$ and $\Omega(D(x'))$ are fixed for the points x' on each latitude circle C' on S' but may vary from circle to circle on S'. Let X(x) be the spherical region swept out by $S(x,\xi)$ as x is held fixed and ξ allowed to vary through all of Ξ. Finally, assume that a general radiance distribution of fixed structure is defined at each point within X(x). Then if "$P(S(x,\xi))$" denotes the radiant flux collected by S for a given x and ξ, we make the following plausible assertion with the above conditions in mind: *For every point x in the optical medium X, the radiance distribution* $N(x,\cdot)$ *is equivalent to the radiant flux distribution* $P(S(x,\cdot))$ *in the sense that there is a one-to-one integral operator* $E(S,x)$ *such that:*

$$P(S(x,\cdot)) = N(x,\cdot)\, E(S,x) \qquad (76)$$

The preceding assertion clearly contains the irradiance assertions above as special cases. For example, let S be a plane circular surface of positive area, with unit inward normal ξ and center x. Let S' be one side of S such that $D(x') = E(\xi)$ for every x' in S'. Then under the conditions of the preceding assertion, we have:

$$P(S(x,\xi)) = H(x,\xi)\ A(S) \quad ,$$

so that, according to (70) and (76):

$$E(S,x) = 2\pi C(x)\ A(S) \quad ,$$

where A(S) is the area of the plane circular surface S.

These examples do not exhaust the possibilities inherent in (70) and (76); however, they will suffice for the present to show that there is an infinite class of radiometric functions each member of which is equivalent to the radiance function in the sense of there being a one-to-one linear transformation between the vector spaces of radiance distributions and radiant flux distributions of such functions. Let us say that an arbitrary convex surface S is a *radiometrically adequate collector* in an optical medium X if its associated radiant flux distribution $P(S(x,\cdot))$ is equivalent, in the sense of the present example, to $N(x,\cdot)$ for every point x in X. We close this example with the following problem directed to interested readers: *Characterize the most general class of radiometrically adequate collectors. (In other words: give the necessary and sufficient conditions that a surface* S *be a radiometrically adequate collector.)* We have shown in the present example that plane circular surfaces, and more generally, have conjectured that surfaces of revolution such as cylinders, spheres, hemispheres, spherical caps, prolate and oblate spheroids, etc., can be radiometrically adequate collectors. It is certainly clear, at least intuitively, that the class of radiometrically adequate collectors is quite large and could, under suitable qualifications, contain surfaces not necessarily surfaces of revolution, such as the Platonic "solids", rectangular parallelepipeds, convex surfaces, and even certain non convex surfaces. However, non convex surfaces introduce self-interreflection complications which cannot be handled until the interaction principle (Chapter 3) has been studied, and therefore for the present at any rate, will be omitted from the problem stated above.

2.12 Transition from Radiometry to Photometry

The concepts of classical photometry, to which we turn our attention in this section, are designed to give quantitative measures of the capability of radiant flux to evoke the sensation of brightness in human eyes. These measures all rest in the single concept of the *standard luminosity function* the key concept in the science of photometry. Photometry is principally concerned with the precise description of and the deductions from the relative visibility of monochromatic radiant flux as a function of wavelength and as embodied in the

standard luminosity function. The depth to which we shall study photometry will be only so far that the reader may gain an insight into the principal features of the subject and a competence in working with photometric concepts, in the forms they commonly occur in the study of applied hydrologic optics. Such interesting problems as the physiological basis of color vision, which lie at the base of the subject, transcend the scope of the present discussion.

We shall initially motivate the transition to the photometric concepts by means of hypothetical experiments designed to acquaint the reader with the main empirical features of photometry. The experiments described are to be understood as didactic devices and as such omit the wealth of detail required for the implementation of their real counterparts. Once the essential idea of the transition has been explained and the transition made from the concept of radiance to that of its photometric counterpart, luminance, then we shall embark on a systematic transition to geometrical photometry and compile our results in compact tabular form suitable for convenient reference.

The Individual Luminosity Functions

Figure 2.43 depicts an observer viewing a screen in a well-lighted room. The screen is divided into two equal areas, and is devised so that on the left half a radiance of fixed amount $N(\lambda_o)$ of fixed wavelength λ_o is constantly displayed throughout the experiment. The magnitude of $N(\lambda_o)$ is chosen comparable to daylight radiances. The right half of

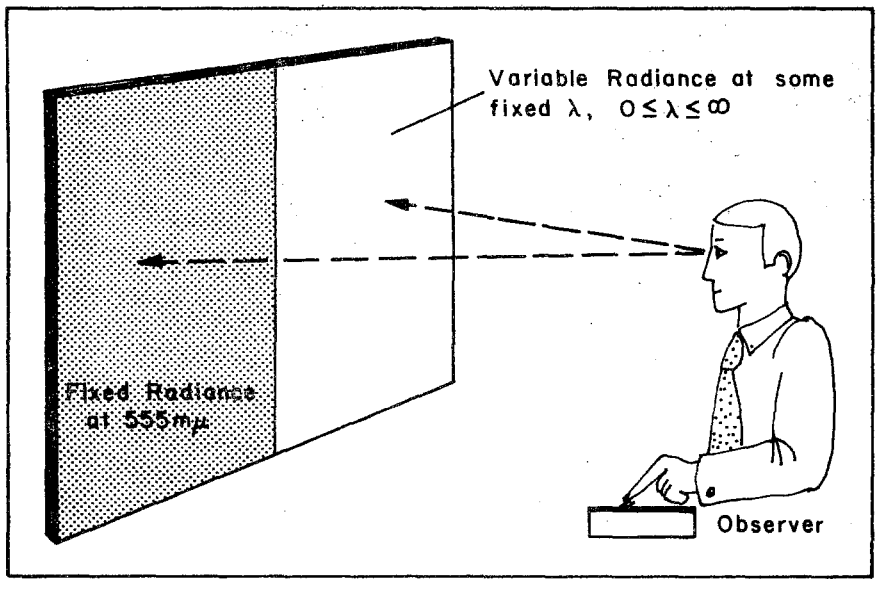


FIG. 2.43 A schematic setting for the empirical determination of the individual luminosity function.

the screen displays a radiance of variable amount $N(\lambda_j)$ of fixed wavelength λ_j in some set L of wavelengths. The observer begins the experiment with $N(\lambda_j)=0$ and, by means of a controlling device, slowly increases $N(\lambda_j)$ from 0 *until that magnitude* $N(\lambda_j)$ *is attained for which he believes the brightness sensation produced in his brain by* $N(\lambda_j)$ *is equal to that produced by* $N(\lambda_o)$. This decision process may require some preliminary trials on the observer's part. Soon, however, he makes a decision that $N(\lambda_j)$ and $N(\lambda_o)$ are of equal "brightness" and presses a button, thereby recording $N(\lambda_j)$. This procedure is now repeated for the other wavelengths over the electromagnetic spectrum.

The experiment with observer a_i just described results in a set of radiance values $N(\lambda_j)$ with λ_j in the set L. These values, it should be noted, are associated with the particular observer a_i used in the preceding experiment. To occasionally point up this fact let us write "$N(a_i,\lambda_j)$" for the radiance that matches $N(\lambda_o)$ as judged by observer a_i in some set A of observers which have performed the experiment. As a result of these experiments, to each observer a_i we may assign his particular *luminosity function* defined as follows. We write:

$$\text{"}\bar{y}(a_i,\lambda_j)\text{"} \quad \text{for} \quad N(\lambda_o)/N(a_i,\lambda_j) \tag{1}$$

for every λ_j in L, and call $\bar{y}(a_i,\cdot)$ the *luminosity function for observer* a_i. The value $\bar{y}(a_i,\lambda_j)$ is called the *luminosity* of the wavelength λ_j, as judged by observer a_i.

Matters have been arranged (on the basis of earlier experiments with observer a_i, not recorded here) so that wavelength λ_o was the wavelength of maximum luminosity for observer a_i. To see what this means, recall that $N(a_i,\lambda_j)$ is chosen to be of such a magnitude as to match $N(\lambda_o)$ in its capability of evoking the sensation of brightness. Since $N(\lambda_o)$, the radiance with wavelength λ_o of maximum luminosity is fixed in magnitude, all other radiances $N(a_i,\lambda_j)$ must then be increased to give the same brightness sensation to a_i as did the radiance $N(\lambda_o)$. Hence a plot of $\bar{y}(a_i,\lambda_j)$ versus λ_j for each observer a_i in A will have a graph of the general form in Fig. 2.44. At $\lambda_j = \lambda_o$, $y(a_i,\lambda_o) = 1$. For every other λ_j, $\bar{y}(a_i,\lambda_j) < 1$. To point up the fact that λ_o varies from observer to observer, let us write, alternatively, "$\lambda_o(a_i)$" for the λ_o of observer a_i.

Once each observer in the experimental group A has been assigned a luminosity function, this information could be used to predict the subjective sensation of brightness of a given sample of monochromatic radiant flux in the following sense. Suppose that observer a_i encounters a radiance of magnitude $N(\lambda_j)$. Then by (1) we can predict that this radiance would appear to him to have the same "brightness" as a sample of radiant flux of wavelength $\lambda_o(a_i)$ and radiance:

$$N(\lambda_j)\bar{y}(a_i,\lambda_j) \quad . \tag{2}$$

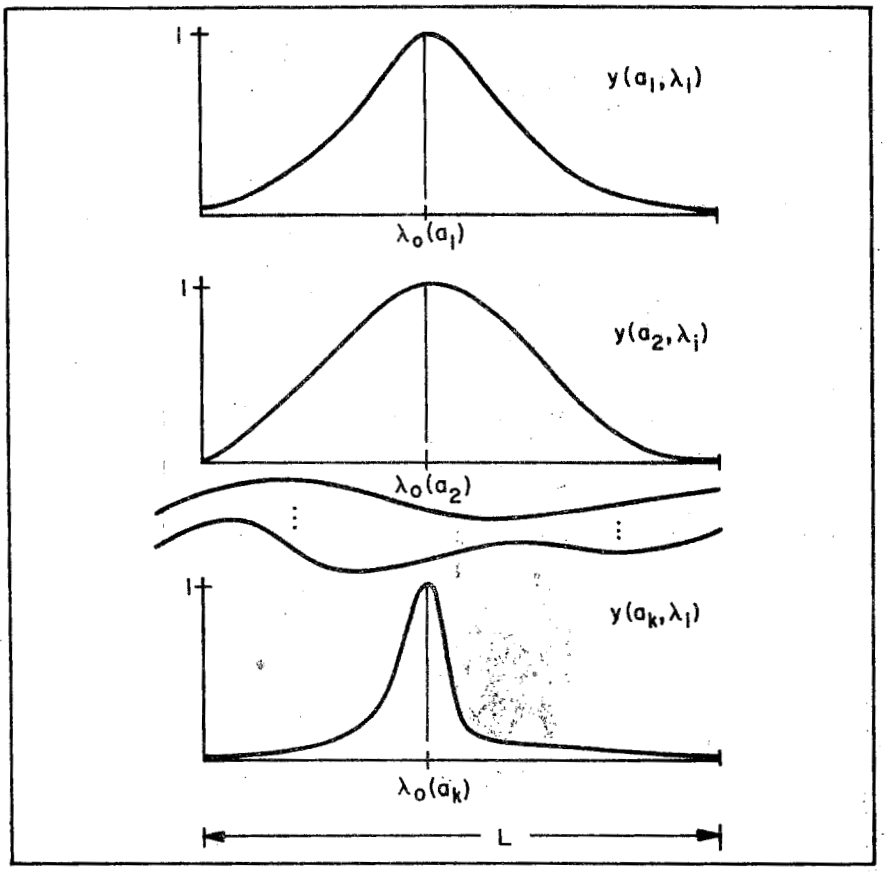


FIG. 2.44 Some individual luminosity functions (schematic only).

The term follows very simply and logically from (1). But the *interpretation of this term, as just stated, is not compelled to follow from (1) by the laws of algebra.* To make this interpretation we must first make an assumption (preferably explicitly) that the subjective sensation of brightness that can be produced by a radiance $N(\lambda_j)$ *varies linearly* with the magnitude of $N(\lambda_j)$. Thus if we were to double $N(\lambda_j)$, then the sensation would be the same as that produced by viewing radiant flux of wavelength λ_o and of double the radiance $N(\lambda_j)\bar{y}(a_i,\lambda_j)$. The reasonableness of this assumption rests critically on the stability of a_i's luminosity curve with respect to the absolute magnitude of $N(\lambda_o(a_i))$ used in the experiment, and on the general lighting level within the experimental room. Actual experimental evidence indicates that the luminosity function for a_i is dependent to a measurable degree on both $N(\lambda_o(a_i))$ and the background radiance. The description

of the hypothetical experiment above was careful to note that the experiment took place in a well-lighted room. To point up this fact the resultant luminosity curves in Fig. 2.44 are called *photopic luminosity* curves. When the observers view the screen in a darkened room with dark-adapted vision, it is found that the luminosity curves shift en masse 50-60 mμ to the left with very slight overall change in shape. The resultant curves are called *scotopic luminosity* curves.

To summarize, a consistent, workable interpretation of the meaning of (2) requires an explicit assumption of the linearity of subjective brightness sensations with respect to the magnitude $N(\lambda_j)$. This assumption might be at slight variance with experimental fact over some ranges of values of $N(\lambda_j)$, but it has the virtue of leading the way to a scientific basis of photometry. A science, it will be recalled, is an organized body of knowledge sustained within the webwork of a set of generally accepted conventions. To raise photometry to the level of a science, for better or worse, requires the explicit statement of adopted conventions of the kind just discussed.

We return now to experimental subject a_i, whose photopic luminosity function has been determined, and attempt to predict a new kind of response of a_i to radiant flux. Suppose now that a_i is confronted with a radiance in the right half of the screen in Fig. 2.43 which consists of a radiance which is a superimposed mixture of two monochromatic radiances $N(\lambda_j)$, $N(\lambda_k)$ from the set L, say of distinct wavelengths λ_j and λ_k. Were he confronted with each separately, we would be able to predict the equivalent sensation producing radiance of the wavelength $\lambda_o(a_i)$ by performing twice the operation in (2): once for λ_j and then again for λ_k. In an attempt to predict the sensation producing capabilities of radiance of wavelength $\lambda_o(a_i)$ equivalent to that of the radiance mixture $N(\lambda_j)+N(\lambda_k)$ we are tempted by simple energy-addition arguments to say that:

$$N(\lambda_j)\bar{y}(a_i,\lambda_j) + N(\lambda_k)\bar{y}(a_i,\lambda_k) \qquad (3)$$

is the requisite radiance. However, there appears to be no experimental evidence to substantiate this attempt, although practical calculations based on (3), and physiological eye-mechanisms tend to lend some support of (3). In the absence of such experimental evidence and in the presence of a desire to progress to a scientific discipline, we must make an explicit assumption to the effect that: *the radiance of wavelength* $\lambda_o(a_i)$, *capable of producing the same sensation of brightness as a mixture of two radiances of wavelengths* λ_j *and* λ_k, *is given by (3) above.* Clearly this is a generalization of the linearity assumption above, the earlier form being obtained by setting $\lambda_j = \lambda_k$.

Once the preceding assumption--(or *definition* of equivalent radiance of wavelength λ_o, as it should preferably be called)--is made, the path toward a sound basis for the science of photometry is cleared of one further obstacle. Indeed, it is but a formal step from (3) to the following general definition for the relative luminance distribution

associated with a radiance distribution at a point x in an optical medium: Let $N(x,\cdot,t,\lambda)$ be the radiance distribution at x at time t for wavelength λ. Then the associated *relative luminance distribution* with respect to observer a_i is the function:

$$\int_0^\infty N(x,\cdot,t,\lambda)\bar{y}(a_i,\lambda)\, d\lambda \qquad (4)$$

which assigns to each ξ at x at time t, the *relative luminance*, with respect to a_i, of the *integrated radiance* distribution $\int_0^\infty N(x,\cdot,t,\lambda)\, d\lambda$. We shall denote the latter by "$N(x,\cdot,t,\Lambda)$".

A minor technical point should be noted here before going further, a point which concerns the integration of radiance with respect to *wavelength* λ rather than *frequency* ν. It will be recalled that the basis for integrating radiance over the spectrum of frequencies was established in Sec. 2.3, and that the possibility of such an operation is guaranteed by the additivity and continuity properties of Φ with respect to frequency (cf., (1) and (2) in Sec. 2.2). By noting that $\nu\lambda = \nu$ implies $d\nu = -(\nu/\lambda^2)\, d\lambda$, each integration with respect to ν can be cast into an integration with respect to λ. (See note (c) to Table 3 below.) Whenever such a change of variables from ν to λ is made, we assume that the factor $-(\nu/\lambda^2)$ is suitably absorbed in the radiometric symbol, and the dimension of the radiometric concept, e.g., radiance, as far as the frequency component is concerned, is tacitly changed from "per unit frequency length" to "per unit wavelength".

Returning now to (4), we attempt to interpret (4) after the fashion of the interpretation of (3). A straightforward extension of the interpretation of (3) is the following: *for a given direction ξ, (4) is the amount of monochromatic radiance of wavelength $\lambda_0(a_i)$ which would produce an equivalent sensation of brightness in the brain of observer a_i as would the integrated radiance* $N(x,\xi,t,\Lambda)$, where Λ is the entire wavelength (or frequency) spectrum. In view of the preceding observations, in the definitions (4) of Sec. 2.5, one can replace "F" by "Λ" and have:

$$N(x,\xi,t,\Lambda) = \int_0^\infty N(x,\xi,t,\lambda)\, d\lambda \quad ,$$

by virtue of (4), Sec. 2.3. It is to be particularly noted that the preceding italicized interpretation is a *formal* interpretation with no known empirical basis--except for the single case where the given radiance distribution is monochromatic.

With the preceding interpretation of (4) in mind, we next return to (1) and emulate that definition in the present heterochromatic setting of (4). Thus, we write:

$$\text{"}\bar{y}(a_i)\text{"} \quad \text{for} \quad \frac{1}{N(\Lambda)} \int_0^\infty N(\lambda)\bar{y}(a_i,\lambda)\, d\lambda \qquad (5)$$

and where for brevity we have written:

$$\text{"}N(\lambda)\text{"} \quad \text{for} \quad N(x,\cdot,t,\lambda)$$

and

$$\text{"}N(\Lambda)\text{"} \quad \text{for} \quad N(x,\cdot,t,\Lambda) \quad .$$

We call $\bar{y}(a_i)$ the *relative luminosity* of the radiance $N(\cdot)$ over Λ for the observer a_i. In this way we come to one of the principal definitions of photometry:

Let $\mathscr{R}$ be any radiometric concept of geometrical radiometry (radiance, irradiance, radiant intensity, etc.), defined over the part M of the spectrum Λ. The *relative luminosity of $\mathscr{R}$ over M for an observer* a_i is the number $\bar{Y}(\mathscr{R},M,a_i)$ where we have written:

$$\text{"}\bar{Y}(\mathscr{R},M,a_i)\text{"} \quad \text{for} \quad \frac{\left[\int_M \mathscr{R}(\lambda)\bar{y}(a_i,\lambda)\,d\lambda\right]}{\int_M \mathscr{R}(\lambda)\,d\lambda} \tag{6}$$

The Standard Luminosity Functions

We now re-examine the family of individual relative luminosity functions, depicted in Fig. 2.44, and attempt to define a single luminosity function which is representative of the entire set A of individual observers. There are several ways to go about this. For example in one method, we can go through the set of graphs of Fig. 2.44, note each $\lambda_o(a_i)$ and make a histogram, over λ in Λ, of the number of observers whose maximum luminosity occurred at wavelength λ. A typical histogram that would result is shown schematically in part (a) of Fig. 2.45. All indications in real experiments and theoretical considerations point to a gaussian distribution for the ideal limit of such histograms as the number of members in the set A increases indefinitely. The peak of the distribution is found in actual experiments to occur near a λ of 555 mμ. Next, a general wavelength λ is selected and the graphs of Fig. 2.44 are combed through with the specific goal in mind of finding the spread of values of $\bar{y}(a_i,\lambda)$ over the a_i in A. This spread of values is then split up into intervals. Part (b) of Fig. 2.45 depicts a typical histogram with the abscissas locating the observed values $\bar{y}(a_i,\lambda)$ occurring over the selected set of intervals, and the ordinates giving the number of a_i in each interval. Part (b) of Fig. 2.45 is adapted from Fig. 3.03a of Moon's treatise on Illuminating Engineering (Ref. [185]), which in turn is derived from actual experimental results by Coblentz and Emerson who gathered data from a set A of 125 observers. By means of (b) of Fig. 2.45, the relative luminosity value of 0.1750 is assigned to the standard observer for $\lambda = 640$ mμ.

By going through the entire spectrum in this way--i.e., by repeating the process summarized in (b) of Fig. 2.45, now

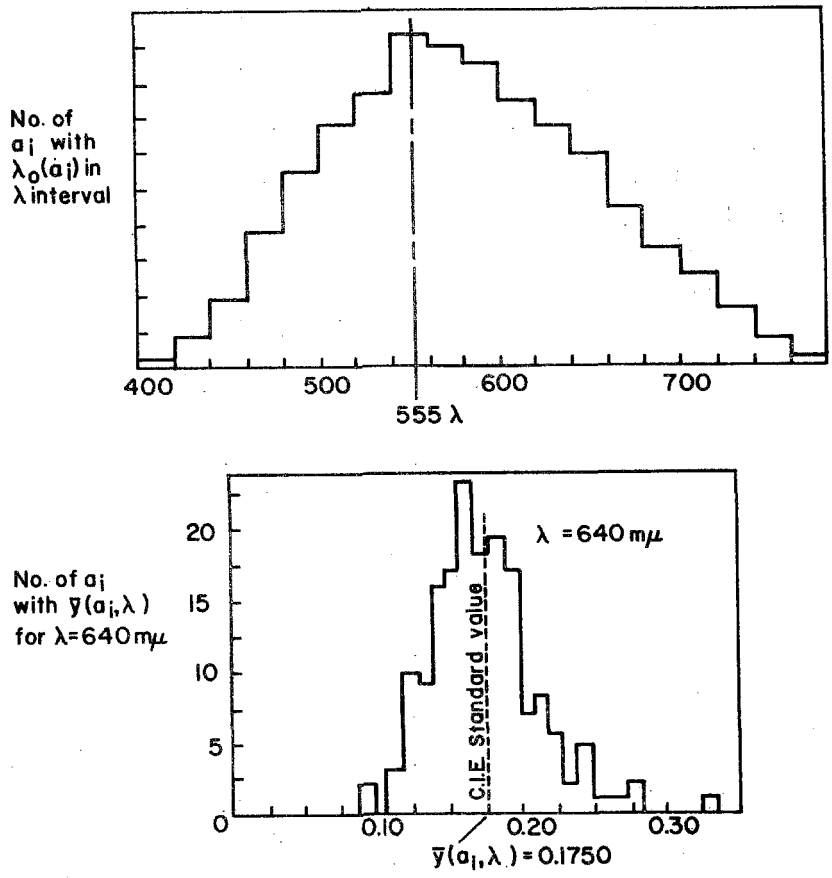


FIG. 2.45 Towards determining the standard luminosity function $\bar{y}$. (From [185], by permission)

for each λ in a selected range of λ's through A--the desired standard luminosity function is obtained. A graph, to scale, of the *standard photopic luminosity function* $\bar{y}(\cdot)$ is given in Fig. 2.46, and a tabulation of $\bar{y}(\cdot)$ is given in Table 1. A more detailed tabulation of the values $\bar{y}(\lambda)$ over the visible spectrum is given in Ref. [50].

Now, all that we did in the preceding discussion by means of individual luminosity functions $\bar{y}(a_i,\cdot)$ can be repeated line for line for the *standard observer* a. Thus, wherever "$\bar{y}(a_i,\cdot)$" appeared, we can write "$\bar{y}(a,\cdot)$" or, more simply, "$\bar{y}(\cdot)$", for the standard photopic luminosity function, and where "a" stands for the hypothetical standard observer (a creature who shares the same corner of conceptual reality with

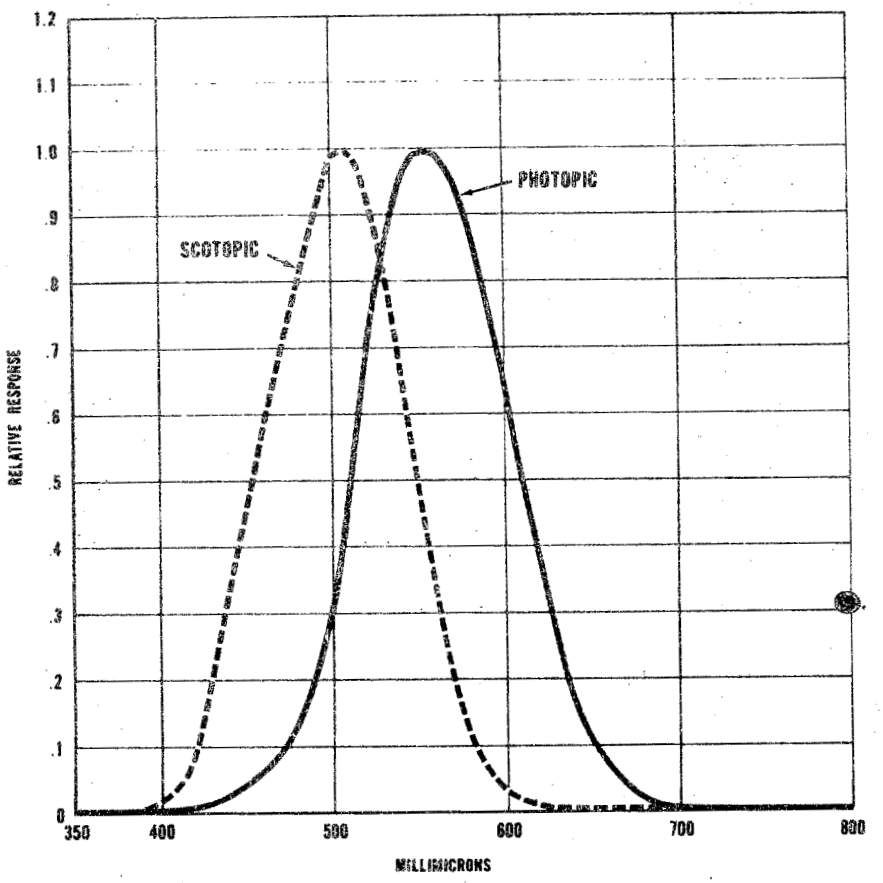


FIG. 2.46 The solid curve depicts the standard photopic luminosity function for daylight adaptation. The standard scotopic luminosity function (for dark adaptation) is shown dashed.

such entities as the "average American male, age 30"). Specifically, we can now make the following definition which is one of the principal definitions of photometry:

Let $\mathscr{R}$ be any radiometric concept of geometrical radiometry (radiance, irradiance, radiant intensity, etc.) defined on the part M of the spectrum Λ. The *relative luminosity of* $\mathscr{R}$ *over* M *for the standard observer* is the number $\overline{Y}(\mathscr{R},M)$ where we have written:

$$\text{"}\overline{Y}(\mathscr{R},M)\text{"} \quad \text{for} \quad \frac{\left[\int_M \mathscr{R}(\lambda)\overline{y}(\lambda)\,d\lambda\right]}{\int_M \mathscr{R}(\lambda)\,d\lambda} \,. \qquad (7)$$

TABLE 1

The Standard Photopic Luminosity Function $\bar{y}(\cdot)$ and its Indefinite Integral

$\lambda(m\mu)$	$\bar{y}(\lambda)$	$\int_{390}^{\lambda} \bar{y}(\lambda')\, d\lambda'$
390	$1. \times 10^{-4}$	0×10^{-3}
400	4.	1
410	12.	5
420	40.	17
430	116.	57
440	230.	173
450	380.	403
460	600.	783
470	910.	1,383
480	1,390.	2,293
490	2,080.	3,683
500	3,230.	5,763
510	5,030.	8,993
520	7,100.	14,023
530	8,620.	21,123
540	9,540.	29,743
550	9,950.	39,283
560	9,950.	49,233
570	9,520.	59,183
580	8,700	68,703
590	7,570.	77,403
600	6,310.	84,973
610	5,030.	91,283
620	3,810.	96,313
630	2,650.	100,123
640	1,750.	102,773
650	1,070.	104,523
660	610.	105,593
670	320.	106,203
680	170.	106,523
690	82.	106,693
700	41.	106,775
710	21.	106,816
720	11.	106,837
730	5.	106,848
740	3.	106,853
750	1.	106,856
760	0.	106,857

Photometric Bedrock: the Lumen

We now take the final step in the transition from radiometry to photometry. This step consists in reaching an agreement on how to assign to every sample of radiant flux of given spectral composition a quantitative measure of the sample's capability of producing within the standard observer the associated sensation of brightness. Now that the concept of the standard observer has been fixed, the remaining task consists in finding a suitable standard radiant-flux emitter whose wavelength dependence over the spectrum Λ is uniquely defined within a rigid experimental and theoretical framework and which is precisely reproducible in practice. Once such a standard is found it is assigned a preselected number of units of "brightness"-producing capability and all other radiant flux samples can then be given their amounts of brightness-producing capability relative to the standard.

Such candidates as laser beams of given monochromaticity, various flames of burning liquids or solids, incandescent gases of known spectral decomposition, the surface of the sun, the surfaces of various molten metals--all these are possible candidates which can serve as photometric standards. The traditional standard was a candle flame--the candle having been manufactured, set to burn, and observed in a rigidly controlled manner. The current standard is the surface of a pool of platinum which is at the precisely determinable temperature (see [51]) of its change from the solid to the liquid phase (2042° Kelvin). Once the metal has reached this temperature within some thermally stable enclosure, its radiant emittance W_b is precisely computable for each wavelength in the spectrum using the laws of blackbody thermal radiation. The surface radiance distribution of the platinum is uniform of magnitude N over all emergent directions from the surface at a point. Hence the relation $N_b = W_b/\pi$ exists between N_b and W_b (cf., closing remarks of Sec. 2.4).

The key step is now taken: it is agreed that the surface radiance of the surface of freezing platinum is to be assigned a brightness sensation producing capability, a *luminance*, of 6×10^5 *lumens per square meter per steradian*. Thus, the unit of brightness-producing capability of radiant flux is called a *lumen*. The lumen is the photometric counterpart to the radiometric *watt*. This convention is translated into practical working formulas as follows: we observe that if $N_b(\lambda)$ is the radiance of the standard platinum surface as given by the blackbody thermal radiation laws, then on the one hand, by (4), the *relative* (standard) luminance of the platinum surface is:

$$\int_0^\infty N_b(\lambda)\bar{y}(\lambda)\, d\lambda \tag{8}$$

and on the other hand, by fiat, the *absolute* (standard) luminance of the platinum surface is:

$$6 \times 10^5 \ \frac{\text{lumens}}{\text{m}^2 \times \text{steradian}} \tag{9}$$

The units of (8) are *watts/(m² × steradian)*. Our agreement leads us to equate (8) and (9), after introducing a numerical constant which will balance units in the resulting equation. Let us denote this number by "K_m". Then we agree to write:

$$6 \times 10^5 = K_m \int_0^\infty N_b(\lambda)\bar{y}(\lambda)\, d\lambda \quad . \tag{10}$$

The number K_m so defined has units: lumens/watt. Its magnitude is determined by explicitly introducing the functional form for the surface radiance N_b of the surface of a blackbody (a *complete* radiator or *Planckian* radiator) at temperature T:

$$N_b(\lambda) = c_1 \lambda^{-5} / \pi (e^{c_2/(\lambda T)} - 1)$$

in which we have set:*

$$c_1/\pi = 2c^2h = 1.1909 \times 10^{-16} \text{ watts m}^2\text{/steradian}$$

$$c_2 = hc/k = 1.4380 \times 10^{-2} \text{ m °(Kelvin)}$$

$$T = 2042° \text{ °(Kelvin)}$$

It follows, on numerical integration of $N_b(\lambda)\bar{y}(\lambda)$ over Λ, that:

$$\int_0^\infty N_b(\lambda)\bar{y}(\lambda)\, d\lambda = 884 \text{ watts/(m}^2 \times \text{steradian)} \quad .$$

Hence, from (10):**

$$K_m = 6 \times 10^5/884$$

$$= 680 \text{ lumens/watt} \quad . \tag{11}$$

---

*The units of c_1 are determined by specifically using the spectral density part of the dimensions of radiance. Thus dim [N] = watts/(m² × steradian × m), using wavelengths.

**Uncertainties in the measured values of c_1, c_2 and in the numerical integrations leading to the value 884 watts/(m² × steradian) lead to a corresponding uncertainty of K_m of about 5 or 6 units in the last digit. See, e.g., [51], [153].

The only seemingly arbitrary feature in this final step from radiometry to photometry is the choice of the magnitude 6×10^5. Actually, the choice of this particular magnitude is not completely arbitrary; it is tied to the historical precedent set for a lumen by the early international candle standard. The historical details of these matters may be found in the standard treatises on photometry, Refs. [185], [311], or in Ref. [50]. See also [206], [51]. The current standard unit of luminous intensity (defined formally below) is the *candela* which by definition is *1 lumen per steradian*. Hence the convention in (9) may be read as *600,000 candelas per square meter*.

Luminance Distributions

The magnitude of the transition factor K_m having been determined, we can go on to give a precise definition of the requisite "measure" of the capability of a given sample of radiant flux to evoke the brightness sensation. Thus let $N(x,\xi,t,\cdot)$ be the radiance function which assigns to every λ in the spectrum Λ a radiance $N(x,\xi,t,\lambda)$ at a fixed point x in the fixed direction ξ at given time t. Then we call the number:

$$K_m \int_0^\infty N(x,\xi,t,\lambda)\bar{y}(\lambda)\, d\lambda$$

the *luminance* associated with the radiance function $N(x,\xi,t,\cdot)$, and write:

$$\boxed{\text{"}B(x,\xi,t)\text{"} \quad \text{for} \quad K_m \int_0^\infty N(x,\xi,t,\lambda)\bar{y}(\lambda)\, d\lambda} \tag{12}$$

If x and t are fixed but ξ allowed to vary in $N(x,\xi,t,\lambda)$ then the resultant function $B(x,\cdot,t)$ is called the *luminance distribution* at x, at time t. Often the time t, or x, or even ξ are understood (as occurred e.g., in the radiometric context) and so may be dropped from the notation provided no confusion results. Thus we agree that we can occasionally write:

$$\text{"}B(x,\xi)\text{"} \quad \text{or} \quad \text{"}B(\xi)\text{"} \quad \text{or "}B\text{"} \quad \text{for} \quad B(x,\xi,t) \quad .$$

These definitions serve to fix $B(x,\cdot)$ as the photometric counterpart to the radiometric function $N(x,\cdot)$ studied in earlier sections of this chapter. The units of $B(x,\cdot)$ are *lumens/(m^2 × steradian)*.

Table 1 of Sec. 2.4 can be used to construct a corresponding table of radiance by assuming that the surfaces S referred to in Table 1 of Sec. 2.4 have *uniform* radiance. Then the desired radiances are found by dividing each irradiance in the right hand column of that table by π. A similar table for the general order of magnitude of *luminance* of common

natural objects can be constructed. A sample of such a table is given below which is partly constructed from Fig. 1.12.

TABLE 2

Source	Luminance (lumens/(m² × steradian))
Surface of sun	2×10^9
Clear sky	3200
Surface of moon*	6200 (full), 1400 (half), 0 (new) } 2800 (average)

Further illustrative examples of luminance are easily constructed: suppose a source of monochromatic radiant flux has a radiance N of 1000 watts/(m² × steradian) per meter wavelength for each λ over an interval $\Delta\lambda$ of 10 mμ centering on wavelength λ = 555 mμ, and of zero radiance outside this interval. What is the luminance B of this source? Returning to (12) we see that in this case:

$$B = K_m \int_0^\infty N(\lambda)\bar{y}(\lambda)\, d\lambda$$

$$= K_m N(555)\bar{y}(555)\Delta\lambda$$

$$= 680 \times 1000 \times 1 \times 10^{-8} \quad (1\ m\mu = 10^{-9} m)$$

$$= 6.8 \times 10^{-3} \text{ lumens/(m}^2 \times \text{ steradian)} \quad .$$

As another example, consider a source of radiance N = 1000 watts/(m² × steradian) per millimicron wavelength at λ = 450 mμ over an interval $\Delta\lambda_1$ of 10 mμ about this wavelength, and of radiance N = 500 watts/(m² × steradian) per millimicron wavelength at λ = 600 mμ over an interval $\Delta\lambda_2$ of 5 mμ about the latter wavelength. What is the associated luminance of this source? By (12) we have:

*These luminances are computed directly from the full and half phase illuminances produced by the moon, as given in Fig. 1.12. For half phase, the solid angle of the luminous surface was taken as 3×10^{-5} steradians. Standard references give 2500-3000 lumens/(m² × steradian) for the moon's luminance. The lighting geometry on the porous and craggy lunar surface is partly involved in this spread of values.

$$B = K_m \int_0^{\infty} N(\lambda)\bar{y}(\lambda)\, d\lambda$$

$$= K_m N(450)\bar{y}(450)\Delta\lambda_1 + K_m N(600)\bar{y}(600)\Delta\lambda_2$$

$$= 680(1000 \times .038 \times 10 + 500 \times .631 \times 5)$$

$$= 680(380 + 1580)$$

$$= 1.34 \times 10^6 \text{ lumens}/(\text{m}^2 \times \text{steradian}) \quad .$$

Here $\bar{y}(450)$ and $\bar{y}(600)$ may be estimated from Fig. 2.46 or taken directly from Table 1. Note the two different ways of specifying the spectral density of radiance in these two examples.

As a final example, let a source be of constant radiance N (per millimicron wavelength) over the region of the spectrum from 390 mμ to 760 mμ (the part of the spectrum over which $\bar{y}(\lambda)$ is defined in Table 1, and zero outside this region. What is the luminance of the source? By (12) we have:

$$B = K_m \int_{390}^{760} N\bar{y}(\lambda)\, d\lambda$$

$$= K_m N \int_{390}^{760} \bar{y}(\lambda)\, d\lambda$$

$$= 680 \times 107 \times N = 7.3 \times 10^4\, N$$

Here we have integrated $\bar{y}(\cdot)$ over the range $\lambda = 390$ mμ, to $\lambda = 760$ mμ in steps of 10 mμ using the values of Table 1 above. The result is:

$$\int_{390}^{760} \bar{y}(\lambda)\, d\lambda = 106.857 \quad \text{(millimicrons)}$$

This value may, for all practical purposes, be taken as the integral of $\bar{y}(\cdot)$ from $\lambda = 0$ to $\lambda = \infty$.

Transition to Geometrical Photometry

The transition from geometrical radiometry to geometrical photometry has so far been made between two points, i.e., between the radiance and luminance concepts by means of (12), and with the help of (10) and (11). This choice of the radiance-luminance bridge rather than any other means was governed

by the relative visualizability of these concepts as contrasted with other radiometric-photometric pairs, say with the visualizability of hemispherical irradiance and its counterpart hemispherical illuminance (to be defined below). But now that the bridge has been constructed with suitable attention to intuitive motivations and visualization, we return to its site and start anew with the purpose in mind of constructing the bridge once again, but now in a logically more satisfying way. By undertaking this reconstruction we are given the opportunity to re-emphasize and make formal the additivity assumption we had encountered on our way to the relative luminance distribution in (4). This formalized additivity assumption will subsequently take its place among the other basic assumptions of radiometry which we isolated for the radiant flux function in the discussions of Sec. 2.3.

The transition from radiance to luminance, as summarized in (12), may now be emulated systematically for each radiometric concept. That is, for every part M of the spectrum Λ we first define a general integral the *radiometric-photometric transition operator* by writing:

$$\text{"}Y(\cdot,M)\text{"} \quad \text{for} \quad K_m \int_M [\ ]\bar{y}(\lambda)\, d\lambda \quad . \tag{13}$$

Then it follows from (12) that:

$$B = Y(N,\Lambda) \quad , \tag{14}$$

where "B" and "N" are the abbreviated names for the given luminance and radiance functions in (12). But we need not stop at (12). Indeed, let us go on and write:

$\text{"}F^{\pm}(S,D,t)\text{"}$	for	$Y(P^{\pm}(S,D,t,\cdot),\Lambda)$	(*luminous flux*, (3) of Sec. 2.3; cf.,(17) of Sec. 2.4)	(15)
$\text{"}E(x,\xi,t)\text{"}$	for	$Y(H(x,\xi,t,\cdot),\Lambda)$	(*illuminance* (11), (17) of Sec. 2.4)	(16)
$\text{"}L(x,\xi,t)\text{"}$	for	$Y(W(x,\xi,t,\cdot),\Lambda)$	(*luminous emittance*, (22) of Sec. 2.4)	(17)
$\text{"}B^{\pm}(x,\xi,t)\text{"}$	for	$Y(N^{\pm}(x,\xi,t,\cdot),\Lambda)$	(*luminance*, (30), (31) of Sec. 2.5)	(18)
$\text{"}I^{\pm}(S,\xi,t)\text{"}$	for	$Y(J^{\pm}(S,\xi,t,\cdot),\Lambda)$	(*luminous intensity*, (7),(10) of Sec. 2.9)	(19)

These are the definitions of the first five principal photometric concepts under both the surface (+) and field (-) interpretations. The names of the concepts are given to the right of each definition and reference is made to the appropriate radiometric ancestor of each concept. Thus, e.g., surface luminous flux $F^{+}(S,D,t)$ is derived from surface radiant flux

$P^+(S,D,t)$, which in turn is defined in (3) of Sec. 2.3, and with the surface interpretation of $P^+(S,D,t)$ given in (17) of Sec. 2.4. In this way the logical ancestry of each of the preceding eight photometric concepts is traceable back to the primitive radiometric function Φ. The bridge to the ancestor in each case is the integral operator $Y(\cdot,\Lambda)$, defined in (13). This integral operator now permits, without the necessity of any further geometrical arguments, all the radiometric connections among P, H, W, N and J developed in the preceding sections, to be carried over directly to the photometric context. As an example, (8) of Sec. 2.5 is carried over to the photometric context by applying $Y(\cdot,\Lambda)$ to each side of that equation. Thus, by (16) above, (31) of Sec. 2.5, and (8) of Sec. 2.5:

$$E(x,\xi) = Y(H(x,\xi),\Lambda) = Y\left(\int_{\Xi(\xi)} N^-(x,\xi)\xi\cdot\xi'\, d\Omega(\xi'),\Lambda\right)$$

$$= \int_{\Xi(\xi)} Y(N^-(x,\xi'),\Lambda)\xi\cdot\xi'\, d\Omega(\xi')$$

$$= \int_{\Xi(\xi)} B^-(x,\xi')\xi\cdot\xi'\, d\Omega(\xi') \qquad (20)$$

Whenever either "+" or "-" is understood, or an equation is valid under both the field and surface interpretations, then these signs may be dropped, if desired. For example, in the case of (20), we know from (21) of Sec. 2.5 that H and N^- go together, so that dropping "-" on the right sides of the equations in (20), no confusion can result. Hence, every occurrence of the signs "-" may be dropped from (20) and left implicitly understood.

The roll-call of principal photometric concepts is continued as follows. We shall write:

$Q^{\pm}(X,t)$"	for	$Y(U^{\pm}(X,t,\cdot),\Lambda)$	(*luminous energy* in region X, at time t, (12) of Sec. 2.7)	(21)
"$Q^{\pm}(S,T)$"	for	$Y(U^{\pm}(S,T,\cdot),\Lambda)$	(*luminous energy* across surface S over time interval T, (17) of Sec. 2.7)	(22)
"$q^{\pm}(x,t)$"	for	$Y(u^{\pm}(x,t,\cdot),\Lambda)$	(*luminous energy density*, (2) of Sec. 2.7)	(23)
"$e(x,t)$"	for	$Y(h(x,t,\cdot),\Lambda)$	(*scalar illuminance*, (3) of Sec. 2.7)	(24)

"l(x,t)" for $Y(w(x,t,\cdot),\Lambda)$ (*scalar luminous emittance*, (19) of Sec. 2.7) (25)

We illustrate again the fact that any linear relation between two radiometric quantities has a carbon copy in the photometric context. Thus, consider (14) of Sec. 2.7; assuming v is independent of λ in X and applying the operator $Y(\cdot,\Lambda)$ to each side we have:

$$\begin{aligned} Q(X,t) = Y(U(X,t,\cdot),\Lambda) &= Y((u/v)V(X),\Lambda) \\ &= (V(X)/v)Y(u,\Lambda) \\ &= (q/v)V(X) \quad . \end{aligned} \tag{26}$$

There remains to be defined certain of the photometric concepts such as the vector counterpart **E** to **H**,**I** to **J**, etc. However, instead of going on to explicitly exhaust all these transitions, which are quite numerous, we state below a general definition-scheme which covers all transitions just made, and any yet unmade.

Let $\mathscr{R}$ be a radiometric function defined on Λ. Then $Y(\mathscr{R},\Lambda)$ is the *photometric counterpart* to $\mathscr{R}$. Let "$\mathscr{P}$" denote this photometric counterpart. Then the following statement is a definitional identity:

$$\boxed{\mathscr{P} = Y(\mathscr{R},\Lambda)} \tag{27}$$

A *definitional identity* is a statement of the form "A = B" where "A" and "B" are the names of one and the same object arising from a definition. Thus, e.g., "Q(X,t)" and "$Y(U(X,t,\cdot),\Lambda)$" are names of one and the same object, namely the number:

$$\int_0^\infty U(X,t,\lambda)\bar{y}(\lambda)\, d\lambda$$

and so:

$$Q(X,t) = Y(U(X,t,\cdot),\Lambda)$$

and alternatively:

$$Q(X,t) = \int_0^\infty U(X,t,\lambda)\bar{y}(\lambda)\, d\lambda$$

are definitional identities. For example, definitional identities were used to start and end the series of deductions summarized in (20) and (26). The significance of (27) is

simply this: every photometric concept $\mathscr{P}$ is the image, under $Y(\cdot,\Lambda)$, of some radiometric concept $\mathscr{R}$; thus, to define a new photometric concept, first find its radiometric progenitor $\mathscr{R}$. Then $\mathscr{P}$ is the desired concept $Y(\mathscr{R},\Lambda)$.

General Properties of the Radiometric-Photometric Transition Operator

The integral operator $Y(\cdot,M)$ defined in (13) has several properties built into it which are of critical importance in establishing the science of theoretical photometry. To recognize and understand these properties is to recognize and understand the role of photometry as a descriptive science. Therefore we devote some attention to the isolation of these properties.

Let $\mathscr{R}_1$ and $\mathscr{R}_2$ be any two radiometric functions defined on a subset M of Λ and let c_1 and c_2 be any two real numbers such that the sum $(c_1\mathscr{R}_1 + c_2\mathscr{R}_2)$ is defined. Then by (13) and the linearity of the mathematical integration process:

$$Y(c_1\mathscr{R}_1 + c_2\mathscr{R}_2,\Lambda) = c_1Y(\mathscr{R}_1,\Lambda) + c_2Y(\mathscr{R}_2,M) \tag{28}$$

This is the *linearity property* of $Y(\cdot,M)$, the formal vestige of the associated property of $\overline{y}(a_i,\lambda)$ discussed in (2).

Next, for every radiometric function $\mathscr{R}$ defined on Λ,

$$Y(\mathscr{R},M_1 \cup M_2) = Y(\mathscr{R},M_1) + Y(\mathscr{R},M_2) \tag{29}$$

for every pair of disjoint subsets M_1 and M_2 of Λ. This is the *additive property* of $Y(\mathscr{R},\cdot)$ and is the formal vestige of the property of $\overline{y}(a_i,\lambda)$ discussed in (3). Finally, for every radiometric function $\mathscr{R}$ defined on Λ,

$$\text{If} \quad l(M) = 0, \quad \text{then} \quad Y(\mathscr{R},M) = 0 \tag{30}$$

which is the M-*continuity property* of $Y(\mathscr{R},\cdot)$ for continuous spectra. The length measure l and its general use was defined in (4) of Sec. 2.3.

The Mathematical Basis for Geometrical Photometry

Properties (29) and (30) may be added to the set of six additivity and continuity properties of Φ discussed in Sec. 2.3. In fact, in an axiomatic development of the mathematical theory of photometry, statements (28), (29) and (30)

would constitute the essential starting point of the construction of the theory, just as the properties of Φ in Sec. 2.3 constitute the essential starting point of the theory of geometrical radiometry. Indeed, for any radiometric function defined on Λ, we may deduce from (29) and (30) alone the existence of a function $\bar{y}'(\cdot)$ on Λ such that:

$$Y(\mathscr{R},\Lambda) = \int_{\Lambda} \mathscr{R}(\lambda)\bar{y}'(\lambda)\, d\lambda \quad . \tag{31}$$

Evidently $\bar{y}'(\cdot)$ will turn out to be $K_m\bar{y}(\cdot)$ discussed above. The complete details of the mathematical justification of this assertion lie beyond the scope of this work. Some of the mathematical background of (31) will be covered as a matter of course in Sec. 3.16. The requisite mathematical basis of the assertion may be found in part in Sec. 56, in particular theorem D, of Ref. [103]. The general measure-theoretic approach to foundations of radiative transfer theory, introduced in Ref. [216], can now, by (31), be systematically extended to the domain of photometry. Hence, as far as the mathematical structure of photometry is concerned, it rests on three pillars: (28), (29), and (30), and its framework can be erected by means of the theorems of modern measure theory and without the necessity of any further reference to physical constructs. In other words, the epistemological content of classical photometry rests in but three postulates, the statements of the linearity, M-additivity and M-continuity of Y introduced above. We note in closing that the preceding observations apply immediately to the representations of colors by the tristimulus procedure of colorimetry; all that has been said for the function $\bar{y}$, now applies, without essential change, to the other two tristimulus functions $\bar{x}$ and $\bar{z}$ (cf., Sec. 1.7). The mathematical setting in the colorimetric case would be a three-dimensional vector space, and the measure-theoretic aspects will be elevated from the scalar to the vector level.

Summary and Examples

The present discussion of radiometry and photometry will be brought to a close with a summary of the main concepts introduced in this chapter. The units and dimensions of the concepts will be tabulated, discussed and illustrated, and a few further illustrative examples will be given.

Table 3 lists the main radiometric concepts by name, symbol, units, dimensions, and reference to its definition in the present work. A similar Table 4 lists the main photometric concepts in an exactly analogous way, as far as possible. Explanatory notes are appended to each table.

TABLE 3

RADIOMETRIC CONCEPTS

NAME	BASIC SYMBOL	DIMENSIONS	MKS UNITS	DEFINITION REFERENCES
RADIANT FLUX (general)	Φ	$\mathbf{P}^{\pm}$	WATT	Sec. 2.1; (17) and (18) of Sec. 2.4
RADIANT FLUX (spectral)	P	$\mathbf{P}_{\lambda}^{\pm}$	WATT/mμ	(3) of Sec.2.3 (17) and (18) of Sec. 2.4
RADIANCE (all radiometric concepts here and below may be either *general* or *spectral*)	N	$\mathbf{P}^{\pm}\mathbf{A}^{-1}\Omega^{-1}$	WATT/(m²×sr) (See note (c) below)	(1), (4), of Sec. 2.5
IRRADIANCE	H	$\mathbf{P}^{-}\mathbf{A}^{-1}$	WATT/m²	(1), (17) of Sec. 2.4
VECTOR IRRADIANCE	**H**	$\mathbf{P}^{-}\mathbf{A}^{-1}$	WATT/m²	(2) of Sec.2.8
SCALAR IRRADIANCE	h	$\mathbf{P}^{-}\mathbf{A}^{-1}$	WATT/m²	(3) of Sec.2.7
RADIANT EMITTANCE	W	$\mathbf{P}^{+}\mathbf{A}^{-1}$	WATT/m²	(18), (22) of Sec. 2.4
VECTOR RADIANT EMITTANCE	**W**	$\mathbf{P}^{+}\mathbf{A}^{-1}$	WATT/m²	See note (d) below
SCALAR RADIANT EMITTANCE	w	$\mathbf{P}^{+}\mathbf{A}^{-1}$	WATT/m²	(19) of Sec. 2.7
RADIANT INTENSITY	J	$\mathbf{P}^{\pm}\Omega^{-1}$	WATT/sr	(1), (10) of Sec. 2.9
VECTOR RADIANT INTENSITY	**J**	$\mathbf{P}^{\pm}\Omega^{-1}$	WATT/sr	(22) of Sec. 2.9
SCALAR RADIANT INTENSITY	j	$\mathbf{P}^{\pm}\Omega^{-1}$	WATT/sr	See note (d) below

TABLE 3 (Continued)

NAME	BASIC SYMBOL	DIMENSIONS	MKS UNITS	DEFINITION REFERENCES
RADIANT ENERGY	U	$P^{\pm}T$	WATT-SECOND or JOULE	(12), (17) of Sec. 2.7
RADIANT DENSITY	u	$P^{\pm}TV^{-1}$	WATT-SECOND/m^3 or JOULE/m^3	(2) of Sec.2.7
(RADIANT) PATH FUNCTION	N_*	$P^{\pm}V^{-1}\Omega^{-1}$	WATT/(m^3×sr) or HERSCHEL/m	(2) of Sec.3.12 (8) of Sec.3.14 (3) of Sec.13.3
PATH RADIANCE	N^*	$P^{\pm}A^{-1}\Omega^{-1}$	WATT/(m^2×sr) or HERSCHEL	(1) of Sec.3.12 (15) of Sec.3.12 (2) of Sec.13.3

Explanatory Notes for Table 3

(a) The names and basic symbols are drawn, as far as possible, from the current standard in nomenclature, namely that recommended in 1953 by the American Standards Association Section Committee Z-58, sponsored by the Optical Society ([4], [49], also cf., p. 229, Ref. [50]). The basic symbols are used to construct names for various radiometric functions by placing various modifiers after them. Thus, e.g., $\Phi(S,D,t,F)$ is the value of the function Φ which assigns to each set F of frequencies the radiant flux incident on collecting surface S through the set D of directions at time t. Further examples are found throughout the preceding sections of this chapter. It might be well to observe here that the symbols and names for the concepts in such a venerable subject as geometrical radiometry are still in a state of change. However, there is currently some effort being made in the direction of establishing an international standard of terminology in radiometry and photometry (see, e.g., Ref. [130]). It may be noted that the terminology and notation listed in Tables 3 and 4 have withstood the severe tests of use in courses and research studies by the author and his colleagues over the past twenty years, and have been found adequate for the purposes of radiative transfer studies in natural optical media. (See also p. 6, [177].)

It now appears possible to attain a systematic and basic terminology for radiometry and photometry by combining the best features of Table 3 and Table 4 (below) and the suggestions by Jones in Ref. [130]. Toward this end we observe that

Jones extracts the idea of *flux*, as the basic concept whose task is to describe the flow of a generalized 'substance'. The 'substance' may be radiant energy, luminous energy (the photometric counterpart to radiant energy) or even entropy. The suggested term for the 'ometry' which studies general flows is 'phluometry' ("phluo" = "to flow"). There are five such 'ometries' suggested at present:

Name of the Phluometry	Phluometric Modifier	Unit of Flux
Radiometry	Radiant	Watt
Photometry	Luminous	Lumen
Ergometry	Energic	Joule
Ergophotometry	Ergolumic	Lumen-second
Entropometry	Entropic	Watt/degree

(b) The basic radiant flux dimensions P^+, P^- are associated with flux leaving and incident on a surface, respectively. The idea of 'radiant flux' is the central physical idea of geometrical radiometry. However, it is found useful in theory and practice to distinguish between emitted and incident radiant flux. This distinction has been placed into the dimensions for appropriate use, if needed, and its geometrical significance is summarized in Fig. 2.47. (See also Fig. 2.12.) If the distinction is not needed, or is understood, the occurrences of "+" or "-" may be omitted. Further discussion of dimensions is made in note (h).

(c) The spectral radiant flux P has units of WATT/mμ if wavelengths in millimicrons (mμ) are used, or has units of

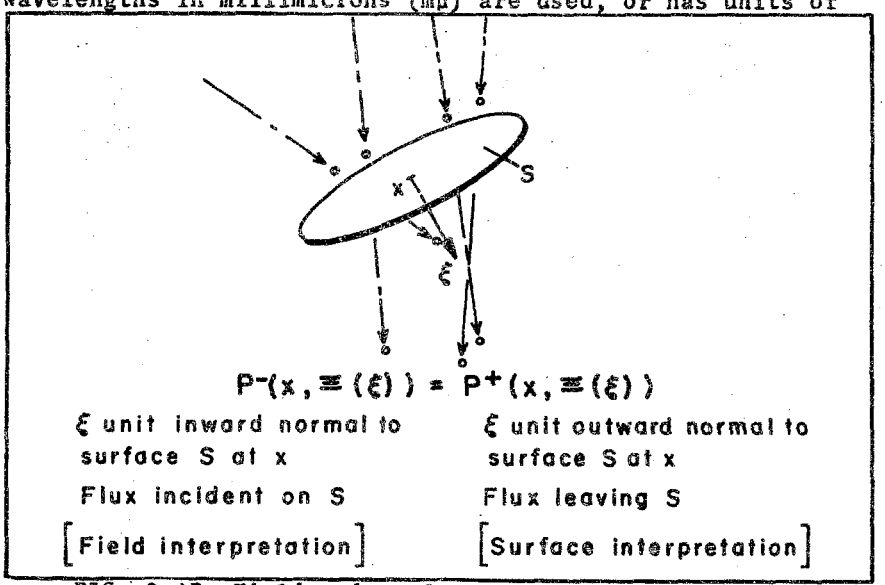


FIG. 2.47 Field and surface interpretation of radiant flux.

WATT/sec^{-1}, if frequency is in units of sec^{-1}. The dimensions "WATT/mμ" are used often in practice; hence their inclusion in the table. The radiant flux dimension **P** of any radiometric quantity below P in the table can be either spectral (hence $\mathbf{P}_\lambda$) or general (hence **P**). For simplicity, only the general radiant flux dimension is given. When working with spectral radiant flux, it is occasionally necessary to explicitly use, during a given discussion, both wavelength and frequency dimensions for radiant flux. The radiometric quantities can then be given a "λ" or a "ν" subscript for the duration of such discussions. In general, however, such explicitness is not needed and the dimension of the spectral flux is understood implicitly, and (except for specific numerical examples) will so be understood throughout this work. In theoretical radiative transfer discussions, e.g., the frequency dimension is usually preferred over wavelength (and this preference is implicit in the notation) because frequency of radiation is invariant along a path with variable index of refraction. The general (definitional) connection between P_ν and P_λ is obtained by writing:

$$\text{"}P_\nu\text{"} \quad \text{for} \quad \frac{d\Phi_\nu}{d\nu}$$

$$\text{"}P_\lambda\text{"} \quad \text{for} \quad \frac{d\Phi_\lambda}{d\lambda} \quad .$$

Whence:

$$P_\lambda = P_\nu \left|\frac{d\nu}{d\lambda}\right| = \frac{\nu}{\lambda^2} P_\nu \quad . \tag{32}$$

(d) Table 3 is divided into five natural groupings of concepts. First in order are the three main concepts--Φ, P, N. Then comes the irradiance group, the radiant emittance group, and the radiant intensity group. These are followed by the energy group, and the radiative transfer group consisting of N_* and N^*. In principle, the irradiance group and the radiant emittance groups may be coalesced into a single group by using explicitly the surface (+) and (-) concepts. However, historical precedent has fixed the distinction between these groups by means of the generic letters "H" and "W", and we see no reason at the present time to change such established notation to "H^+" for W and "H^-" for H merely on the grounds of esthetic reasons. However, esthetic reasons (in particular the desire for symmetry) are responsible for the inclusion of two concepts in Table 3 which--if the practical photometric worker had a say--would normally be omitted. These are the two concepts **W** and **j**. The distinction between **W** and **H** is very fine conceptually and non-existent vectorially. For we define **W**(x) as follows. We write:

$$\text{"}\mathbf{W}(x)\text{"} \quad \text{for} \quad \int \xi N^+(x,\xi)\, d\Omega(\xi) \tag{33}$$

where, as noted, the integral uses the surface radiance. We call **W**(x) the *vector radiant emittance* at x. Thus (33), by

(32) of Sec. 2.5, and (2) of Sec. 2.8, yields the equality $W(x) = H(x)$. Further, whenever S is a surface and N(x) is a uniform radiance distribution at x on S (either field or surface) then we may write:

$$\text{"j(S)"} \quad \text{for} \quad \int_S N(x)\, dA(x) \tag{34}$$

and call j(S) the *scalar radiant intensity*. By including (33) and (34) we round out each of the radiant emittance and radiant intensity families to a full threesome, and maintain the interesting duality between irradiance and radiant intensity brought out in the main discussions above (where j(S) is now the dual to h(x)).

(e) The unit name "herschel" for radiance is adopted from a suggestion by Moon (Ref. [184]). However, the unit, as used here, is left "unrationalized." This means simply that for uniform radiance distribution, we have $H = \pi N$. Hence if $N = 1$ herschel, then $H = \pi$ watts/m^2, and numerical computations are not jeopardized by not remembering what to do with "π". Furthermore, the π serves to keep tabs on the dimensions of H and N in calculations. It is clear that something other than the relatively lengthy "WATT/(m^2 x sr)" is desirable, at least in verbal discussions, where "sr" stands for "steradian", and "m^2" as usual denotes "square meters".

(f) The final group consisting of path radiance N^* and path function N_* is included for convenience of reference. These are the only two additional radiometric concepts needed in the general studies of radiative transfer in natural optical media. Actually these concepts are mutually dependent and only one is needed. The full discussion of this matter is reserved for Chapter III.

(g) The only radiometric concepts omitted from Table 3 and which are of some importance, are the spherical and hemispherical scalar irradiances defined in Sec. 2.7. These concepts, especially the latter, are primarily indigenous to plane-parallel (or one parameter) geometries, whereas all the listed concepts pertain to general geometries. Not defined at all were the spherical and hemispherical scalar emittances. For the sake of completeness (cf., (6) of Sec. 2.7), we write:

$$\text{"}w_{4\pi}(x,t)\text{"} \quad \text{for} \quad \frac{1}{4}\, w(x,t) \tag{35}$$

and (cf. (7) of Sec. 2.7):

$$\text{"w}(x,\xi,t)\text{"} \quad \text{for} \quad \int_{\Xi(\xi)} N(x,\xi',t)\, d\Omega(\xi') \tag{36}$$

and (cf. (8) of Sec. 2.7):

$$\text{"}w_{4\pi}(x,\xi,t)\text{"} \quad \text{for} \quad \frac{1}{4}\, w(x,\xi,t) \tag{37}$$

and these are called, respectively, *spherical radiant*

emittance, hemispherical scalar radiant emittance, and *hemispherical radiant emittance.*

(h) The theory of dimensions of radiometric and photometric concepts has received relatively little systematic attention. We shall devote a few comments to this matter in the present note. The dimensional system chosen for Tables 3 and 4 is constructed from two basic physical dimensions and one basic geometrical dimension. These are the dimensions of radiant flux **P**, time **T**, and length **L**. The general radiant flux function Φ is assigned the dimension **P**; *this dimension is considered irreducible in the radiometric context.* In other contexts, **P** need not be irreducible. Thus, in the electromagnetic context **P** is representable in terms of the dimensions of force, length and time: $(\text{force}) \times (\text{length}) \times (\text{time})^{-1}$, or as $(\text{mass}) \times (\text{length})^2 \times (\text{time})^{-3}$. The "+" and "-" superscripts on "**P**" do not change its dimension; they merely serve as convenient mnemonics for the surface and field interpretations of radiant flux.

As already made clear in note (c) above, the dimensions $\mathbf{P}_\nu$ or $\mathbf{P}_\lambda$ are reducible to $\mathbf{PT}$ or $\mathbf{PL}^{-1}$, respectively. Specifically, in Table 3 we have implicitly written:

$$\text{"}\mathbf{P}_\nu{}^{\pm}\text{"} \quad \text{for} \quad \mathbf{P}^{\pm}\mathbf{T}$$

and for the wavelength case we have explicitly written:

$$\text{"}\mathbf{P}_\lambda{}^{\pm}\text{"} \quad \text{for} \quad \mathbf{P}^{\pm}\mathbf{L}^{-1} \quad .$$

Now just as we find it convenient to append "+" and "-" to the basic symbol "**P**" to denote the geometric sense of the flow of radiant flux, so too is it helpful to distinguish between two types of length in geometrical radiometric discussions. Following Moon [184], we write: "$\mathbf{L}_t$" to denote the dimension of length measured in a direction transverse (i.e., perpendicular) to a given direction ξ; and "$\mathbf{L}_r$" to denote the dimension of length along the given (radial) direction ξ. As in the case of $\mathbf{P}^{\pm}$, attaching "t" and "r" to "**L**" does not change the dimension; rather it serves as a conceptual reminder of the transverse and radial interpretations of length. Then in the table we have written:

$$\text{"}\mathbf{A}\text{"} \quad \text{for} \quad \mathbf{L}_t{}^2$$

$$\text{"}\mathbf{V}\text{"} \quad \text{for} \quad \mathbf{L}_t{}^2\mathbf{L}_r$$

$$\text{"}\Omega\text{"} \quad \text{for} \quad \mathbf{L}_t{}^2\mathbf{L}_r{}^{-2}$$

Thus in the present dimensional system, area has the dimensions of transverse length squared--a most natural dimension within radiometry since we perceive areas as two-dimensional extensions of space in the transverse directions to a line of sight. Volume has dimensions of $\mathbf{AL}_r$, i.e., (transverse) area times (radial) length--again a most natural combination of

dimensions for the radiometrist. Finally, solid angles are measured using the steradian concept. In the present system the dimension of solid angles does not vanish from view, but rather is expressed as the product of L_t^2 and L_r^{-2}, as indicated above. Since L_t and L_r are conceptually distinct, this product is conceptually not dimensionless. In this way the occasionally bothersome problem of the vanishing dimensions of solid angle can be solved. (Ordinarily the dimensions vanish, but like the smile of the Cheshire cat, the units remain.) It should be noted that the conventional dimensions of the radiometric concepts are recovered by dropping "+", "-", "t", and "r" from P and L wherever they occur.

As an illustration of the use of these dimensions, observe that the dimension of path function can be written as $(P^{\pm}A^{-1}\Omega^{-1})L_r^{-1}$, so that the path function concept is seen to have the dimensions of radiance per unit of radial length. The full significance of this interpretation will become clear in Sec. 3.12, wherein the path function concept is formally introduced. On the other hand, we may rearrange the path function dimensions as follows: $(P^{\pm}\Omega^{-1})V^{-1}$, and thereby discern another facet of this concept, namely that it may be viewed as a radiant intensity per unit volume (cf. (7), (10) of Sec. 13.6). The radiance concept itself may be viewed via the dimensional arrangement $(P^{\pm}A^{-1})\Omega^{-1}$ as irradiance (-) or radiant emittance (+) per unit solid angle on the one hand, and via the arrangement $(P^{\pm}\Omega^{-1})A^{-1}$ as field (-) or surface (+) radiant intensity per unit area, on the other hand.

A general guide to the fixing of dimensions of radiometric concepts and their manifold derivates in practice is as follows. Let us refer to "area", "length", "time", etc. by the generic term "measure", and use the generic symbol "m" for a measure. Let us write "dim(m)" for the dimension of m. Thus if A is an area measure, then A(S) is the area of a surface S, and $\dim(A) = L_t^2$. Further, if l is a length measure along paths of sight, then l(p) is the length of a path p, and $\dim(l) = L_r$. Now, according to our development of geometrical radiometry in this chapter, every radiometric concept $\mathscr{R}$ is definable first on the empirical level and then on the theoretical level. The empirical level of definition is simply the level on which the measures are used directly. Thus, e.g., recall that empirical irradiance H(S,D) is P(S,D)/A(S), i.e., the quotient of incident radiant flux over a surface S by the area of S. The corresponding theoretical definition is obtained by going to the appropriate limit (e.g., $S \to \{x\}$ in the case of irradiance). In going from the empirical level to the theoretical level, it is desirable to have the dimensions remain unchanged. Hence the definition on the empirical level already fixes the dimension of a radiometric concept. Suppose then that $\mathscr{R}$ is a radiometric concept and its empirical definition is such that we write:

$$\text{"}\mathscr{R}\text{"} \quad \text{for} \quad \frac{\Phi m_1' \ldots m_a'}{m_1 \ldots m_b}$$

where "m_i'", $i = 1,\ldots,a$, and "m_j", $j = 1,\ldots,b$, denote measures and "Φ" denotes the radiant flux function, which is also a measure with dimension $\dim(\Phi) = P$. Then the *dimension of*

Ris:

$$\frac{(\dim \Phi) \times \dim(m_1') \times \ldots \times \dim(m_a')}{\dim(m_1) \times \ldots \times \dim(m_b)}$$

The preceding reduction of a dimension to simpler terms is facilitated by adopting the following conventions for the dimention operator dim. Let x and y be any two measures or physical concepts. Then:

(i) $\dim(xy) = \dim(x)\ \dim(y)$

(ii) $\dim(x/y) = \dim(x)/\dim(y)$

(iii) If $\dim(x) = \dim(y)$, then $\dim(x) = \dim(x+y)$

(iv) If $\{x_n\}$ is a sequence of terms of common dimension d, and if $\lim_n x_n = y$, then $\dim(y) = d$.

In our development of radiometry, the basic dimensions are **P**, **L**, and **T**. In order to use rules (i)-(iv), we agree that these dimensions obey the same rules of addition and multiplication as real numbers. This is implicitly assumed in the tables and in the various manipulations above. In addition to the four dimensions above, we introduce one more, namely **1**, which has the property that:

$$\mathbf{d1} = \mathbf{1d} = \mathbf{d}$$

for every dimension **d**, and

$$\mathbf{d}_1/\mathbf{d}_2 = \mathbf{1}$$

for every pair of dimensions $\mathbf{d}_1$ and $\mathbf{d}_2$ such that $\mathbf{d}_1 = \mathbf{d}_2$. Thus "**1**" denotes the dimensionless concept.

Explanatory Notes for Table 4

(a) The notes and comments for Table 3 apply also to this table except where explicit references to frequency or wavelength concepts are made. Observe that Tables 3 and 4 correspond item for item, except that there is naturally no luminous counterpart to the general radiant flux function Φ, the primitive radiometric function from which all others spring. The unit of luminance, the (unrationalized) *blondel*, is adapted from a suggestion by Moon (ref. [184]). The luminous counterparts to (35)-(37) are obtained by means of the general definition scheme of (27). In (35) and (37) "w" is replaced by "l", and "radiant" replaced by "luminous", to effect the definitions. We assign to the lumen the basic dimension **F**. Hence, in particular, $\dim(K_m) = \mathbf{FP}^{-1}$. By (15) and property (iv) of the operator dim in note (h) for Table 3, we have, e.g., $\dim(F^{\pm}(S,D,t)) = \mathbf{F}^{\pm}$. In this case, the limit

TABLE 4

PHOTOMETRIC CONCEPTS

NAME	BASIC SYMBOL	DIMENSIONS	MKS UNITS	DEFINITION REFERENCES
LUMINOUS FLUX	F	$F^{\pm}$	LUMEN	(15)
LUMINANCE	B	$F^{\pm}A^{-1}\Omega^{-1}$	LUMEN/(m^2× sr)	(18)
ILLUMINANCE	E	$F^{-}A^{-1}$	LUMEN/m^2	(16)
VECTOR ILLUMINANCE	**E**	$F^{-}A^{-1}$	LUMEN/m^2	(2) of Sec. 2.8 and (27)
SCALAR ILLUMINANCE	e	$F^{-}A^{-1}$	LUMEN/m^2	(24)
LUMINOUS EMITTANCE	L	$F^{+}A^{-1}$	LUMEN/m^2	(17)
VECTOR LUMINOUS EMITTANCE	**L**	$F^{+}A^{-1}$	LUMEN/m^2	(33), (27)
SCALAR LUMINOUS EMITTANCE	l	$F^{+}A^{-1}$	LUMEN/m^2	(25)
LUMINOUS INTENSITY	I	$F^{\pm}\Omega^{-1}$	LUMEN/sr or CANDELA	(19)
VECTOR LUMINOUS INTENSITY	**I**	$F^{\pm}\Omega^{-1}$	LUMEN/sr or CANDELA	(22) of Sec. 2.9 and (27)
SCALAR LUMINOUS INTENSITY	i	$F^{\pm}\Omega^{-1}$	LUMEN/sr or CANDELA	(34), (27)
LUMINOUS ENERGY	Q	$F^{\pm}T$	LUMEN-SECOND or TALBOT	(21), (22)
LUMINOUS DENSITY	q	$F^{\pm}TV^{-1}$	LUMEN-SECOND/m^3 or TALBOT/m^3	(23)
(LUMINOUS) PATH FUNCTION	B_*	$F^{\pm}V^{-1}\Omega^{-1}$	LUMEN/(m^3× sr) or BLONDEL/m	(2) of Sec. 3.12 and (27)
PATH LUMINANCE	B^*	$F^{\pm}A^{-1}\Omega^{-1}$	LUMEN/(m^2× sr) or BLONDEL	(1) of Sec. 3.12 and (27)

operation is that used in the definition of the integral operator (13).

We conclude this discussion of the photometric concepts with a few examples.

Example 1. Using the luminance of the sun as given in Table 2, compute the corresponding illuminance on a plane normal to the rays of the sun. To find the requisite illuminance, recall from (2) of Sec. 2.5 that we can write:

$$H(S,D) = N(S,D)\Omega(D) \quad .$$

Applying the transition operator $Y(\cdot,A)$, as defined in (13), to each side of this equation, we obtain:

$$Y\big(H(S,D),A\big) = Y\big(N(S,D)\Omega(D),A\big)$$

$$= \Omega(D)\ Y\big(N(S,D),A\big)$$

Using (16) and (18) and the general definition scheme (27) to define the empirical counterparts of radiance and irradiance the preceding equation yields:

$$E(S,D) = B(S,D)\Omega(D) \quad ,$$

which is the desired connection between *empirical luminance* and *empirical illuminance*.

From Table 2,

$$B(S,D) = 2\times 10^9 \text{ blondels or candelas/m}^2$$
$$\text{or lumens/m}^2 \times \text{sr}$$

and from Example 1 of Sec. 2.11:

$$\Omega(D) = 6.78\times 10^{-5} \text{ steradians}$$

Hence*:

$$E(S,D) = 2\times 10^9 \times (6.78\times 10^{-5})$$

$$= 136{,}000 \text{ lumens/m}^2$$

Example 2. If the sun in the context of Example 1 is at $\theta = 50°$ from the zenith, and surface S' is the projection on a horizontal plane of the surface S used in Example 1, what is the illuminance E(S',D) produced by the sun's rays on S'? To find this illuminance, recall (15) of Sec. 2.4:

$$H(S',D) = H(S,D)\cos\nu \quad ,$$

where the symbols are explained in detail in Sec. 2.4, and

---

*A relatively recent estimate (Ref. [128]) of E(S,D) is 136,000 lumens/m^2. See also [296] for a survey of measurements of the solar constant.

which now apply directly to the present context. In particular now $\alpha' = 50°$, and D is the cone of directions subtended by the sun's disk. Applying the operator $Y(\cdot,A)$, as defined in (13) to this equation, we have:

$$Y(H(S',D),A) = Y(H(S,D) \cos \alpha', A) \quad .$$

Using the definition scheme (27) now for empirical illuminance, this becomes:

$$E(S',D) = E(S,D) \cos \alpha' \quad .$$

Hence:

$$E(S',D) = 136{,}000 \times \cos 50°$$

$$= 136{,}000 \times .643$$

$$= 87{,}500 \text{ lumens/m}^2 \quad .$$

Example 3. Using the illuminance of Example 1, compute the illuminance produced by the sun's rays normally incident on surfaces in the orbits of the planets Venus and Mars. The requisite illuminances can be found by means of the inverse square law for irradiance deduced in Example 4 of Sec. 2.11. Thus, from (8) of Sec. 2.11, if r and s are two distances from the sphere's center, we deduce that:

$$H_r r^2 = H_s s^2 \quad .$$

Using the operator $Y(\cdot,A)$ and the definition scheme (27), this equation becomes:

$$E_r r^2 = E_s s^2 \quad .$$

Let E_r be the illuminance at the earth as given in the form E(S,D) in Example 1. Hence $r = 93 \times 10^6$ miles. In the case of Venus, $s = 67 \times 10^6$ miles. Hence:

$$E_s = E_r(r/s)^2$$

$$= 136{,}000 \times (93/67)^2$$

$$= 136{,}000 \times 1.93$$

$$= 262{,}000 \text{ lumens/m}^2$$

In the case of Mars, $s = 142 \times 10^6$ miles. Hence:

$$E_s = E_r(r/s)^2$$

$$= 136{,}000 \times (93/142)^2$$

$$= 136{,}000 \times .430$$

$$= 58{,}500 \text{ lumens/m}^2$$

Example 4. Compute the number of lumens F incident on a plane surface S of area A(S), every point of which is illuminated by a luminance distribution of constant magnitude B incident over directions within a conical solid angle D of half angle θ and whose axis is normal to S. The requisite relation for the lumens incident on S is obtained by beginning with (14) in Example 6 of Sec. 2.11:

$$H = \frac{N}{2} \int_0^{2\pi} \sin^2\theta(\phi)\, d\phi \quad ,$$

and applying the operator $Y(\cdot,A)$ to each side to get:

$$Y(H,A) = H\left(\frac{N}{2} \int_0^{2\pi} \sin^2\theta(\phi)\, d\phi, A\right)$$

$$= \frac{1}{2}\left(\int_0^{2\pi} \sin^2\theta(\phi)\, d\phi \right) Y(N,A) \quad ,$$

which, via the definition scheme (27), can be written:

$$E = \frac{B}{2} \int_0^{2\pi} \sin^2\theta(\phi)\, d\phi \quad .$$

In the present case, $\theta(\phi) = \theta$ for every ϕ, $0 \leq \phi \leq 2\pi$. Hence:

$$E = \pi B \sin^2\theta \quad .$$

Next, from (6) of Sec. 2.4 we have:

$$P(S,D,t,\nu) = H(S,D,t,\nu)\, A(S) \quad .$$

Applying the operator $Y(\cdot,A)$, to each side of this equation, we have:

$$Y(P(S,D,t,\cdot),A) = Y(H(S,D,t,\cdot)A(S),A) \quad .$$

From (15) and the definition scheme (27) applied to empirical irradiance, we consequently have:

$$F(S,D,t) = E(S,D,t)A(S) \quad .$$

Considering references to S,D, and t as understood for the present discussion, we distill this to:

$$F = EA \quad .$$

Thus we are led to the desired relation:

$$F = \pi BA \sin^2\theta \quad .$$

As a specific example, let $\theta = 30°$, $B = 120$ blondels, and $A = 4$ m^2. Then:

$$F = (3.14) \times (120) \times 4 \times (1/2)^2$$

$$= 378 \text{ lumens.}$$

Example 5. A red-orange appearing filter is known to have a band pass of 10 mμ, but it is not known precisely what wavelengths of radiant flux it transmits. An experiment is suggested and tried in which it is inferred that an irradiance of 2 watts/m^2 over the transmission interval $\Delta\lambda = 10$ mμ produces an illuminance of 1360 lumens. Can the transmission wavelengths of the filter also be inferred from this information? To answer this, consider the following observations.

From the definitional identity:

$$E = Y(H,\Lambda)$$

$$= K_m \int_0^\infty H(\lambda)\bar{y}(\lambda)\, d\lambda$$

and the fact that $H(\lambda) = 0$ outside the interval $\Delta\lambda$ about the unknown λ, we have very nearly:

$$E = K_m H(\lambda)\bar{y}(\lambda)\Delta\lambda \quad .$$

Hence:

$$\bar{y}(\lambda) = E/(K_m H(\lambda)\Delta\lambda)$$

$$= 1360/(680 \times 2 \times 10)$$

$$= 1/10 = .10 \quad .$$

From Table 1, by linear interpolation, we infer that $\lambda = 472$ mμ or 652 mμ. From the given general appearance of the filter's color, we infer that $\lambda = 652$ mμ.

2.13 Generalized Photometries

We conclude this chapter with a few observations on the necessary forms of certain generalized photometries which arise in an attempt to extend the salient ideas of classical photometry. The directions of extension to which we subject the ideas of photometry in this discussion are toward a more general class of 'luminosity' functions. The class we envision here is to contain not only the classical luminosity functions of human eyes, as briefly discussed in 2.11, but also irradiation-response functions describing photographic, phototransmissive, photovoltaic, photoemissive, and photocurrent phenomena. In short, we attempt to sketch in broad terms certain possible generalizations of the 'lumen' concept

with reference to irradiations which can be measurably effective on both organic and inorganic levels. Our discussion will consider in turn linear and nonlinear generalized photometries.

Linear Photometries

Let us begin with the simpler of the two generalizations: the *linear photometry*. The classical photometry discussed in Sec. 2.12 is an instance of a linear photometry. Using that discussion as a suitable motivation and background, we can initially and broadly define *theoretical linear photometry* to be the study of the properties of the effects $Z(\mathscr{R},M)$, on some physical object, of radiometric causes $\mathscr{R}$ over a wavelength set M, and under the premise that the numbers $Z(\mathscr{R},M)$ have certain postulated general properties. Specifically, for a given physical object (eye, skin, selenium cell, etc.), let $Z(\cdot,\cdot)$ be a function which assigns to each radiometric concept $\mathscr{R}$ and part M of the spectrum Λ a real number $Z(\mathscr{R},M)$ with the following properties*:

(i) $\mathscr{R}$-*Linearity*: For every two radiometric concepts $\mathscr{R}_1$ and $\mathscr{R}_2$ and nonnegative real numbers c_1 and c_2 for which $c_1\mathscr{R}_1 + c_2\mathscr{R}_2$ is defined, and for every part M of the spectrum Λ,

$$Z(c_1\mathscr{R}_1+c_2\mathscr{R}_2,M) = c_1Z(\mathscr{R}_1,M)+c_2Z(\mathscr{R}_2,M)$$

(ii) M-*Additivity*: For every radiometric function and every two disjoint parts M_1 and M_2 of Λ,

$$Z(\mathscr{R},M_1 \cup M_2) = Z(\mathscr{R},M_1)+Z(\mathscr{R},M_2)$$

(iii) M-*Continuity*: For every radiometric function

$$\text{if } 1(M) = 0, \quad \text{then } Z(\mathscr{R},M) = 0 .$$

An example of $Z(\mathscr{R},M)$ would be the amount of reddening (suitably measured) of human skin under irradiation (so that $\mathscr{R}$ can be irradiance H) over a certain portion of the far infrared (so that M consists of all wavelengths from, e.g., $\lambda = 800$ mμ to $\lambda = 850$ mμ). Another example of $Z(\mathscr{R},M)$ would be the rate of oxygen production by a leaf of some type of vegetation under irradiation (so that $\mathscr{R}$ can be scalar irradiance h) and over some part M of the spectrum. Marine biological contexts appear also to present potential areas for generalized photometries.

At any rate, the landmarks of an incipient linear photometry are properties (i), (ii), (iii) above. The concept of a linear photometry is certainly not empty since we have

---

*The footnote to the discussion of (3) and (4) of Sec. 2.3 applies also to the present discussion and should be consulted before proceeding.

classical photometry; and many additional photometric phenomena appear to be linear over given ranges within Λ. One of the useful facts about a linear photometry is the provable existence of a generalized luminosity function $\bar{z}(\cdot)$ within that photometry with the property that:

$$Z(\mathscr{R},M) = \int_M \mathscr{R}(\lambda)\bar{z}(\lambda)\, d\lambda \quad . \qquad (1)$$

We call this the *canonical representation* of $Z(\cdot,\cdot)$ for a linear photometry. The mathematical basis for this fact rests in general measure theory (the Radon-Nikodym theorem), and was alluded to earlier in (31) of Sec. 2.12 in connection with $\bar{y}(\cdot)$.

In sum then, it is possible to carry over to any general linear photometry the useful notion of a general 'luminosity function' which describes a general 'relative luminosity' of $\mathscr{R}$ over M (cf. (6) and (7) of Sec. 2.12). As a result it also appears possible to generate the concept of generalized 'lumens', so that one can initially place on a firm scientific footing such generalized linear photometries.

Nonlinear Photometries

Turning now to consider the prospects of forming a foundation for nonlinear photometries we are faced with the usual arresting fact about nonlinear phenomena: there are so many types of them. Were the world built so that there was only one type of nonlinearity--say of the power-exponential type or the sinusoidal type, etc.--then the problem of representing nonlinear phenomena would long ago have been thoroughly subdued, analytically speaking. However, since man's finite amount of attention must be spread over an apparently infinite class of nonlinear phenomena, this layer of attention must be nearly 'monomolecular' in depth wherever it exists.

To make a small start into the wilderness of nonlinear photometries, let us consider the first and logically the simplest types of departure from linearity. The preceding three statements (i)-(iii), constituting the defining properties of a linear photometry, may not all hold for given photometric phenomena. The three main types of departure from linearity would be:

Type I nonlinearity:	(i) does not hold; (ii) and (iii) hold
Type II nonlinearity:	(ii) does not hold; (i) and (iii) hold
Type III nonlinearity:	(i) and (ii) do not hold: (iii) holds

This choice of classification is based on the plausible feeling that: "if $1(M) = 0$, then $Z(\mathscr{R},M) = 0$" will always hold in any reasonable designed measure $Z(\cdot,\cdot)$ of a radiometric effect. Therefore, if a nonlinearity is encountered, it is likely to

be traceable to a violation of either (i) or (ii), or both. Each of the three types of nonlinearity will now be briefly discussed with the purpose in mind of suggesting possible routes toward linearization.

One very promising mode of approach to Type I nonlinearities is to find a function f which would linearize $Z(\cdot,M)$ for every M. Specifically, we suggest finding a real valued function f, defined on the real numbers, such that:

$$\text{(iv)}\quad f\{Z(c_1\mathscr{R}_1+c_2\mathscr{R}_2,M)\} = c_1 f\{Z(\mathscr{R}_1,M)\}+c_2 f\{Z(\mathscr{R}_2,M)\} \quad .$$

Many logarithmic and power nonlinearities are linearized away in this manner by the time-tested technique of plotting on logarithmic or exponential, or power coordinates. Whenever a linearizing function f can be found so that (iv) holds, then we say that the Type I nonlinearity is *removable*. The functional composition $f\circ Z$ of the linearizer f and the Z suffering a Type I removable nonlinearity, is now linear. Thus (i)-(iii) hold for $f\circ Z$ and so the canonical form (1) is available for use with $f\circ Z$. Summarizing: *whenever a Type I nonlinearity of a photometric measure* $Z(\cdot,M)$ *is removable by a linearizer* f *such that (iv) holds, then the composition* $f\circ Z(\cdot,M)$ *has a canonical representation (1).*

Let us consider now the Type II nonlinearity. We ask: if (ii) does not hold, *in what way* is it most likely not to hold? Imagine an erythemal phenomenon: a bit of living animal tissue is irradiated simultaneously by two distinct sets of radiation of non-overlapping wavelength sets M_1 and M_2. The effect $Z(\mathscr{R},M_1 \cup M_2)$ is noted. Then a biologically equivalent piece of tissue is irradiated in turn by samples of wavelength sets M_1 and M_2, and $Z(\mathscr{R},M_1)$ and $Z(\mathscr{R},M_2)$ are noted. Since M_1 and M_2 are allowed to be active separately, more effect-activity say, may take place in the tissue under each irradiation by M_i than when they act simultaneously. Thus, it may be that while the effects are not additive, they are M-*subadditive*:

$$\text{(v)}\qquad Z(\mathscr{R},M_1 \cup M_2) \leq Z(\mathscr{R},M_1) + Z(\mathscr{R},M_2) \quad .$$

Whenever a Type II linearity is encountered so that (ii) does not hold, it may be the case that M-subadditivity subsists. If subadditivity is indicated in a Type II nonlinearity, then it may be shown (cf. [103]) that for every $\mathscr{R}$ there exists an extended measure $Z^*(\mathscr{R},\cdot)$ which is additive in the sense of (ii). The net result we have reached may be stated as follows: *Every photometric measure* $Z(\mathscr{R},\cdot)$ *which exhibits nonlinearity of Type II and which is subadditive (i.e., (v) holds) may be extended to a linear photometric measure* $Z_*(\mathscr{R},\cdot)$ *for which a canonical representation (1) is possible.*

The immediate attempt at linearization of a Type III nonlinearity is to seek a linearizer f such that (iv) holds. Some Type III nonlinearities will surely succumb to these very general modes of attack. Beyond these few approaches lies an unknown field of potential modes of study of

nonlinear photometries.

2.14 Bibliographic Notes for Chapter 2

This chapter is based in the main on unpublished lecture notes (Refs. [210], [211]) in radiometry and photometry given in 1953 and 1954 at the Visibility Laboratory of the University of California, San Diego. The characterization of the foundations of radiometry in terms of a systematic use of additivity and continuity properties of the radiant flux function Φ, as given in Sec. 2.3, is derived from a similar treatment given in Ref. [251], and which in turn is based on the general measure-theoretic approach to radiometry and radiative transfer theory introduced in [216]. An important paper on photometry is that of Gershun, [98] who introduced and made precise the concept of the light field (our vector illuminance $\mathbf{E}$). Gershun also introduced the operational definition of radiance in the form $N = H/\Omega$ (re: (2) of Sec. 2.5). An important source of photometric wisdom may be found in the writings of Moon. In particular, the radiometric lectures cited above drew inspiration from some of the ideas of Refs. [184] and [185], especially in connection with developing general photometries. An old standard work on photometry and still valuable is Walsh's treatise [311]. The work by Le Grand, Ref. [153], is a relatively modern work on the optical-physiological properties of human vision which may be used to supplement the discussions of Sec. 2.12.

CHAPTER 3

THE INTERACTION PRINCIPLE

Gird up now thy loins like a man;
for I will demand of thee, and answer
thou me.
Where wast thou when I laid the
foundations of the earth? declare, if thou
hast understanding...
Whereupon are the foundations thereof
fastened? or who laid the cornerstone
thereof ...

JOB XXXVIII, 3-6.

3.0 Introduction

Radiative transfer theory is distinguished by the fact that it is one of the branches of theoretical physics that can be made to rest on a single principle from which all the salient structures of the theory may be systematically deduced. In this sense it is a closed subsystem of electromagnetic theory. The principle that permits this mode of construction of radiative transfer theory is called the *interaction principle*. The interaction principle is a distillate of many diverse conceptual constructions concerned with radiative transfer which have arisen during the past seven decades of evolution of the theory. In this chapter we shall state the principle and present various instances of it for a selected range of physical situations customarily encountered in practical applications of radiative transfer theory. It will be demonstrated that these physical situations can all be formulated within the theory in a uniform manner using a method which we call *the method of the interaction principle*. By means of examples we shall verify, on the one hand, that the salient theoretical structures of the theory do indeed fall under the domain of the principle, and, on the other hand, we shall prepare the groundwork for the various applications of the principle in the subsequent chapters of this work.

The principle of interaction in its essential form is a statement of the linearity of the classical radiative transfer processes. Thus radiative transfer theory, a complex webwork of deductions following from the principle, is at its core a linear theory of the interaction of light with matter on a phenomenological level. The linearity of the theory arises from the confluence of two main points of view adopted

by its principal developers and investigators since the turn of the century.

The first of these views is that the theory is concerned in the main either with radiant energy phenomena within the relatively hair-thin visible wavelength interval from 4×10^2 mμ to 7×10^2 mμ, or within a wider band of wavelengths from 10 to 10^5 mμ. Radiant energy phenomena within this four-orders of magnitude spread of the electromagnetic spectrum are, as we shall see below, associated with energies which barely tap the electronic energy levels of common atomic structures. The resultant interactions of radiant energy with matter are thereby limited essentially to elastic scatter activity, photoelectric effects, and simple absorption-emission phenomena. Inelastic scatter interactions of photons with matter are virtually ruled out within the $10\text{-}10^5$ mμ range of wavelengths. Within this domain the radiant energy interactions are manifestly linear, and thereby set one part of the stage for the linear structure of the interaction principle.

The second viewpoint adopted by the founders of the theory is that the interaction of light with matter is to be viewed on the phenomenological level, i.e., on the macroscopic level, with instruments which mimic normal human vision in its essential geometric characteristics. Therefore the delicate effects of wave phenomena, such as diffraction, interference, and other coherence activities are automatically excluded, by fiat, from the domain of classical radiative transfer theory. (See problems I-V of Sec. 141, Ref. [251]). In adopting this approach, we have 'shut our eyes completely' and have thought about all that we have seen. The linearities resulting from this predominantly geometrical viewpoint form the basis for the various additive and continuity properties of radiant flux discussed and developed at length in Chapter 2. These two views, one physical, the other geometrical, combine to act as effective linearization forces on the formulations of the concepts designed to describe radiative transfer processes in geophysical optics and great stretches of astrophysical optics.

The Physical Basis of the Linearity of the Interaction Principle

Before going on to state and illustrate the interaction principle, it will be instructive to examine in more detail the preceding physical assertions about the types of radiative processes limited to the purview of radiative transfer theory. In contemplating the consequences of the modern view that radiant energy is carried by quantized electromagnetic fields--i.e., by photons--we encounter a great number of possible types of interactions of photons with matter. Adopting a suggestion by Fano [90], we can usefully classify all of these variations into five main types of photon *interactions:*

I Interactions with atomic electrons.

II Interactions with atomic nucleons (protons, neutrons).

III Interactions with electric fields around charged atomic particles (electrons, charged nucleons).

IV Interactions with meson fields surrounding nucleons.

V Interactions with other photons.

The effects of these interactions are also greatly varied. But again for our present purposes, we need distinguish only three broad types of *effects:*

A. Outright absorption

B. Elastic scatter

C. Inelastic scatter

A word or two on the meaning of these terms is in order. Suppose we picture a photon as a small colored fuzzy ball, and an atom or a molecule of an optical medium as a relatively large complex spherical maze of thin, widely spaced fuzzy wires (electronic orbits or electron bonds) with tiny relatively dense central cores. Then in the case of effect A, the colored ball either zooms into the wire cage and becomes enmeshed in the maze of wires or is captured by a dense core, there to stay for a period of time far greater than that normally required to traverse the diameter of the cage at its initial speed. If it is ultimately released, we say an *emission process* has occured. In this captured state the ball, in effect, has been *absorbed* by the atom, and loses its identity as such, resulting momentarily in a higher orbit of one of the atom's electrons or in a higher stationary energy state of a molecule or in an increase in kinetic heat energy of the atom, or some combination of these. In the case of effect B, the colored ball caroms off (or skims through) the electronic shells of the atom, the net effect being a change of direction of travel of the photon with no change of its color, and we say that the photon is *scattered without change in wavelength.* In the final case, C, the ball becomes very briefly enmeshed in the electronic shell, or glances off the dense core, with greater or lesser wavelength than before, the net effect being a change of color and direction of travel, and we refer to the photon as *scattered with change in wavelength.*

Returning now to the interactions and their effects, we see that there are, in the present view, five possible types of interaction of a photon with matter and three possible types of effect. There are then in all fifteen possible interaction-effect pairs we can form: IA, IB, IC, IIA, IIB, IIC,..., VC. We shall call any of these fifteen interaction-effect pairs a *radiative process.* In Table 1 the fifteen general radiative processes are displayed by their characteristic interaction energies and by name whenever possible. For example, the class of processes we know as Rayleigh scatter is subsumed by the process IB. In this process a photon interacts with an atomic electron with the effect that it is scattered elastically. The inequalities that are indicated in the entries of the Table specify the interaction energies for

TABLE 1

GENERAL RADIATIVE PROCESSES

Photons Interacting with	Outright Absorption A	Elastic Scatter B	Inelastic Scatter C
Atomic Electrons I	Photoelectric Effect $\leq$ 0.1 Mev	Rayleigh Scatter $\leq$0.1 Mev	Compton Scatter $\geq$0.1 Mev
Atomic Nucleons II	Nuclear Photoelectric Effect $\geq$ 10 Mev	Nuclear Scatter $\geq$ 10 Mev	Nuclear Resonance Scatter $\geq$ 10 Mev
Electric field around Electrons, Nucleons III	Pair Production $\geq$1 Mev	Delbruck Scatter $\geq$ 3 Mev	Delbruck Resonance Scatter $\geq$ 3 Mev
Meson field around Nucleons IV	Meson Production $\geq$ 150 Mev	$\geq$ 150 Mev	$\geq$ 150 Mev
Other Photons V	Pair Production $\geq$ 1 Mev	$\geq$ 1 Mev	$\geq$ 1 Mev

which the associated process takes place. For example, "$\leq$0.1 Mev" means that the associated process takes place at 0.1 million electron volts or lower. Further, "$\geq$0.1 Mev" means that the associated process takes place at 0.1 million electron volts or higher. The unnamed processes and some of the other processes (IIIC, IVB, IVC, and the photonic interactions) have not been observed at this time of writing.

It will be instructive to correlate the Mev means of specifying the energy of a photon with its associated wavelength. By doing so, we shall be able to see clearly where the interaction energies common to radiative transfer theory stand in the arena of all this activity. To facilitate comparisons, we convert Mev units to wavelength units. The transition from Mev to wavelength is made by first recalling that the basic quantum of energy E associated with a photon of frequency ν is

$$E = h\nu$$

where the frequency is related to wavelength λ by:

$$\lambda\nu = v$$

and where "v" denotes the speed of light. If we let $v = c = 3 \times 10^{8}$ m/sec, and recall that $h = 6.625 \times 10^{-27}$ ergsec, then from the preceding relations:

$$\lambda = hc/E \quad \text{meters}$$

or

$$\lambda = \frac{1.24 \times 10^{-9}}{E} \quad \text{m}\mu \quad ,$$

where E is in Mev units. Thus if $E = 1$, then the associated energy is one million electron volts. The form in which we require this formula is:

$$\boxed{E = \frac{1.24 \times 10^{-3}}{\lambda} \quad \text{Mev}}$$

where λ is in mμ, i.e., millimicrons (10^{-9} meter), or as they are also called, *nanometers*. Assuming that our present interests lie mainly with processes in the wavelength range $10 \leq \lambda \leq 10^{5}$ mμ, we can now estimate the associated energies of interaction. Then by looking over the table of processes we can judge which of the areas of the main interaction arena are of primary interest. Thus we are interested in the energy range:

$$\frac{1.24 \times 10^{-3}}{10^{5}} \leq E \leq \frac{1.24 \times 10^{-3}}{10} \quad ,$$

i.e.,

$$1.24 \times 10^{-8} \leq E \leq 1.24 \times 10^{-4} \quad \text{Mev} \; .$$

In particular, green light (555 mμ) is on the order of 2×10^{-6} Mev.

What a tiny corner of the interaction arena we find ourselves in. A glance at the table shows that our world of radiant phenomena lies well within classes IA and IB. We shall call IA and IB the *classical radiative processes*. The classical radiative processes are, of course, replete with special radiative processes which include the various well-known absorption and scattering processes such as Raman, Rayleigh, Tyndall and resonant scatter; also fluorescence, and phosphorescence.

The simple calculation just performed shows that we need not be overly concerned in this work with such phenomena as Compton scatter--a relativistic phenomenon; pair production --a quantum electrodynamics phenomenon; or scattering of light by light--a quantum relativistic phenomenon. Even if we extend our interests down three orders of magnitude to

wavelengths of the order of 10^{-2} mμ, we still remain essentially within parts IA, IB, and IC of the interaction arena. The classical radiative process region is the domain of the classical Maxwell equations. We need not at present use any other models of the light field such as the Schrödinger or Dirac models, or those of general relativity to describe the activity in that part of the interaction arena in which we have found our current interests to lie.

It follows from the preceding analysis that the Maxwell equations in quantized, special relativistic form will suffice for most conceivable applications of radiative transfer theory in geophysical settings. Actually, it has been found that the classical (non quantified, non relativistic, linear) Maxwellian theory of electromagnetic fields may for all usual purposes encountered at present, serve as the nearest point on the mainland of physics to which radiative transfer theory may be adjoined when desired (see Chap. XIV of Ref. [251]). In this way is set the predominantly linear cast of the phenomenological theory, at the base of which may be found the interaction principle.

Plan of the Chapter

The plan of the remaining part of this chapter is as follows: we present in the next section a preliminary example of the interaction principle. This will serve to focus attention on a relatively concrete but yet typical instance of the use of the principle. From the example we shall extract the essence of the principle and state and discuss the result in Sec. 3.2. Beginning with Sec. 3.3, further examples of the interaction principle will be given. The examples of application will proceed in a systematic manner from relatively simple cases to progressively more complex cases until all the main tools of radiative transfer, as needed in the present work, have been formed.

Thus in Sections 3.3 to 3.5 we apply the interaction principle to the development of the reflectance and transmittance operators for plane and curved surfaces, with detailed examples presented to help fix the main ideas of the derivations and applications. In Sections 3.6 and 3.7 the reflectance and transmittance operators for plane-parallel media are developed and applications are given. The next step in the ascending scale of applications is taken in Sections 3.8 and 3.9 in which the interaction operators for general media are defined, functional relations governing the resulting operators are derived, and applications of the operators illustrated. Then the sequence of five sections 3.10-3.14 goes on to apply the preceding theory to the problem of constructing the basic inherent optical properties and radiance functions of radiative transfer theory (volume attenuation function, volume scattering function, path function, path radiance) and in Sec. 3.15 these are all assembled into the fundamental integral equation for radiance. At this point all the main tools of radiative transfer theory will have been constructed by means of the methodical use of the interaction principle. This use of the interaction principle is systematized and

summarized in Sections 3.16-3.18 in such a way as to aid the student of radiative transfer theory in attempting further applications and development of the method.

Throughout all the examples of this chapter--regardless of their level of complexity--runs a common thread of method: the *method of the interaction principle*. This method begins to form in Example 1 of Sec. 3.4; crystallizes in Example 2 of that section; and then recurs repeatedly, in the manner just outlined, through all the remaining illustrations of the chapter.

3.1 A Preliminary Example

We shall develop an example of the interaction principle in this section with the purpose in mind of fixing, on a relatively simple intuitive level, the salient features of the principle preparatory to stating the principle in its full form.

Empirical Reflectances and Transmittances for Surfaces

A prerequisite for the development of the example is the definition of the empirical reflectance of a small plane surface S. Figure 3.1 depicts such a surface S with unit outward normal **k**, which is irradiated at each point by radiant flux* through a narrow solid angle D', the flux passing through a hypothetical collecting surface S' on its way to S. The observed (empirical) field radiance of the incident flux is N(S',D') and the observed (empirical) surface radiance--arising from reflection of N(S',D') by S in a narrow solid angle D--is N(S',D';S,D). We write:

$$\text{"r(S',D';S,D)"} \quad \text{for} \quad \frac{N(S',D';S,D)}{N(S',D')\Omega(D')} \ ,$$

and call r(S',D';S,D) the (empirical) *reflectance* of surface S for the incident and reflected directions D' and D, respectively. Here S' is the projection of S on a plane perpendicular to a direction ξ', the central direction of D'. The function which assigns to (S',D') and (S,D) the number r(S',D';S,D) is called the (empirical) *reflectance function* for S. For the purpose of the present example, we assume r(S',D';S,D) is known for all pairs (D',D) of incident and response (reflected)

---

*For simplicity in exposition, throughout this work all radiant flux quantities will be assumed unpolarized, unless specifically stated otherwise. For an outline of the task of extending all results below to the polarized context, see Chapter XII of [251]. The interaction principle, however, holds implicitly for the polarized case. For the relative mathematical consistency of the assumption of the unpolarized light field with respect to the complete theory of the polarized field, see Sec. 13.11.

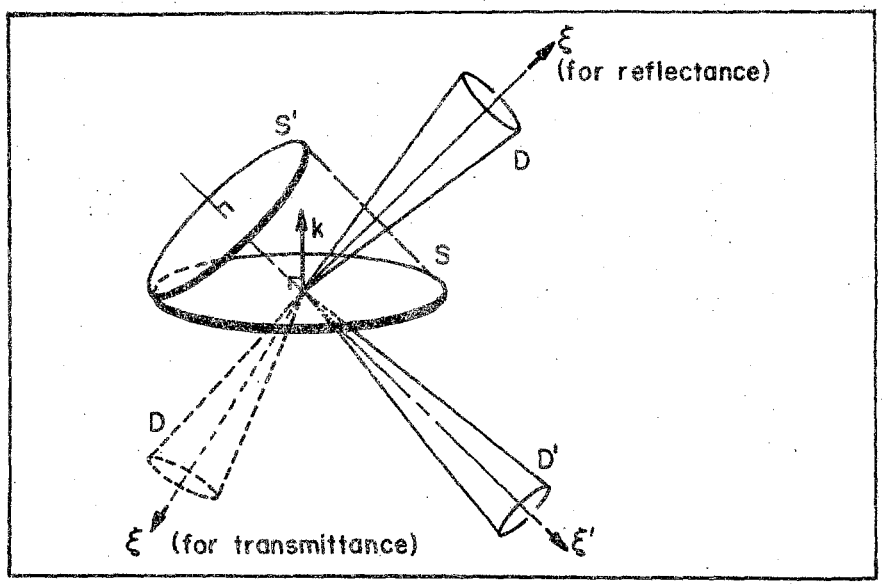


FIG. 3.1 Setting for empirical reflectances and transmittances of surfaces.

direction sets such that $\xi' \cdot \mathbf{k} < 0$ and $\xi \cdot \mathbf{k} > 0$, respectively. Here ξ' and ξ are arbitrary central directions of the sets D' and D, respectively. These two conditions merely require D' and D to lie on opposite sides of S, as in Fig. 3.1.

Now the essential property of the response radiance $N(S',D';S,D)$ is that it is additive with respect to D'. More precisely, experimental evidence indicates that we may assert the following property of $N(S',D';D,S)$. In each case let the sets D,D' of directions be circular, conical sets with central directions ξ,ξ', respectively. Then:

(i) *(D'-Additivity)* If S is a surface in an optical medium X and S is irradiated in turn by radiances $N(S_1',D_1')$ and $N(S_2',D_2')$, with $N(S_1',D_1';S,D)$ and $N(S_2',D_2';S,D)$ as the respective observed response radiances, then $N(S_1',D_1';S,D)+N(S_2',D_2';S,D)$ is the observed radiance of the S under simultaneous irradiation.

Furthermore:

(ii) *(D'-Continuity)* Let the geometric setting be defined as in (i). If $\Omega(D') = 0$ and $\xi' \neq \xi$, then $N(S',D';S,D) = 0$.

By letting D lie on the same side of S as D', these two empirically-based properties of reflected radiant flux are readily turned into the corresponding laws for

transmitted flux (see dotted direction cone in Fig. 3.1). Observe that, by virtue of (i), $\tau(S',D';S,D)$ is independent of the magnitude of $N(S',D')$. It should be particularly noted that (i) and (ii) are new laws which are independent of the D-additivity and D-continuity properties of Φ in Sec. 2.3. The present laws are intended to characterize a particular type of interaction of radiant flux with matter, whereas the earlier laws were intended to characterize certain intrinsic radiometric (principally geometric) properties of radiant flux regardless of its interaction with matter. These two properties permit a limiting process to culminate in rigorously defined reflectance and transmittance functions for surfaces. The details of such definitions will be considered in (6)-(9) of Sec. 3.3. For the present we use (i) and (ii) as they stand to help solve the following radiometric interaction problem.

The Problem

Two plane surfaces, S_1 and S_2, in a vacuum are mutual point sources. In addition, they are mutually visible and are irradiated by sources of radiance N_1^o and N_2^o over solid angles, D_{o1} and D_{o2}, respectively, as shown in Fig. 3.2. These incident radiances initiate an interreflection process between S_1 and S_2 with resultant surface radiances $N(S_1,D_{12})$ and $N(S_2,D_{21})$. Here D_{12} is the set of directions from a point in S_1 to every point in S_2. Since S_1 and S_2 are mutual point sources (i.e., each is a point source as seen from the points of the other), D_{12} does not vary appreciably as location is varied over S_1, and so may be assumed constant over S_1.

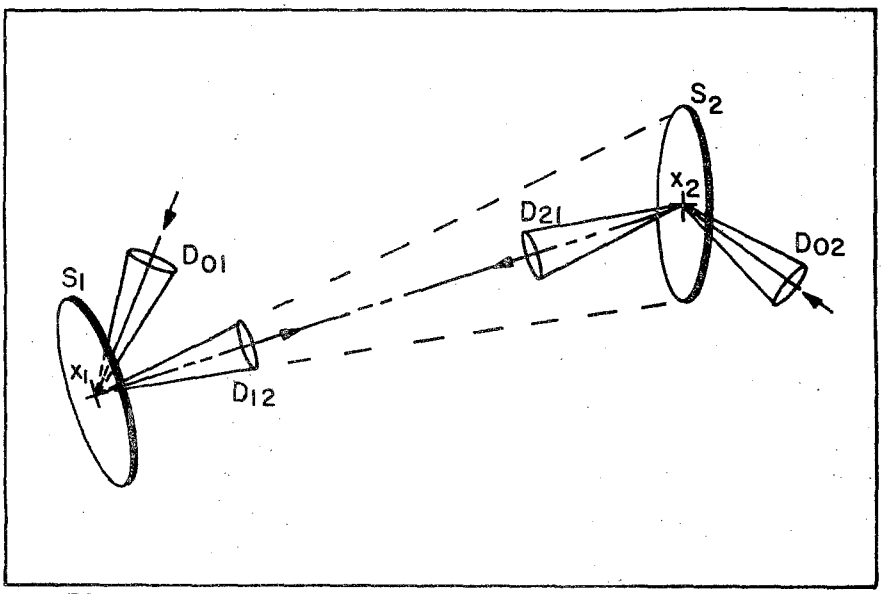


FIG. 3.2 Setting up an interaction calculation for surfaces S_1 and S_2.

Similarly D_{21} is the set of directions from a fixed point of S_2 to every point of S_1; and D_{21} has the same general property as D_{12}. With these preliminaries out of the way, we can now state the present problem: *Given:* S_1 *and* S_2, *as above, with known empirical reflectance functions,* r_1 *and* r_2, *and initial irradiations,* N_1^0, N_2^0. *Required: the steady state empirical radiances* $N(S_1, D_{12})$ *and* $N(S_2, D_{21})$.

The Present Instance of the Interaction Principle

To facilitate the present discussion let us write:

"N_{12}" for $N(S_1, D_{12})$

"N_{21}" for $N(S_2, D_{21})$

"r_{212}" for $r_1(S_{12}', -D_{12}; S_1, D_{12})$

"r_{121}" for $r_2(S_{21}', -D_{21}; S_2, D_{21})$

"r_{012}" for $r_1(S_{01}', D_{01}; S_1, D_{12})$

"r_{021}" for $r_2(S_{02}', D_{02}; S_2, D_{21})$

where S_{12}', e.g., is the projection of S_1 on the plane normal to the direction from x_1 to x_2. Similarly, for S_{21}', S_{01}', and S_{02}'. In the case of S_{01}', e.g., imagine an external source point x_0. The set $-D_{12}$ consists of all negatives of directions in D_{12}. Thus if ξ is in D_{12}, then $-D_{12}$ contains $-\xi$. Now by virtue of the definition of empirical reflectance, the D'-additivity property (i) above, and the fact that the intervening space between S_1 and S_2 is a vacuum (so that the radiance invariance law can be used) we have:

$$N_{12} = N_1^0 r_{012}\Omega_{01} + N_{21} r_{212}\Omega_{12} \tag{2}$$

$$N_{21} = N_2^0 r_{021}\Omega_{02} + N_{12} r_{121}\Omega_{21} \tag{3}$$

where we have written:

"Ω_{0i}" for $\Omega(D_{0i})$, $i = 1,2$

and

"Ω_{ij}" for $\Omega(D_{ij})$, $i,j = 1,2$

For later purposes it is convenient to make one final set of definitions. We write:

"$\sum_{ij}^0$" for $r_{0ij}\Omega_{0i}$ $i,j = 1,2$

"$\sum_{iji}$" for $r_{iji}\Omega_{ji}$ $i,j = 1,2$.

Then (2) and (3) become:

$$N_{12} = N_1^o\Sigma_{12}^o + N_{21}\Sigma_{212} \tag{4}$$

$$N_{21} = N_2^o\Sigma_{21}^o + N_{12}\Sigma_{121} \tag{5}$$

Equations (4) and (5) together constitute the algebraic core of the statement of the present form of the interaction principle. In the present case we have two relatively small plane surfaces which are interacting radiometrically. Each surface S_i ($i=1$ or 2) is irradiated by two incident parcels of radiant flux in the form of the empirical radiances, N_i^o and N_{ji}, and S_i itself has a resultant surface radiance N_{ij}. *To the sets of such incident radiances,* N_i^o *and* N_{ji} *and response radiances* N_{ij} *associated with* S_i *($i=1$ or 2), there correspond four interaction operators (numbers in this case), namely* Σ_{ij}^o *and* Σ_{jij}, *such that (4) and (5) hold.* The main role of the interaction principle in the present case would be to assert the existence of these operators and to yield the interaction equations (4), (5).

Solution of the Problem

The interaction principle formulation (4), (5) of the present problem leads to the solution of the problem by means of the theory of simultaneous algebraic equations. Thus, multiplying each side of (4) by Σ_{121}:

$$N_{12}\Sigma_{121} = N_1^o\Sigma_{12}^o\Sigma_{121} + N_{21}\Sigma_{212}\Sigma_{121}$$

and using this representation of $N_{12}\Sigma_{121}$ in (5):

$$N_{21} = N_2^o\Sigma_{21}^o + \left(N_1^o\Sigma_{12}^o\Sigma_{121} + N_{21}\Sigma_{212}\Sigma_{121}\right)$$

whence:

$$N_{21} = \frac{N_2^o\Sigma_{21}^o + N_1^o\Sigma_{12}^o\Sigma_{121}}{1 - \Sigma_{212}\Sigma_{121}} .$$

The radiance N_{12} can be found by permuting the symbols "1" and "2" in this equation. The complete solution of the system (4), (5) is then:

$$N_{12} = \frac{N_1^o\Sigma_{12}^o + N_2^o\Sigma_{21}^o\Sigma_{212}}{1 - \Sigma_{121}\Sigma_{212}} \tag{6}$$

$$N_{21} = \frac{N_2^o\Sigma_{21}^o + N_1^o\Sigma_{12}^o\Sigma_{121}}{1 - \Sigma_{121}\Sigma_{212}} \tag{7}$$

Discussion of Solution

A sufficient condition that N_{12} and N_{21} are determinable via (6) and (7) is that the product $\Sigma_{121}\Sigma_{212}$ is less than 1. We shall now show that a sufficient condition that this latter property holds is that at least one of Σ_{121} and Σ_{212} is strictly less than 1. An examination of these definitions of Σ_{121} and Σ_{212} shows that this condition may be based on a particular form of the principle of conservation of energy. To see this in the case of Σ_{212}, we need only systematically unfold the definitions leading to it. Thus, Σ_{212} is:

$$r_{212}\Omega_{12}$$

The quantity r_{212} is:

$$r_1(S'_{12},-D_{12};S_1,D_{12}) \quad .$$

This in turn is the value of the empirical reflectance function for S_1. By (1), and the fact that this value of r_1 is independent of the magnitude of the irradiating flux, we can select any incident radiance, say N^-_{12} over $-D_{12}$, and let N^+_{12} be the response (reflected) radiance over D_{12}. Then:

$$r_1(S'_{12},-D_{12};S_1,D_{12})\Omega_{12} = \frac{N^+_{12}}{N^-_{12}} \quad .$$

If $A(S'_{12})$ is the projected area of S_1 on a plane normal to the direction from x_1 to x_2 (see Fig. 3.2), and $P(S_1,-D_{12})$ and $P(S_1, D_{12})$ are the radiant fluxes associated with $N^{\pm}_{12}$, then the incident radiant flux is given by:

$$P(S_1,-D_{12}) = N^-_{12}A(S'_{12})\Omega_{12}$$

and the surface (response) radiant flux is given by:

$$P(S_1,D_{12}) = N^+_{12}A(S'_{12})\Omega_{12} \quad .$$

Here we have used the fact that $\Omega(-D_{12}) = \Omega(D_{12}) = \Omega_{12}$, and also the operational definition of surface radiance (Sec. 2.6). Hence:

$$r_1(S'_{12},-D_{12};S_1,D_{12})\Omega_{12} = \frac{P(S_1,D_{12})}{P(S_1,-D_{12})}$$

At this point we choose the *energy conservation principle* in the form which states that: *if* P^- *is the total radiant flux incident on a given surface* S *and* P^+ *is the total radiant flux leaving the surface* S *and* P^+ *and* P^- *are independent of time, then* $P^+ \leq P^-$. We shall assume this statement is true. From this we deduce in particular that:

$$r_1(S'_{12},-D_{12};S_1,D_{12})\Omega_{12} \leq 1 \quad ,$$

so that:

$$\Sigma_{212} \leq 1 \quad .$$

A similar inequality now follows for Σ_{121}. These inequalities are the most we can say, without further qualifications, about any reflectance (or transmittance) operator occurring in the theory of radiative transfer. *Thus in a particular geometrical situation we must explicitly postulate or demonstrate that at least one of Σ_{121} and Σ_{212} in (6) and (7) is strictly less than 1;* and as our analysis has now made clear, this is a sufficient condition that (6) and (7) uniquely determine N_{12} and N_{21}.

Related Problems and their Solutions

The solutions (6) and (7) of the problem considered above can be used to solve related problems centering on the radiometric interaction of S_1 and S_2. Suppose, for example, we require the surface radiance of S_1 in some set $D_{1\beta}$ of directions other than D_{12}. Here "β" is associated with point $x\beta$ which may be any point in the surrounding medium either in or not in S_1 or S_2. Toward this end we write:

$$\text{"}N_{1\beta}\text{"} \quad \text{for} \quad N(S_1, D_{1\beta})$$

$$\text{"}r_{21\beta}\text{"} \quad \text{for} \quad r_1(S'_{12}, -D_{12}; S_1, D_{1\beta})$$

$$\text{"}r_{01\beta}\text{"} \quad \text{for} \quad r_1(S'_{01}; D_{01}; S_1, D_{1\beta}) \quad .$$

Then by the D'-additive property (i) above and the radiance invariance law we have:

$$N_{1\beta} = N_1^o r_{01\beta}\Omega_{01} + N_{21} r_{21\beta}\Omega_{12} \quad .$$

In an exactly similar manner we arrive at the surface radiance of S_2:

$$N_{2\beta} = N_2^o r_{02\beta}\Omega_{01} + N_{12} r_{12\beta}\Omega_{21} \quad .$$

Once again we can contract these solutions into a fixed form which clearly reveals the underlying unity of the interaction concept. Thus by writing:

$$\text{"}\Sigma^o_{i\beta}\text{"} \quad \text{for} \quad r_{oi\beta}\Omega_{oi} \quad , \; i = 1,2$$

and

$$\text{"}\Sigma_{ij\beta}\text{"} \quad \text{for} \quad r_{ij\beta}\Omega_{ji} \quad , \; i,j = 1,2$$

the preceding equations become:

$$N_{1\beta} = N_1^o \Sigma_{1\beta}^o + N_{21}\Sigma_{21\beta} \tag{8}$$

$$N_{2\beta} = N_2^o \Sigma_{2\beta}^o + N_{12}\Sigma_{12\beta} \tag{9}$$

where N_{12} and N_{21} are as given in (6) and (7). For the purposes of later comparison with the general statement of the interaction principle we observe that: *to the incident radiances* N_i^o *and* N_{ji} *on* S_i *(i = 1 or 2) and response radiance* $N_{i\beta}$ *there correspond four interaction operators (numbers in this case), namely* $\Sigma_{i\beta}^o$ *and* $\Sigma_{ij\beta}$ *such that (8) or (9) hold.* The index β in (8) and (9) (and in the equations below) may be replaced by distinct indices, if desired. In other words, surfaces S_1 and S_2 may give off radiances in distinct directions, which may be computed by (8), (9) by replacing index β in (9), say, by a new index γ.

An Alternate Form of the Principle

We now abruptly change our conceptual orientation in Fig. 3.2 from that of two radiometrically interacting surfaces S_1 and S_2 to that of a single subset S of the optical medium irradiated from without by radiant flux. This change in orientation can be encouraged by imagining S_1 and S_2 in Fig. 3.2 to be encircled by a closed dashed curve and to think of the curve as holding a *single subset* S of space (that is, S is a disconnected subset which happens to consist of two separate surfaces, S_1 and S_2). This subset S is irradiated at two places by incident radiances N_1^o and N_2^o, and the response of S is imagined in the form of two streams of flux characterized by $N_{1\beta}$ and $N_{2\beta}$. This conceptual compression of S_1 and S_2 into a single radiometrically responsive entity can be expressed symbolically as follows. We first write the system (8) and (9) in matrix form (replacing "β" in (9) by "γ", for generality):

$$(N_{1\beta}, N_{2\gamma}) = (N_1^o, N_2^o)\begin{pmatrix} \Sigma_{1\beta} & 0 \\ 0 & \Sigma_{2\gamma} \end{pmatrix} + (N_{21}, N_{12})\begin{pmatrix} \Sigma_{21\beta} & 0 \\ 0 & \Sigma_{12\gamma} \end{pmatrix}.$$

Further, from (6) and (7) we can write:

$$(N_{21}, N_{12}) = (N_1^o, N_2^o)\begin{pmatrix} \Psi_{11}^o & \Psi_{12}^o \\ \Psi_{21}^o & \Psi_{22}^o \end{pmatrix},$$

where, in turn, we have written:

$$\text{"}\Psi^{o}_{12}\text{"} \quad \text{for} \quad \frac{\Sigma^{o}_{12}}{(1-\Sigma_{121}\Sigma_{212})}$$

$$\text{"}\Psi^{o}_{21}\text{"} \quad \text{for} \quad \frac{\Sigma^{o}_{21}}{(1-\Sigma_{121}\Sigma_{212})}$$

$$\text{"}\Psi^{o}_{11}\text{"} \quad \text{for} \quad \frac{\Sigma^{o}_{12}\Sigma_{121}}{(1-\Sigma_{121}\Sigma_{212})}$$

$$\text{"}\Psi^{o}_{22}\text{"} \quad \text{for} \quad \frac{\Sigma^{o}_{21}\Sigma_{212}}{(1-\Sigma_{121}\Sigma_{212})} \quad .$$

Then going one step further and writing:

$$\text{"}\Psi^{o}\text{"} \quad \text{for} \quad \begin{pmatrix} \Psi^{o}_{11} & \Psi^{o}_{12} \\ \Psi^{o}_{21} & \Psi^{o}_{22} \end{pmatrix}$$

and:

$$\text{"}\Psi^{o}_{\beta\gamma}\text{"} \quad \text{for} \quad \begin{pmatrix} \Sigma^{o}_{1\beta} & 0 \\ 0 & \Sigma^{o}_{2\gamma} \end{pmatrix}$$

and:

$$\text{"}\Psi_{\beta\gamma}\text{"} \quad \text{for} \quad \begin{pmatrix} \Sigma_{21\beta} & 0 \\ 0 & \Sigma_{12\gamma} \end{pmatrix} \quad ,$$

we arrive at the following alternate representation of the system (8) and (9):

$$(N_{1\beta},N_{2\gamma}) = (N^{o}_{1},N^{o}_{2})(\Psi^{o}_{\beta\gamma} + \Psi^{o}\Psi_{\beta\gamma}) \quad .$$

Let us write:

$$\text{"}\Phi^{o}_{\beta\gamma}\text{"} \quad \text{for} \quad \Psi^{o}_{\beta\gamma} + \Psi^{o}\Psi_{\beta\gamma}$$

and thereby arrive at the desired form of the system (8), (9):

$$\boxed{(N_{1\beta},N_{2\gamma}) = (N^{o}_{1},N^{o}_{2})\Phi^{o}_{\beta\gamma}} \quad . \qquad (10)$$

The significance of (10) may be discerned as follows: for the given subset S we have shown that *to an arbitrary pair of incident radiances* (N^{o}_{1},N^{o}_{2}) *and response radiances* $(N_{1\beta},N_{2\gamma})$ *there corresponds a unique interaction operator (a 2×2 matrix*

of real numbers in this case), namely $\Phi^{o}_{\beta\gamma}$, *such that (10) holds.*

Equation (10) constitutes an alternate form of the interaction principle to that displayed in (4), (5) or in (8), (9). This alternate form is designed to give an indication of the potential internal complexity of an object S to which the interaction principle may assign an interaction operator. It is not too great an extension of ideas from the setting of (10) to the setting of an arbitrary finite number of interacting surfaces. However, the systematic study of such systems of interacting surfaces or solids is the domain of discrete space radiative transfer theory and lies far beyond our present concerns. For those interested in pursuing this matter further, we observe that the complete theory of such systems is developed in Ref. [251].

The Natural Mode of Solution

We conclude this preliminary example of the interaction principle by displaying an alternate mode of solution of the problem of the radiometric interaction of the two surfaces S_1 and S_2 considered above. Our purpose is to show that this alternate mode of solution and the interaction principle mode of solution are equivalent. As our developments proceed into the next chapter, we shall also see that each mode of solution possesses a valuable conceptual kernel which is capable of extension to quite wide domains of application in radiative transfer theory in general, and hydrologic optics in particular. This alternate mode of solution we call the *natural mode* of solution, for it appears to be conceptually the simplest and most natural approach to interreflection problems.

The natural mode of solution may be described quite briefly as follows. We imagine a hyper-fast camera filming the radiometric interaction of two surfaces, S_1 and S_2. The filmed episode begins the instant the incident radiances N_1^o and N_2^o simultaneously impinge on S_1 and S_2, respectively. In a playback of the filmed episode in slow motion, we see part of N_1^o reflected from S_1 and start to travel toward S_2. This reflected flux eventually reaches S_2 and part of it is redirected back toward S_1. In the meantime N_2^o has been reflected at S_2 and part of the reflected flux moves on to S_1, there to be reflected and to have some flux begin to return to S_2. As the film continues, the sources N_1^o and N_2^o continue to steadily pour flux on S_1 and S_2. After a while S_1 is being irradiated by photons, some of which come directly from N_1^o, some of which are making their first arrival from S_2, and some their second arrival from S_2, etc. By and by the fluxing and interfluxing reaches a measurable steady state (while, in principle, however, there will always be some interreflection number which has not yet been attained). The following argument develops the symbolic representation of this steady state interreflection process.

Retaining the notation of the preceding discussions, let us go on to write:

$$\text{"}N_{12}^{1}\text{"} \quad \text{for} \quad N_1^0 \Sigma_{12}^0$$

and

$$\text{"}N_{21}^{1}\text{"} \quad \text{for} \quad N_2^0 \Sigma_{21}^0 \quad .$$

Further, for every $j = 2,3,\ldots$, we write:

$$\text{"}N_{12}^{j}\text{"} \quad \text{for} \quad N_{21}^{j-1} \Sigma_{212}$$

$$\text{"}N_{21}^{j}\text{"} \quad \text{for} \quad N_{12}^{j-1} \Sigma_{121} \quad .$$

By recalling the moving-picture allusion it is easy to see that N_{12}^{j} is interpretable as the surface radiance of S_1 in the directions of S_2 consisting of radiant flux having undergone precisely j reflections. Again, by means of the analogy, we are led to write:

$$\text{"}N_{12}\text{"} \quad \text{for} \quad \sum_{j=1}^{\infty} N_{12}^{j} \tag{11}$$

$$\text{"}N_{21}\text{"} \quad \text{for} \quad \sum_{j=1}^{\infty} N_{21}^{j} \tag{12}$$

The numbers N_{12} and N_{21} obtained in this way are called the *natural solution* of the present problem of the radiometrically interacting surfaces S_1 and S_2. That N_{12} and N_{21} are indeed solutions of the steady-state interreflection problem associated with S_1 and S_2 will now be shown. By starting with the definitional identity arising from (11):

$$N_{12} = \sum_{j=1}^{\infty} N_{12}^{j} \quad ,$$

we deduce the following chain of equalities:

$$
\begin{aligned}
N_{12} &= \sum_{j=1}^{\infty} N_{12}^{j} \\
&= N_{12}^{1} + \sum_{j=2}^{\infty} N_{12}^{j} \\
&= N_{12}^{1} + \sum_{j=2}^{\infty} \left(N_{21}^{j-1} \Sigma_{212} \right) \\
&= N_{12}^{1} + \left(\sum_{j=1}^{\infty} N_{21}^{j} \right) \Sigma_{212} \\
&= N_{1}^{0} \Sigma_{12}^{0} + N_{21} \Sigma_{212} \qquad (13)
\end{aligned}
$$

The last equality follows from (12) and the preceding definition of N_{12}^{1}. By comparing (13) and (4) we see that the natural mode of solution implies the interaction mode of solution. Evidently the steps in (13) are reversible, so that the interaction mode of solution implies the natural mode of solution. Thus the two modes of solution are equivalent in this case. Since the interaction mode of solution clearly represents the solution of the interreflection problem of S_1 and S_2, the natural mode of solution therefore is also, by virtue of the preceding equivalence, a solution of the interreflection problem. This equivalence actually holds in very general settings and has been established in detail for these settings, in Ref. [251]. We shall have occasion to study and use once again this equivalence of the two techniques later in the present work. Finally, we observe that the sums in (11) and (12), being reducible to a simple geometric series with ratio $\Sigma_{121}\Sigma_{212}$ and initial term of the form $(N_{1}^{0}\Sigma_{ij}^{0} + N_{j}^{0}\Sigma_{ji}^{0}\Sigma_{jij})$ (i=1, j=2 for (11); i=2, j=1 for (12)), are readily evaluated; these sums are given by (6) and (7).

3.2 The Interaction Principle

With the preliminary example complete, we turn now to the statement of the central principle of radiative transfer theory:

The Interaction Principle: For every X, S, A, B, m *and* n, *if* X *is an optical medium and* S *is a subset of* X, *and* A $(= (A_1, \ldots, A_m))$ *is a class of sets* A_i *consisting of incident radiometric functions on* S, *and* B $(= (B_1, \ldots, B_n))$ *is a class of sets* B_j *consisting of response radiometric functions on* S, *and* m *and* n *are positive integers, then there exists a unique set* $\{s_{ij}: i=1,\ldots,m,\ j=1,\ldots,n\}$ *of linear (interaction)*

operators s_{ij} *with domain* A_i *and range* B_j *with the property that for every element* $(a_1,\ldots,a_m)$ *of* A *there exists an element* $(b_1,\ldots,b_n)$ *of* B *such that:*

$$b_j = \sum_{i=1}^{m} a_i s_{ij}$$

or in matrix form:

$$b = as$$

where we have written:

"a" for $(a_1,\ldots,a_m)$

"b" for $(b_1,\ldots,b_n)$

"s" for $\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ s_{m1} & s_{m2} & \cdots & s_{mn} \end{pmatrix}$

Discussion of the Interaction Principle

We shall discuss in some detail the meanings of the various terms in the interaction principle. First of all, the meaning of the term "optical medium" as used in the statement is quite broad and, for example, is intended to have as real designata such parts of the world as lakes, oceans and various portions of the atmosphere. From the mathematical point of view, "optical medium" may be interpreted simply as part of Euclidean three-dimensional space such as the region between two infinite parallel planes or the interior of a sphere, etc., in which we assume that the principles of geometric optics hold, in particular, Fermat's principle. There will eventually evolve, as the studies progress and the basic constructs assume their final form, a relatively technical version of what we mean by the term "optical medium" in the fully developed theory (re: Def. 5 of Sec. 9.1). However, for the present the term may have either of the simple meanings suggested above.

The meanings of the terms A, B, and s_{ij} in the principle can be illustrated using the preliminary example of Sec. 3.1. Let us return to the setting summarized by Eqs. (4) and (5) of Sec. 3.1. In that setting the optical medium was some (physically) vacuous region X of Euclidean three-space containing two plane surfaces S_1 and S_2. We concentrate attention on S_1. Then S_1 is an instance of S in the principle. Consider the set of all incident radiances like N_1^o on S_1.

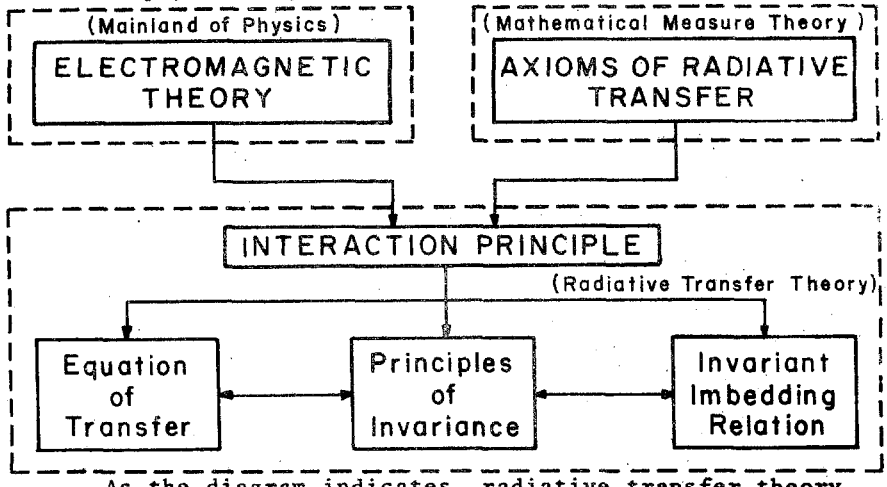
This set of incident radiances becomes the set A_1 in the principle. Consider the set of all incident radiances like N_{21} on S_1. This becomes the set A_2 in the principle. Together, (A_1,A_2) constitute the incident class A in the principle, so that $m = 2$. It should be noted that A_1 and A_2 are each closed under the operations of forming sums and products by nonnegative numbers *(linear closure)*. Thus if N_1 and N_2 are in A_1, then so is $cN_1 + dN_2$ where c and d are nonnegative numbers. This feature of A_1 and A_2 comes automatically with the requisite linearity of the s_{ij}. The class B of response functions S_1 consists of one set B_1, with N_{12} as a typical element. Therefore in the case of S_1 we have $m = 2$, and $n = 1$, with Σ^o_{21} and Σ_{212} as the present instances of s_{11} and s_{21}, respectively. Hence one invocation of the interaction principle for the case of S_1 yields (4). Another and distinct invocation in the case of S_2 yields (5).

The alternate example summarized in (10) of Sec. 3.1 provides a further illustration of the principle's linear algebraic statement. In (10) of Sec. 3.1, X is the same space as above. Now, however, S_1 and S_2 are considered parts of one and the same subset, say S of X. Consider the set of all ordered pairs of incident radiance on S like (N^o_1,N^o_2). This becomes A_1 in the principle. Consider the set of all ordered pairs of response radiances of S like $(N_1\beta,N_2\gamma)$. This becomes B_1 in the principle. Therefore in the present case of S, we have $m = n_o = 1$, and s_{11} is $\Phi^o_{\beta\gamma}$. As we select any new incident pair (N^o_1,N^o_2), there corresponds the associated response pair $(N_1\beta,N_2\gamma)$ given by (10). Clearly (10) is the present instance of the matricial form of the principle's algebraic statement.

As we progress along the line of examples of the interaction principle we shall be gradually less explicit in pointing out the particular parts of the current form of the interaction principle, leaving the details of correlation more to the reader as he becomes familiar with the principle. In all the subsequent uses of the principle, we shall look upon it as a convenient working principle, i.e., a rule of action for the formulation of subordinate principles, the various laws, and everyday problems of radiative transfer theory. The practical uses of the principle are directed to determining the light field in natural optical media by finding the interaction operator s_{ij}, supplied by the basic principle, for a given medium. The determination of the structure of the operators s_{ij} and the various functional equations they satisfy constitutes one of the more interesting and challenging problems of modern radiative transfer theory. We shall begin the investigation of these operators in the present chapter and continue it in Chapter 7.

The Place of the Interaction Principle in Radiative Transfer Theory

It is not intended that the interaction principle categorically replace all classical instances of itself such as the principles of invariance and the invariant imbedding relation, or other classical instances that occur in the literature or that arise during the subsequent developments below. Rather, it is intended that the principle be viewed by its users simply as a working principle of radiative transfer theory, and to be used (and perhaps refined) by those students of the subject who prefer to envision the theory as governed by and derivable from a single idea. The place of the interaction principle in radiative transfer theory and in the mainstream of physics may be summarized by the following diagram:



As the diagram indicates, radiative transfer theory may join the mainland of physics via electromagnetic theory (see, e.g., Chapter XIV, Ref. [251]) or the theory may be made completely autonomous using an axiomatic formulation made elsewhere (Chapter XV, Ref. [251]). Direct interconnections also exist between the three principal parts of the theory (indicated in the diagram below the interaction principle). In fact the internal ties on the level of the general equation of transfer, the general principles of invariance, and the general invariant imbedding relation are so strong that these ties are effectively logical equivalences. The details of the pursuit of these connections are mainly mathematical and are beyond the scope of the present work. For further details on this matter, the reader is referred to the various chapters of Ref. [251].

Levels of Interpretation of the Interaction Principle

The great practical range and depth of the interaction principle arises from the levels of interpretation on which it may be applied. There are generally four main levels of

interpretation of the principle: the *point, line, surface,* and *space* levels. Of these, the surface and space levels of interpretation are operationally the most meaningful. The point and line interpretations are special theoretical artifices which increase the range of the principle in specific settings. The preliminary example above is an instance of the surface level of interpretation. In general, the *surface level* interpretation of the interaction principle subsists when one interprets the subset S of a space X as a subset of one less dimension than X. For three-dimensional spaces X, S would have two dimensions. For two-dimensional spaces X (which arise in certain mathematical models) S would have one dimension, etc. In general the *space-level interpretation* of the interaction principle subsists when one interprets the subset S of a space X as a subset of the same dimension as X. Plane-parallel slabs, spherical solids in Euclidean three space are settings for the space-level interpretation. For two-dimensional spaces X, the subset S would have two dimensions, etc.

Of the remaining two levels of interpretation of the principle, the point level interpretation is the more widely used. In fact the point-level interpretation covers so much ground that it is convenient to regard it from two separate aspects. The *general point-level interpretation* of the interaction principle subsists when X is a general space whose points are arbitrary. The general point-level interpretation is of most use in the development of general discrete-space theory (Ref. [251]). The *special point-level interpretation* of the interaction principle subsists when S is a point or an optically small three-dimensional subset of space (i.e., e.g., a point source) in which single scattering processes are to be dominant relative to multiple scattering processes. This special interpretation is commonly used to establish in an intuitive fashion the concept of the volume scattering function, which plays a key role in the theory (see Sec. 13.4). An alternate establishment of the volume scattering function could take place strictly and rigorously in the space-level interpretation (see Sec. 3.14). The special point-level interpretation is also a useful and defensible ploy in setting up radiative transfer theory and is thereby retained and given a special status. (See, e.g., Example 1, Sec. 3.17.)

The final level of interpretation to be discussed is the line-level interpretation of the interaction principle. The *line-level interpretation* subsists when one interprets the subset S of a space X as a one-dimensional subset of X. The line-level interpretation is not operationally meaningful as are the surface, space and special point-level interpretations. However, it is retained because it favors useful mathematical artifacts, as does the special point-level interpretation. Furthermore, like the special point-level interpretation, the use of the line-level interpretation is rigorously defensible by means of limit aguments starting with the space-level interpretation; for that reason it is retained as a useful technical device. We shall use it below in viewing the path radiance as the response of a path in real optical medium to the incident path function radiances along the path. (Example 2, Sec. 3.17.)

Unless specifically noted otherwise, we shall henceforth mean by "optical medium" any three-dimensional part X of Euclidean three-dimensional space. This then will automatically set the dimensionality of S in the various interpretations of the interaction principle. (A formal definition of optical media, as they are studied in radiative transfer theory, is given in Sec. 9.1.)

3.3 Reflectance and Transmittance Operators for Surfaces

In this section we begin the sequence of constructions of the concepts needed for the description of the manifold radiative transfer phenomena encountered in the practice of radiative transfer theory. In particular in this section we shall use the interaction principle as a base for the construction of the more commonly used surface reflectance and transmittance concepts. Some work has already been done in this direction in Sec. 3.1. In fact the empirical reflectance function was defined in that section as a necessary prerequisite for the construction of the preliminary example of the interaction principle. We now return to that setting for the purpose of establishing systematic definitions for the family of reflectance and transmittance operators for surfaces.

Geometrical Conventions

Figure 3.3 (a) depicts a general surface Y in an optical medium X and a relatively small part S of Y about point x on Y. We are interested in the reflectance and transmittance of Y in the region S about x. Now the terms "transmittance" and "reflectance" become meaningful only after adequate reference frames have been established at given points x of Y within which one can unambiguously establish conventions about the notions of "inwardness", "outwardness", "upwardness", "downwardness", "forwardness", "backwardness", etc. Suppose then we affix to point x of Y a unit vector $\mathbf{k}(x)$ and call it the *unit outward normal* to Y at x. Perhaps some readers would prefer to call $-\mathbf{k}(x)$ the unit outward normal to Y at x. This is perfectly admissible for our present purposes, and the reader may therefore turn around the arrows in parts (a)-(d) of Fig. 3.3 and read the following discussion as it stands. The point being made here is that what one calls "outward", etc., is immaterial. What does matter is what one subsequently does with the concept and that, within a given discussion, a measure of consistency is sustained in the use of the concept once the convention is made.

During the present discussion, let "D'" and "D" denote narrow circular conical solid angles of central directions ξ' and ξ, respectively. S is a small collecting surface on Y, and x is a point of Y in S. Let "S'" denote the projection of S on a plane normal to ξ'. (See parts (c) and (d) of Fig. 3.3.) D' is the set of *incident directions*; D is the set of *response directions*. Both D' and D will always lie completely within $\Xi_+(\mathbf{k}(x))$ or $\Xi_-(\mathbf{k}(x))$ where $\Xi_+(\mathbf{k}(x))$ is the set of all directions ξ' such that $\xi' \cdot \mathbf{k}(x) > 0$, and $\Xi_-(\mathbf{k}(x))$ is the set

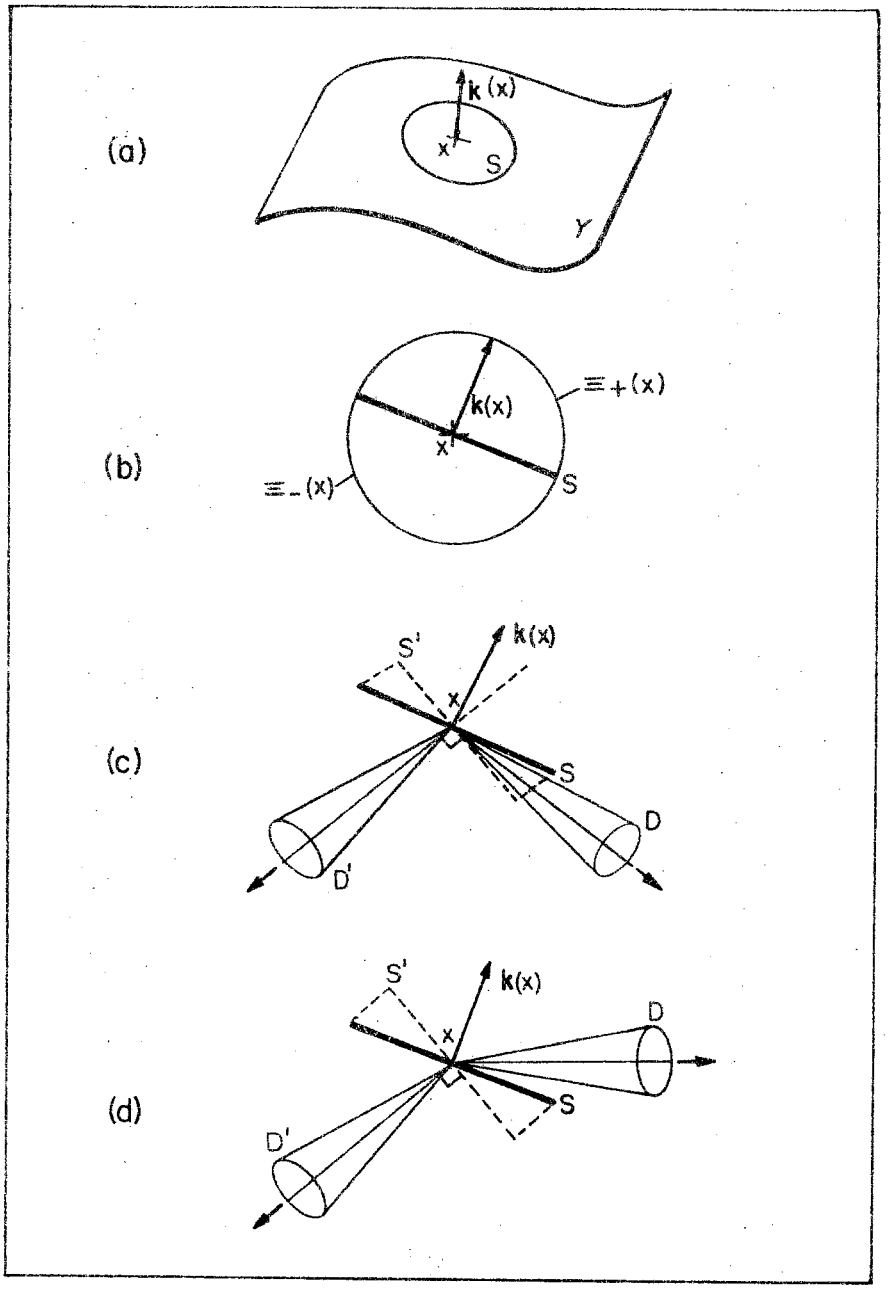


FIG. 3.3 Setting for reflectance and transmittance operators for surfaces.

of all directions ξ' such that $\xi' \cdot \mathbf{k}(x) < 0$. (See part (b) of Fig. 3.3 and compare with Sec. 2.4, so that $\Xi_+(\mathbf{k}(x)) = \Xi(\mathbf{k}(x))$ and $\Xi_-(\mathbf{k}(x)) = \Xi(-\mathbf{k}(x))$. We shall also write for brevity:

$$\text{"}\Xi_\pm(x)\text{"} \quad \text{for} \quad \Xi_\pm(\mathbf{k}(x)) \quad .$$

The notation "$\Xi_\pm(x)$" finds its best use when specific surfaces are under consideration, while the notation "$\Xi(\pm\mathbf{k}(x))$" finds its greatest use when (as in Sec. 2.4) purely radiometric arguments are in effect as no specific surfaces are being discussed.

The Empirical Reflectances and Transmittances

With these preliminaries established we can define with some measure of precision the empirical reflectance and transmittance function. Emulating (1) of Sec. 3.1 we write:

$$\text{"}s(S',D';S,D)\text{"} \quad \text{for} \quad \frac{N(S',D';S,D)}{N(S',D')\Omega(D')} \tag{1}$$

where all terms on the right side of the definition are as described in Sec. 3.1, but now with x, ξ', ξ, S', D', S, and D as specified above. The notation in (1) does not tell us specifically on which side of S the sets D' and D lie. By specifying this information, the values $s(S',D';S,D)$ take on the characteristics of reflectances and transmittances. Thus let us write:

$$\text{"}r_+(S',D';S,D)\text{"} \quad \text{for} \quad s(S',D';S,D), \text{ if } D' \subset \Xi_+(x) \text{ and } D \subset \Xi_-(x) \tag{2}$$

$$\text{"}r_-(S',D';S,D)\text{"} \quad \text{for} \quad s(S',D';S,D), \text{ if } D' \subset \Xi_-(x) \text{ and } D \subset \Xi_+(x) \tag{3}$$

$$\text{"}t_+(S',D';S,D)\text{"} \quad \text{for} \quad s(S',D';S,D), \text{ if } D' \subset \Xi_+(x) \text{ and } D \subset \Xi_+(x) \tag{4}$$

$$\text{"}t_-(S',D';S,D)\text{"} \quad \text{for} \quad s(S',D';S,D), \text{ if } D' \subset \Xi_-(x) \text{ and } D \subset \Xi_-(x) \tag{5}$$

Here "$D' \subset \Xi_+(x)$" is an inclusion statement which means that D' is contained in $\Xi_+(x)$. Similar interpretations hold for the other three inclusion statements. For example, part (c) of Fig. 3.3 depicts the geometrical arrangement for $t_-(S',D';S,D)$, and part (d) of Fig. 3.3 depicts the arrangement for $r_-(S',D';S,D)$. Definitions (2) and (4) cover the *outward* (or *upward* or *forward*) empirical *reflectance* and *transmittance* of Y over S. Properties (i) and (ii) in Sec. 3.1 hold for the $r_\pm$ and $t_\pm$ just defined.

We could have arrived at the preceding four empirical reflectance and transmittance functions just above by direct appeal to the interaction principle. Thus, with X and S as given, let $m = n = 1$ and A be the set of all outward directed incident radiances $N(S',D')$ on S (i.e., D' contained in $\Xi_+(x)$), and let B be the set of all inward directed response

radiances of S (i.e., D contained in $\Xi_-(x)$). Then the interaction principle asserts the existence of a linear interaction operator s_{11}--call it "$r_+(\cdot,\cdot;\cdot,\cdot)\Omega(\cdot)$"--with the property that for every N(S',D') in A there exists an N(S,D) in B such that:

$$N(S,D) = N(S',D')r_+(S',D';S,D)\Omega(D') \quad .$$

Hence the interaction operator in this instance is a real valued function of four variables (S',D',S,D) which assigns to each choice of these variables a number--the reflectance of Y over S under the indicated conditions. If instead of incident radiances, we chose incident scalar irradiances over D' for the set A, then $r_+(\cdot,\cdot;\cdot,\cdot)$ itself would have been obtained. If we had chosen incident irradiances instead, then $r_+(\cdot,\cdot;\cdot,\cdot)/\xi'\cdot k$ would have been obtained. This shows the potential flexibility of the principle in supplying a great variety of "reflectances", depending on what set of radiometric quantities are chosen for A and for B.

The Theoretical Reflectances and Transmittances

By letting S approach {x}, D' approach {ξ'}, and D approach {ξ} in the limit, definitions (2)-(5) yield definitions of the corresponding *theoretical reflectances and transmittances* of Y at x. Thus by performing the indicated limit operations, we arrive at:

$$r_+(x;\xi';\xi) \quad \text{if} \quad \xi' \in \Xi_+(x) \quad \text{and} \quad \xi \in \Xi_-(x) \tag{6}$$

$$r_-(x;\xi';\xi) \quad \text{if} \quad \xi' \in \Xi_-(x) \quad \text{and} \quad \xi \in \Xi_+(x) \tag{7}$$

$$t_+(x;\xi';\xi) \quad \text{if} \quad \xi' \in \Xi_+(x) \quad \text{and} \quad \xi \in \Xi_+(x) \tag{8}$$

$$t_-(x;\xi';\xi) \quad \text{if} \quad \xi' \in \Xi_-(x) \quad \text{and} \quad \xi \in \Xi_-(x) \tag{9}$$

Here "$\xi' \in \Xi_+(x)$" means that ξ' is a direction in $\Xi_+(x)$, etc.

It is a simple matter to show how these theoretical reflectance and transmittance functions for surfaces follow directly from the interaction principle. The technique of obtaining $r_\pm$ or $t_\pm$ is similar to that discussed in Sec. 2.13 for obtaining the generalized luminosity function $\bar{z}(\cdot)$. Specifically, we would use the interaction principle to supply a positive linear function with the property that it acts on incident radiance distributions and yields reflected or transmitted radiance distributions. Interested mathematical readers may pursue this matter further in Sec. 3.16. To develop this application of the interaction principle in the present section would be to digress too far from the chosen scope of the present discussions. We give only the results of such an excursion into measure theory. Thus, we write:

$$\text{"}r_{\pm}(Y)\text{"} \quad \text{for} \quad \int_{\Xi_{\pm}(Y)} [\ ]\, r_{\pm}(\cdot;\xi';\xi)d\Omega(\xi') \tag{10}$$

$$\text{"}t_{\pm}(Y)\text{"} \quad \text{for} \quad \int_{\Xi_{\pm}(Y)} [\ ]\, t_{\pm}(\cdot;\xi';\xi)d\Omega(\xi'). \tag{11}$$

These are the general *reflectance* and *transmittance* integral operators associated with an arbitrary surface Y with outward unit normal $\mathbf{k}(x)$ at each point x of Y. The domain of integration in each operator is of the form $\Xi_+(Y)$ or $\Xi_-(Y)$ and is known once x in Y is specified. Thus, if $N(x,\cdot)$ is an inward incident radiance distribution at x in Y, then:

$$\int_{\Xi_-(x)} N(x,\xi')r_-(x;\xi';\xi)d\Omega(\xi') \tag{12}$$

is the outward reflected radiance at x in the direction ξ in response to $N(x,\cdot)$. In general, if $N(x,\cdot)$ and $r_{\pm}$ and $t_{\pm}$ are defined over just part a of Y, then we use "$N_-(a)$" to denote inward incident or response radiance distributions over part a, and "$N_+(a)$" to denote outward incident or response radiance distributions over part a. For example, if x is a point of a and ξ is an outward direction, then $N_+(a)$ assigns to x and ξ the response radiance $N(x,\xi)$. If we let x in (12) range over all points of part a of Y, then we see that (12) defines the response function $N_+(a)$ of a. Hence $N_+(a)$ in this instance is a general reflected radiance distribution resulting from operating on $N_-(a)$ by $r_-(a)$. This fact we write in the form:

$$N_+(a) = N_-(a)r_-(a) \tag{13}$$

where we have written:

$$\text{"}N_-(a)r_-(a)\text{"} \quad \text{for} \quad \int_{\Xi_-(a)} N(\cdot,\xi')r_-(\cdot;\xi',\xi)d\Omega(\xi') \tag{14}$$

The radiance distribution appearing in the integral is (by noting that the range of integration is $\Xi_-(a)$) an inward radiance distribution incident on a at a general (unspecified) point. The definition (14) can be repeated for the three other general cases associated with a, namely $N_+(a)r_+(a)$, $N_-(a)t_-(a)$, $N_+(a)t_+(a)$. Equation (13) gives the integration operation an algebraic appearance, a feature which, as we shall see, is most conducive to rapid and creative manipulations during theoretical radiative transfer computations. This algebraization of radiative transfer theory is fostered by the interaction principle whose salient character is itself basically algebraic (rather than, say, analytic or geometric).

Variations of the Basic Theme

Some attention will next be given to the possible variations the preceding definitions of $r_\pm$ and $t_\pm$ may undergo as shifts are made in the choice of types of radiometric incident and response quantities. A few specific instances will suffice to show the potentially great number of variations possible.

To begin, suppose that the radiometric quantities in the incident set A are to be *irradiances* and those in the response set B to be *radiances*. Then, e.g., in the expanded rendition of (13):

$$N(x,\xi) = \int_{\Xi_-(x)} N(x,\xi')r_-(x;\xi';\xi)d\Omega(\xi') \quad ,$$

we rearrange matters so:

$$N(x,\xi) = \int_{\Xi_-(x)} N(x,\xi')|\xi'\cdot\mathbf{k}(x)| \frac{r_-(x;\xi';\xi)}{|\xi'\cdot\mathbf{k}(x)|} d\Omega(\xi') \qquad (15)$$

with the result that the new reflectance operator has a kernel with values of the form:

$$\frac{r_-(x;\xi';\xi)}{|\xi'\cdot\mathbf{k}(x)|} \quad . \qquad (16)$$

We shall not devise notation to cover this case or the multitude of alternate cases possible. The notation is best settled by those who must work repeatedly with the specialized concepts. A semblance of order and universality is attained in such matters, however, if some set of functions such as those defined via (6)-(9) is taken as a fixed base of operations from which to proceed to new territory.

Reflectance functions of the form displayed in (16) are used in practice where the surfaces under study are often considered ideally or nearly *uniform* (or *lambert*) *reflectors*. For suppose a surface Y at x has the property that there is a real number r_- such that:

$$\frac{r_-(x;\xi';\xi)}{|\xi'\cdot\mathbf{k}(x)|} = \frac{r_-}{\pi} \qquad (17)$$

for every ξ' in $\Xi_-(x)$ and every ξ in $\Xi_+(x)$. Then (15) becomes:

$$N(x,\xi) = \frac{r_-}{\pi}\int_{\Xi_-(x)} N(x,\xi')|\xi'\cdot\mathbf{k}(x)|d\Omega(\xi')$$

$$= \frac{r_-}{\pi} H(x,\Xi_-(x)) \quad .$$

Hence the reflected radiance distribution $N(x,\cdot)$ is uniform (independent of ξ) of magnitude $N(x)$, say. Then the associated radiant emittance is:

$$W(x,\Xi_-(x)) = \pi N(x)$$

$$= r_-H(x,\Xi_-(x)) \quad ,$$

as one would expect by the way r_- is defined. If the incident radiance distribution itself was uniform, of magnitude $N'(x)$ then

$$H(x,\Xi_-(x)) = \pi N'(x) \quad .$$

From this and the preceding equation we have:

$$N(x) = r_-N'(x) \quad ,$$

again as one would expect of the new version of the reflectance function and a lambert reflector.

As another example, suppose that the incident radiometric quantities in A are radiances and those in B are radiant emittances. Specifically, let (15) be used as starting point and operate on each side of (15) with an integration of the kind:

$$\int_{\Xi_+(x)} N(x,\xi)\xi\cdot\mathbf{k}(x)d\Omega(\xi) =$$

$$= \int_{\Xi_+(x)}\left[\int_{\Xi_-(x)} N(x,\xi')r_-(x;\xi';\xi)d\Omega(\xi')\right]\xi\cdot\mathbf{k}(x)d\Omega(\xi) \quad . \quad (18)$$

It is clear that the integral on the left yields the requisite radiant emittance $W(x,\Xi_+(x))$ (cf. (22) of Sec. 2.4) which thus is obtained by operating on the incident radiance distribution $N(x,\cdot)$ with the integral operator

$$\int_{\Xi_+(x)}\left[\int_{\Xi_-(x)} [\,]r_-(x;\xi';\xi)d\Omega(\xi')\right]\xi\cdot\mathbf{k}(x)d\Omega(\xi) \quad .$$

Now it is quite natural when using irradiance and radiant emittance in this way for us to assign to the quotient

$$W(x,\Xi_+(x))/H(x,\Xi_-(x))$$

the meaning of a reflectance (an *albedo*) of the surface Y at x. Thus if we write:

"$r_-(x)$" for

$$\frac{1}{H(x, \Xi_-(x))} \cdot \int_{\Xi_+(x)} \left[\int_{\Xi_-(x)} N(x,\xi') r_-(x;\xi';\xi) d\Omega(\xi') \right] \xi \cdot k(x) \, d\Omega(\xi)$$

then we can go on to rearrange (18) into the form:

$$W(x, \Xi_+(x)) = H(x, \Xi_-(x)) r_-(x) \quad . \tag{19}$$

This definition of $r_-(x)$ (and the three analogous definitions $r_+(x)$, $t_\pm(x)$) is motivated by the need for working with numerical irradiances and radiant emmittances, and numerical reflectances rather than the analogous functional and operatorial concepts which must be used in certain full treatments of interreflection problems. In the next section, we shall illustrate in more detail the use of (13) and (19).

3.4 Applications to Plane Surfaces

In this section we shall illustrate the application of the reflectance and transmittance operators for surfaces, constructed in Sec. 3.3, for several types of frequently encountered plane-surface settings in radiative transfer theory. Throughout this section and, indeed, the remainder of this chapter, one of the principal goals is the demonstration of the systematic use to which the interaction principle may be put in formulating the concepts and problems of radiative transfer theory.

Example 1: Irradiances on Two Infinite Parallel Planes

Let "a" and "b" denote two infinite parallel plane surfaces separated by a vacuum, as in Fig. 3.4. The coordinate system used is the terrestrial system defined in Sec. 2.4. Each plane has assigned reflectance and transmittance functions as developed in Sec. 3.3 which are to be constant over a and b. However, the directional structures of the reflectance and transmittance functions are otherwise arbitrary. An interreflection process between a and b is initiated and sustained by a steady downward field radiance distribution $N^0(b)$ on plane b which has the same structure at all points of b. Our present goal is to compute the resultant steady state irradiances on a and b, that is the upward irradiance $H_+(a)$ on a and the downward irradiance $H_-(b)$ on b.

The interaction principle applied to a and b in turn yields the requisite irradiance reflectance operators. Thus for a the set A of incident radiometric functions consists of irradiances like $H_+(a)$, and the set B of response radiometric functions of a consists of downward radiant emittances $W_-(a)$ which by the hypothesized vacuum between a and b have magnitudes equal to $H_+(a)$. (Cf. also example 12 of Sec. 2.11 [illegible] with r in the [illegible] of infinite

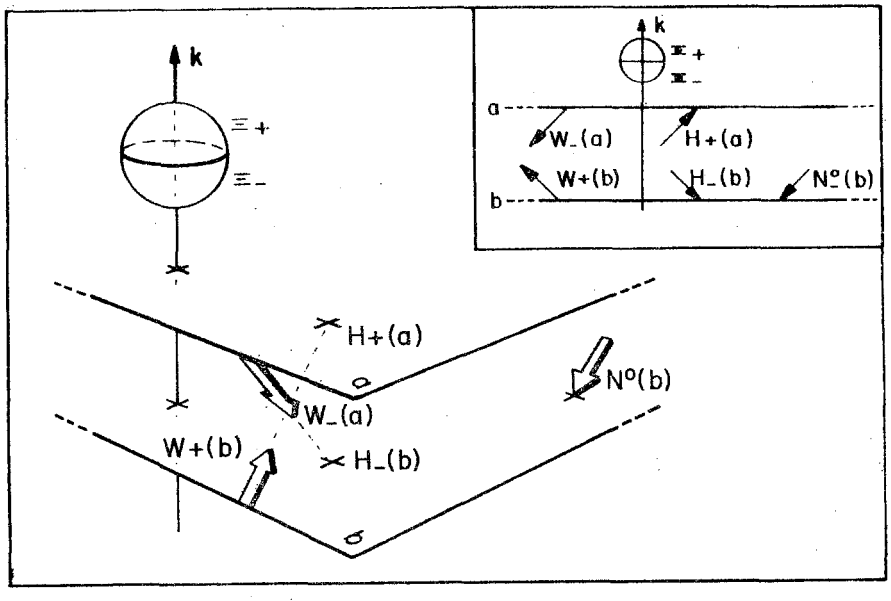


FIG. 3.4 Two interreflecting parallel planes.

planes). The interaction principle then asserts the existence of a reflectance (a number) $r_+(a)$ such that:

$$W_-(a) = H_-(b) = H_+(a)r_+(a)$$
$$= W_+(b)r_+(a) \qquad (1)$$

The last equality uses the radiance invariance law which implies that $W_+(b) = H_+(a)$. The closing example of Sec. 3.3 shows the necessary form of $r_+(a)$. Thus, following the pattern (19) of Sec. 3.3 we have written:

$$\text{"}r_+(a)\text{"} \quad \text{for} \quad \frac{1}{H_+(a)}\int_{\Xi_-}\left[\int_{\Xi_+} N(x,\xi')r_+(x;\xi';\xi)\,d\Omega(\xi')\right]\xi\cdot(-\mathbf{k})\,d\Omega(\xi)$$

and:

$$\text{"}H_+(a)\text{"} \quad \text{for} \quad \int_{\Xi_+} N(x,\xi')\xi'\cdot\mathbf{k}\;d\Omega(\xi')$$

where $N(x,\cdot)$ is now the upward surface radiance distribution of b at x which, with $r_+(x;\xi';\xi)$, is independent of x. By noting that the iterated integration amounts to finding $W_-(a)$, the downward radiant emittance of a, we see that we are simply writing:

$$\text{"}r_+(a)\text{"} \quad \text{for} \quad \frac{W_-(a)}{H_+(a)} \quad .$$

However, $r_+(a)$ is now precisely determinable as shown in the iterated integration whenever $N(x,\xi)$ and $r_+(x;\xi';\xi)$ are known for every x,ξ', and ξ. Even if the surface radiances $N(x,\xi)$ of b (and also a) are not known in absolute magnitude, but only in relative magnitude (i.e., its shape but not the size is known) the present goal can be attained, as we shall see.

Continuing, we apply the interaction principle to plane b, which has two sets of incident functions and one set of response functions. For, the given downward surface radiance $N_-^o(b)$ on b gives rise to a known incident irradiance $H_-^o(b)$. Irradiances like $H_-^o(b)$ comprise the set A_1 of incident radiometric functions for b. Irradiances like $H_-(b)$ comprise the set A_2 of incident radiometric functions for b. The set B_1 of response functions of b consists of radiant emittances $W_+(b)$ numerically equal to $H_+(a)$,(via the radiance invariance law once again). The interaction principle then yields two reflectances (numbers) $r_-^o(b)$ (for A_1 and B_1) and $r_-(b)$ (for A_2 and B_1) such that:

$$W_+(b) = H_+(a) = H_-^o(b)r_-^o(b) + H_-(b)r_-(b) \quad . \qquad (2)$$

The numbers $r_-^o(b)$ and $r_-(b)$ are defined exactly analogously to $r_+(a)$. Equations (1) and (2) together determine $H_+(a)$. Thus, using (1) to eliminate $H_-(b)$ from (2), we have:

$$H_+(a) = H_-^o(b)r_-^o(b) + (H_+(a)r_+(a))r_-(b) \quad ,$$

whence:

$$\boxed{W_+(b) = H_+(a) = H_-^o(b)r_-^o(b)/[1-r_+(a)r_-(b)]} \qquad (3)$$

and so:

$$\boxed{W_-(a) = H_-(b) = H_-^o(b)r_-^o(b)r_+(a)/[1-r_+(a)r_-(b)]} \qquad (4)$$

These solutions exist provided that the product $r_+(a)r_-(b)$ is less than 1. This provision is reminiscent of a similar provision for Σ_{121} and Σ_{212} encountered in the preliminary example of Sec. 3.1, and may also be handled via the energy conservation law if desired. It is clear that (3) and (4) are usable in practice once reasonable estimates of $r_+(a)$ and $r_-(b)$ are made. Such estimates can be based either on empirical data in the form of measured ratios such as $W_-(a)/H_+(a)$, or by means of integral computations knowing the values $r_+(x;\xi';\xi)$ and the *shape* of the reflected radiance distributions. For example one can assume the perennial favorite: a uniform radiance distribution, or other readily integrated products of the form $N(x,\xi')r_+(x;\xi';\xi)$.

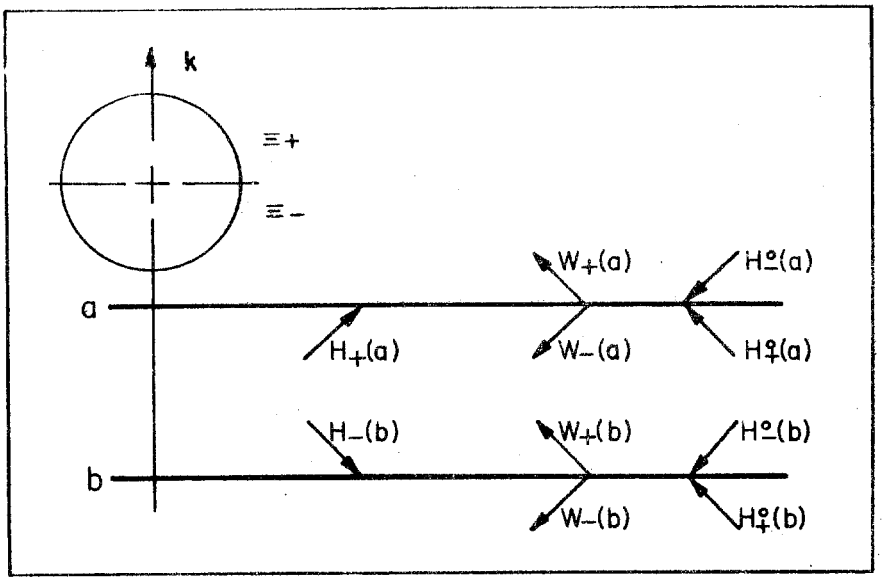


FIG. 3.5 Systematic details for an interreflection calculation between two parallel planes.

Example 2: Irradiances on Two Infinite Parallel Planes, Reexamined

In this example, we systematize the procedure and results of Example 1. In that example the radiometric details were kept at an absolute minimum so that the algebraic workings of the interaction principle could be readily followed. Now that the algebraic details of the interaction formulation have been demonstrated, we return to that simple setting and pull out nearly all the radiometric stops and turn on all the lights--so to speak. Specifically, we now let plane a be irradiated by two external sources, (i.e., origins of flux other than a and b) which produce downward $H^{o}_{-}(a)$ and upward $H^{o}_{+}(a)$ irradiances; similarly, b is irradiated by two external sources which produce $H^{o}_{-}(b)$ and $H^{o}_{+}(b)$ as schematically shown in Fig. 3.5. Our present goal is to use the interaction principle to formulate the equations governing the four quantities: $W_{\pm}(a)$, $W_{\pm}(b)$, i.e., the upward (+) and downward (-) radiant emittances of a and b induced by the interreflection interaction between a and b and the incident external sources on a and b. We direct attention first to plane a and list all possible incident radiometric quantities on a:

A_1: all irradiances like $H^{o}_{-}(a)$

A_2: all irradiances like $H^{o}_{+}(a)$

A_3: all irradiances like $H_{+}(a)$

A_1 and A_2 are self explanatory; A_3 is the set of irradiances induced by the presence of plane b below a. Next, the set of all response radiometric quantities of a are enumerated as follows:

B_1: all radiant emittances like $W_+(a)$

B_2: all radiant emittances like $W_-(a)$

Thus, in the case of plane a, $m = 3$, $n = 2$, and the six abstract interaction operators s_{ij} supplied by the interaction principle are in the form of reflectance and transmittance numbers as follows:

$$s_{11} \text{ -- } r_-^o(a)$$

$$s_{12} \text{ -- } t_-^o(a)$$

$$s_{21} \text{ -- } t_+^o(a)$$

$$s_{22} \text{ -- } r_+^o(a)$$

$$s_{31} \text{ -- } t_+(a)$$

$$s_{32} \text{ -- } r_+(a)$$

The six numbers $r_-^o(a),\ldots,r_+(a)$ are defined exactly analogously to $r_+(a)$ in Example 1 and come ultimately from the interaction principle as outlined in Sec. 3.3. The superscripts "o" set off the external incident sources from the internal sources. Then, according to the interaction principle $W_+(a)$ and $W_-(a)$ are given by:

$$W_+(a) = H_-^o(a)r_-^o(a) + H_+^o(a)t_+^o(a) + H_+(a)t_+(a) \qquad (5)$$

$$W_-(a) = H_-^o(a)t_-^o(a) + H_+^o(a)r_+^o(a) + H_+(a)r_+(a) \ . \qquad (6)$$

By repeating this process of application of the interaction principle to plane b we arrive at the analogous pair of statements:

$$W_+(b) = H_+^o(b)t_+^o(b) + H_-^o(b)r_-^o(b) + H_-(b)r_-(b) \qquad (7)$$

$$W_-(b) = H_+^o(b)r_+^o(b) + H_-^o(b)t_-^o(b) + H_-(b)t_-(b) \ . \qquad (8)$$

When we append the following two equations:

$$W_+(b) = H_+(a) \qquad (9)$$

$$W_-(a) = H_-(b) \ , \qquad (10)$$

which follow from the hypothesized vacuum between a and b and the radiance invariance law, the resulting system (5)-(10) is self-contained and in principle solvable. In particular when $W_-(a)$ and $W_+(b)$ in (6) and (7) are eliminated via (9) and (10), the resultant pair of equations is autonomous:

$$H_-(b) = A_- + H_+(a)r_+(a) \tag{11}$$

$$H_+(a) = B_+ + H_-(b)r_-(b) \tag{12}$$

where we have written:

"A_-" for $H^o_-(a)t^o_-(a) + H^o_+(a)r^o_+(a)$

"B_+" for $H^o_+(b)t^o_+(b) + H^o_-(b)r^o_-(b)$.

Using (11) eliminate $H_-(b)$ from (12):

$$H_+(a) = B_+ + [A_- + H_+(a)r_+(a)]r_-(b) ,$$

we readily solve for $H_+(a)$:

$$\boxed{H_+(a) = \left[B_+ + A_-r_-(b)\right]\Big/\left[1 - r_+(a)r_-(b)\right]} . \tag{13}$$

Then by (9) we obtain $W_+(b)$. From (13) and (11) we find $H_-(b)$ and so by (10), $W_-(a)$. Equations (5) and (8) then yield $W_+(a)$ and $W_-(b)$. In this way all four radiant emittances $W_\pm(a)$ and $W_\pm(b)$ are determined.

A First Synthesis of the Interaction Method

This example is valuable in pointing up the systematic use to which the interaction principle may be put in formulating and solving a radiative transfer problem associated with a subset S of an optical medium X. The essential steps of this method exhibited by the preceding example are as follows:

(i) Isolate the subset S of the optical medium X.

(ii) Enumerate the incident radiometric quantities a_i on S. This determines $A_1,\ldots,A_m$.

(iii) Enumerate the requisite response radiometric quantities b_j of S. This determines $B_1,\ldots,B_n$.

(iv) Enumerate the mn operators s_{ij}, $i = 1,\ldots,m$, $j = 1,\ldots,n$ supplied by the interaction principle.

(v) Write the interaction equation $b_j = \sum_{i=1}^{m} a_i s_{ij}$ for $j = 1,\ldots,n$.

(vi) Append auxiliary equations connecting various chosen a_i and b_j, in as much detail as required to solve the system in (v) for the b_j.

Step (vi) in the present example occurred in (9) and (10) above. Invariably, the additional auxiliary equations in (vi) are equations which match radiances on adjoining subsets of X and use one or the other of the following laws:

(a) The radiance invariance law

(b) The equality of field and surface radiance at a given point and for a given direction.

The six steps (i)-(vi) together with (a) and (b) above will be used time and again in the following examples. These steps appear to lead to systematic formulations of radiative transfer problems in a manner similar to that used in the formulation of the problems of statics and dynamics in mechanics, i.e., by using the technique which begins with the time-honored injunction to: "isolate the body", then categorically adding up all forces on the isolated body, and finally applying one or all of the three basic Newtonian laws of mechanics to the isolated system. It is somewhat amusing and perhaps of interest to observe that the three Newtonian laws even appear to have their explicit radiometric counterparts in the form of (a) above for the First Law, (v) above for the Second law, and (b) above for the Third law. We shall call the method of formulation summarized in (i)-(vi) and (a), (b) above the *method of the interaction principle*, or simply the *interaction method.**

Example 3: Irradiances on Finitely Many Infinite Parallel Planes

What we have done above for two plane surfaces we can in principle do again for any finite number and even an infinite number of plane surfaces. We now consider the case of finitely many parallel planes mainly for the novel problems of solution it presents subsequent to the invocation of the interaction principle. This will serve to show that the

---

*In studies of linear hydrodynamics subsequent to the completion of the present work, I have found that the interaction method is capable of unifying this field in an elegant and practical manner, and that it leads to detailed numerical descriptions of scattered fields of surface water waves. Further studies in water wave-guide theory show similarities to the scattering matrix method in e.m. wave-guide propagation. All of this is not surprising, as the wording of the interaction principle is quite wide, and will apply to these other contexts by changing "radiometric" along with "optical medium" appropriately. See, e.g., Preisendorfer, R.W., "Surface-wave transport in nonuniform canals," Report NOAA/JTRE-80, Hawaii Institute of Geophysics, 1972.

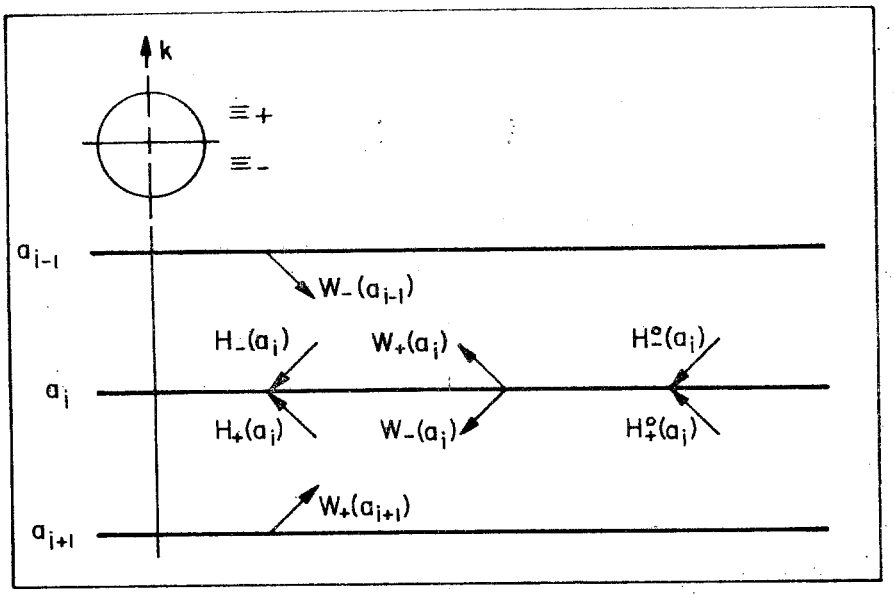


FIG. 3.6 Interaction calculation details for finitely many parallel planes.

interaction principle can lead even the most assiduous investigator only so far: there will always be a need for effective solution procedures of the more complex formulations supplied by the method of the interaction principle.

Figure 3.6 depicts three adjacent, parallel planes a_{i-1}, a_i, and a_{i+1} in a family of p parallel planes separated by vacua. Hence $1 < i < p$ in the Figure. For the moment, p may be either finite or infinite. Each plane a_i, $i = 1,\ldots,p$ is generally irradiated by an external source and the irradiation is constant over the extent of each a_i. The reflectance and transmittance functions of the a_i are also independent of x over a_i. Let "$H^o_\pm(a_i)$" denote the upward (+) and downward (-) external source irradiance on a_i. Furthermore, for every internal plane a_i, $1 < i < p$, there are irradiances $H_\pm(a_i)$ produced by flux from its lower neighbor (a_{i+1}) and its upper neighbor (a_{i-1}). Therefore for each isolated internal plane a_i there are four sets of incident radiometric quantities:

$1 < i < p$

A_1: all irradiances like $H^o_-(a_i)$

A_2: all irradiances like $H^o_+(a_i)$

A_3: all irradiances like $H_-(a_i)$

A_4: all irradiances like $H_+(a_i)$

The set of all response radiometric quantities of a_i are enumerated as follows:

$1 < i < p$

B_1: all radiant emittances like $W_+(a_i)$

B_2: all radiant emittances like $W_-(a_i)$

Thus for the case of subset a_i of X, $1 < i < p$, we have $m = 4$ and $n = 2$, and the eight abstract interaction operators s_{ij} supplied by the interaction principle are specifically of the following forms:

$$s_{11} \text{ -- } r_-^o(a_i)$$

$$s_{12} \text{ -- } t_-^o(a_i)$$

$$s_{21} \text{ -- } t_+^o(a_i)$$

$$s_{22} \text{ -- } r_+^o(a_i)$$

$$s_{31} \text{ -- } r_-(a_i)$$

$$s_{32} \text{ -- } t_-(a_i)$$

$$s_{41} \text{ -- } t_+(a_i)$$

$$s_{42} \text{ -- } r_+(a_i)$$

The eight numbers $r_-^o(a_i),\ldots,r_+(a_i)$ for a given a_i are defined exactly as $r_+(a)$ was defined in Example 1. Then, according to step (2) of the method of the interaction principle (as outlined in Example 2), we have the following two equations for $W_\pm(a_i)$:

$$W_+(a_i) = H_-^o(a_i)r_-^o(a_i) + H_+^o(a_i)t_+^o(a_i) + H_-(a_i)r_-(a_i) + H_+(a_i)t_+(a_i) \tag{14}$$

$$W_-(a_i) = H_-^o(a_i)t_-^o(a_i) + H_+^o(a_i)r_+^o(a_i) + H_-(a_i)t_-(a_i) + H_+(a_i)r_+(a_i) \tag{15}$$

$1 < i < p$

The interaction principle is now applied to planes a_1 and a_p in turn. In the case of a_1, the five steps of the method of the interaction principle yield:

$$W_+(a_1) = H_-^o(a_1)r_-^o(a_1) + H_+^o(a_1)t_+^o(a_1) + H_+(a_1)t_+(a_1) \tag{16}$$

$$W_-(a_1) = H_-^o(a_1)t_-^o(a_1) + H_+^o(a_1)r_+^o(a_1) + H_+(a_1)r_+(a_1) \tag{17}$$

which are (and should be) identical in form to (5) and (6) of Example 2. As might now be expected the radiant emittance equations for a_p, are identical in form to (7) and (8) of Example 2:

$$W_+(a_p) = H^o_+(a_p)t^o_+(a_p) + H^o_-(a_p)r^o_-(a_p) + H_-(a_p)r_-(a_p) \qquad (18)$$

$$W_-(a_p) = H^o_+(a_p)r^o_+(a_p) + H^o_-(a_p)t^o_-(a_p) + H_-(a_p)t_-(a_p) \qquad (19)$$

In this way, for finite p, we arrive at p pairs of equations for $W_\pm(a_i)$, $i = 1,\ldots,p$. Step (vi) of the interaction principle method (in particular law (a)) yields the following 2(p-1) auxiliary equations:

$$W_+(a_i) = H_+(a_{i-1}) \ , \quad i = 2,\ldots,p \qquad (20)$$

$$W_-(a_i) = H_-(a_{i+1}) \ , \quad i = 1,\ldots,p-1 \qquad (21)$$

Using (20) and (21) in (14), (15), (17), and (18) we arrive at the following set of 2(p-1) equations in 2(p-1) unknowns $H_\pm(a_i)$:

$1<i<p$

$$H_-(a_2) = A_-(a_1) + H_+(a_1)r_+(a_1) \qquad (22)$$

$$H_+(a_{i-1}) = A_+(a_i) + H_-(a_i)r_-(a_i) + H_+(a_i)t_+(a_i) \qquad (23)$$

$$H_-(a_{i+1}) = A_-(a_i) + H_-(a_i)t_-(a_i) + H_+(a_i)r_+(a_i) \qquad (24)$$

$$H_+(a_{p-1}) = A_+(a_p) + H_-(a_p)r_-(a_p) \qquad (25)$$

where for every i, $1 \le i \le p$, we have written:

$$\text{"}A_+(a_i)\text{"} \quad \text{for} \quad H^o_-(a_i)r^o_-(a_i) + H^o_+(a_i)t^o_+(a_i)$$

$$\text{"}A_-(a_i)\text{"} \quad \text{for} \quad H^o_-(a_i)t^o_-(a_i) + H^o_+(a_i)r^o_+(a_i)$$

The system of equations (22)-(25) is very nearly a diagonal system, so can be solved relatively easily by successive elimination, or by other well-known methods of solution for such a system. However, it is interesting to note that a general recursion procedure for solving the system is suggested by the following physical observations. Suppose we conceptually view the system of p planes as a new system of two interacting subsets namely a_1 and the remaining set $\{a_2,\ldots,a_p\}$ of p-1 planes. Considering $\{a_2,\ldots,a_p\}$ as a single unit is reminiscent of considering the two surfaces S_1 and S_2 in the preliminary example of Sec. 3.1 as a single unit. By applying the interaction principle to a_1 we obtain the equations for $W_\pm(a_1)$ whose forms are precisely those of (5), (6). By applying the interaction method to the set $\{a_2,\ldots,a_p\}$, which we will also call by the new name "b", we obtain equations for $W_\pm(b)$ which are precisely those in (7), (8) provided that all internal sources on $\{a_2,\ldots,a_p\}$ are shut off. We defer treatment of the internal-source case for the moment. Then $\{a_2,\ldots,a_p\}$ reacts radiometrically to irradiation as would a single plane. That is, $\{a_2,\ldots,a_p\}$ has its own reflectance and transmittance functions assigned to it by the interaction

principle and hence its associated reflectance and transmittances for irradiance are presumed known. It follows that equation (13) yields $H_+(a_1)$ and hence $W_+(b)$, etc., provided that $r_-(b)$ --i.e., the reflectance of the set $\{a_2,\ldots,a_p\}$ for downward irradiance is known. We do not know $r_-(b)$ as such, but at any rate we have gone down the ladder of complexity one rung: our original task was to find the reflectance and transmittance properties of a set of p parallel planes. Now we have only to find them for a set of p-1 planes. The course of action now before us is clear: we use (13) and its related equations applied to a_{p-1} and a_p to obtain the reflectance and transmittance of the set $\{a_{p-1},a_p\}$. Then we use (13) again to find the reflectance and transmittance of $\{a_{p-2},a_{p-1},a_p\}$, and so on, adding layer by layer until we return to the system comprised of a_1 and $\{a_2,\ldots,a_p\}$ and perform the final step. The preceding technique has been hastily sketched rather than developed in detail because it is only a special case of a more general problem, the complete details of which have been worked out elsewhere using the interaction principle and thus need not be repeated here. The reader may consult Chapters IX and X of Ref. [251] for a categorical analysis of this internal-source problem. Furthermore, for several succinct formulas summarizing the preceding source-free analysis of the system $\{a_1,\ldots,a_p\}$, see Example 6 of Sec. 8.7.

Despite the fact that the system (22)-(25) can be subsumed under a more general completed analysis, the reader should try his own hand at solving the system *de novo*, first for the case of all internal sources zero (i.e., $H^o_+(a_i) = 0$, $i = 1,\ldots,p\text{-}1$ and $H^o_-(a_i) = 0$, $i = 2,\ldots,p$), and then with a general distribution of internal sources. The continuous version of the present problem is considered later (see (44)-(66) of Sec. 8.5). A general solution of the internal-source problem is given in Ex. 3 of Sec. 3.9.

Example 4: Irradiances on Infinitely Many Infinite Parallel Planes

We now consider an infinite set A of infinite parallel planes, sharing common reflectance and transmittance functions. These functions are constant over each plane but may have arbitrary directional structure. This infinite set is depicted schematically in Fig. 3.7. Our purpose in exhibiting this example is to point up again the fact that the interaction principle carries one only as far as to permit a meticulous formulation of the problem. However, the variations in the incisiveness and pertinence of a given formulation is thereafter limited only by the ingenuity of the wielder of the principle.

As a case in point, let the set A be irradiated at its upper level only and in an amount $H^o_-(A)$. It is required to find $W_+(A)$, the resultant upward radiant emittance of the upper surface of the set A. A straightforward application of the interaction method to the subset A yields the equation:

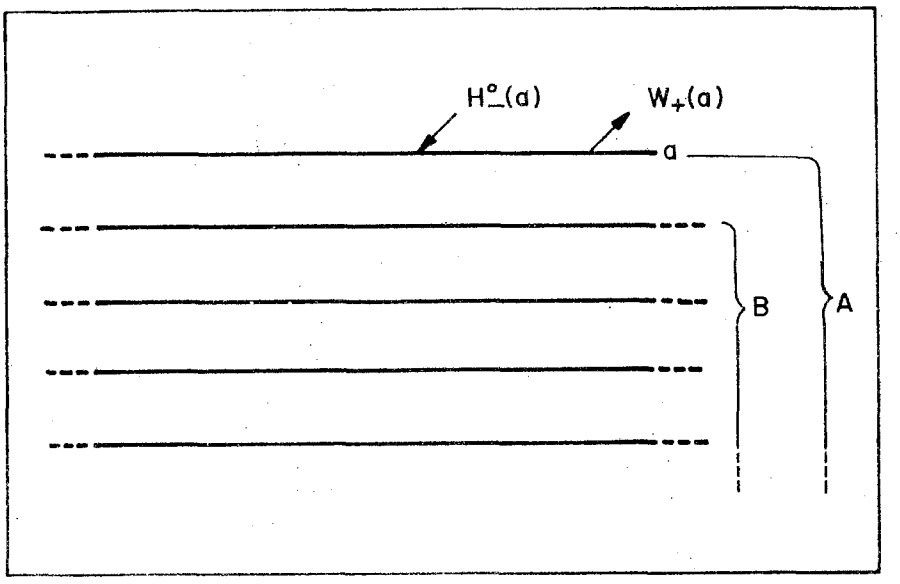


FIG. 3.7 An interaction of infinitely many planes, which illustrates Ambarzumian's principle.

$$W_+(A) = H_-^o(a)r_-(A) \quad . \tag{26}$$

Without any further insight into the problem on the part of a student of the subject, the interaction principle can carry the discussion no further. A significant advance toward the solution occurs, however, if the first plane a of the set A is removed from A and it is noticed that the remaining set B is in all particulars just like A except that it starts one layer lower than its predecessor. This may be viewed alternatively as if A were shoved bodily down one notch and otherwise left unchanged. Thus if a radiance distribution of the same directional structure impinges on B as that associated with $H_-^o(A)$, our hypothesized conditions imply that the reflectance $r_-(B)$ of B is equal to the reflectance $r_-(A)$ of A. Let us explicitly make this assumption to see where such an insight into the physics of the problem leads us.

The system A is now imagined to consist of two interacting subsets a and B. By invoking the interaction principle for each of these subsets, interaction equations are obtained which are formally identical to those for $H_+(a)$ and $W_+(b)$ (with "b" replaced by "B") in example 2. The only salient difference is that in the present case $W_+(B) = H_-^o(B) = H_+^o(B) = 0$. When these incident lighting conditions are put into (13) we obtain:

$$H_+(a) = H_-^o(a)t_-^o(a)r_-(B)/[1 - r_-(a)r_-(B)] \tag{27}$$

By (5) we have:

$$W_+(a) = H^o_-(a)r^o_-(a) + \frac{H^o_-(a)t^o_-(a)t_+(a)r_-(B)}{[1-r_+(a)r_-(B)]} \qquad (28)$$

From a simple operational argument (i.e., imagine taking instrument readings) it is clear that:

$$W_+(a) = W_+(A) \quad . \qquad (29)$$

Now, comparing (26) and (28), and using the physically-based insight:

$$r_-(A) = r_-(B) \qquad (30)$$

we have:

$$\boxed{r_-(A) = r^o_-(a) + \frac{t^o_-(a)t_+(a)r_-(A)}{1-r_+(a)r_-(A)}} \qquad (31)$$

Equation (31) governs the requisite reflectance $r_-(A)$ (for irradiance) of an infinite system A of parallel plane surfaces of common optical properties. In order to arrive at (31) two uses of physical intuition in the form of (29), (30) were needed. We assumed in particular that the structure of the downward incident radiance distribution on B was of the same angular structure as that on A. In the present adventurous spirit we can go on to assume that all radiance distributions within A have the same angular structure (say uniform) so that:

$$r^o_-(a) = r_+(a)$$

$$t^o_-(a) = t_+(a)$$

Let us write "r" for this common reflectance value for each plane in A and "t" for this common transmittance value for each plane in A. Then (31) reduces to:

$$r_-(A) = r + \frac{t^2 r_-(A)}{1 + rr_-(A)} \quad . \qquad (32)$$

This yields the following quadratic equation governing $r_-(A)$:

$$rr^2_-(A) + (1-r^2-t^2)r_-(A) - r = 0 \qquad (33)$$

which has the physically meaningful solution:

$$r_-(A) = \frac{-(1-r^2-t^2) + \sqrt{(1-r^2-t^2)^2 + 4r^2}}{2r} \quad . \qquad (34)$$

To help choose between the two root signs, observe that if $t = 0$, then $r_-(A)$ should be r. By using the sign "+", (34) yields this limiting answer. As an example of $r_-(A)$ for a

particular value of r, let r = 1/2 and t = 1/2, then:

$$r_{-}(A) = .618 \quad .$$

The lesson provided by this particular case of many interacting plane surfaces may be summarized as follows: while the interaction principle always yields the correct formulation corresponding to the wielder's method of analysis of a given radiometric system, there are nevertheless some choices of physical analyses of that system which are more pertinent than others. The methods of devising such analyses defy systematic description and retain radiative transfer theory, in this respect, in the ranks of the arts. The particular insight which was decisive in the present case was that summarized in (30), and is originally due to Ambarzumian. (See Ref. [1], [2].) This insight was the basis for the formulation of the first of the principles of invariance of modern radiative transfer theory (for a historical sketch, see Sec. 49 of Ref. [251]). Now that hindsight and formal principles (such as the interaction principle) are available, we can mechanically reproduce Ambarzumian's insight (30), (see, e.g., (30) of Sec. 7.3). Hence the arguments leading to (34) can be made formally rigorous without superfluous physical assumptions.

Example 5: The Algebra of Reflectance and Transmittance Operators for Planes

In this example we show how the reflectance and transmittance operators $r_{\pm}(Y)$, $t_{\pm}(Y)$, as given in (10), (11) of Sec. 3.3, are used in computations leading to reflected and transmitted radiance distributions on plane surfaces. Furthermore, some important technical concepts, such as the radiometric norm of an integral operator, will be developed, along with the rudiments of the algebra of reflectance and transmittance operators.

We consider first a plane surface a with unit upward normal **k**, as in Fig. 3.8. As usual, "Ξ_{+}" and "Ξ_{-}" will denote the set of all *upward* and *downward* directions with respect to **k**. Let "$N_{-}^{+}(x,\cdot)$" denote the downward (-) surface radiance distribution at x in a. Let "$N_{+}^{+}(x,\cdot)$" denote the upward surface radiance (+) distribution at x in a. In general throughout this example, a signed subscript on "N" will tell whether the radiance distribution is upward (+) or downward (-). A signed superscript on "N" will tell whether the radiance distribution is a surface (+) or field (-) radiance. This convention will serve to help us keep track of the comings and goings of radiant flux over the surface a.

For the moment we are interested in relating the downward field radiance distribution $N_{-}^{-}(x,\cdot)$ to its reflected and transmitted surface radiance distribution at x on surface a. Our goal is to establish the concept of the radiometric norm of radiance distributions and of the reflectance and transmittance operators. According to (10) and (13) of Sec. 3.3:

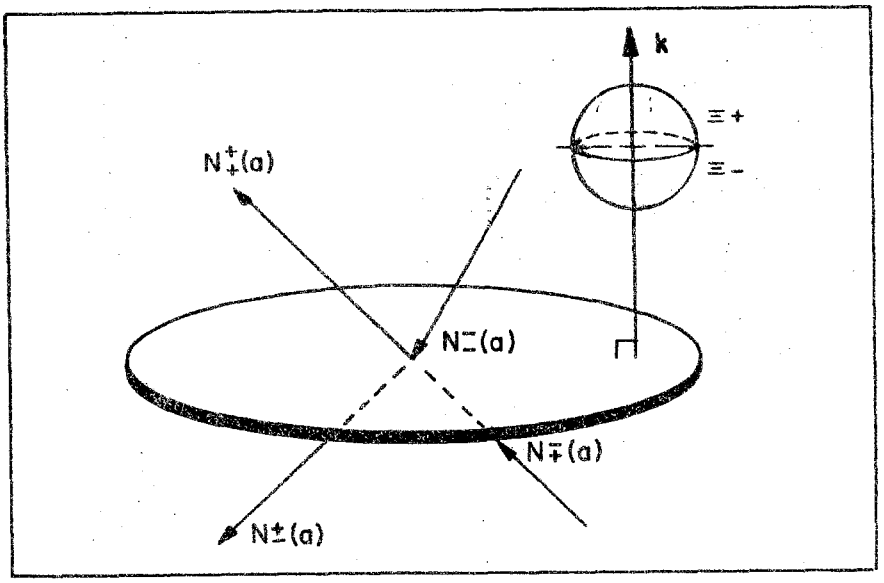


FIG. 3.8 Depicting the notation conventions for incident and response radiances of the field and surface type.

$$N^+_+(x,\xi) = \int_{\Xi_-} N^-_-(x,\xi')r_-(x;\xi';\xi)\, d\Omega(\xi') \quad , \tag{35}$$

and similarly, according to (11) of Sec. 3.3:

$$N^+_-(x,\xi) = \int_{\Xi_-} N^-_-(x,\xi')t_-(x;\xi';\xi)\, d\Omega(\xi') \quad . \tag{36}$$

Here $N^+_+(x,\cdot)$ is the *reflected* and $N^+_-(x,\cdot)$ the *transmitted* radiance distributions at x on a. Integral (35) shows us how to find the value of $N^+_+(x,\xi)$ for every x in a and ξ in Ξ_+, given the values $N^-_-(x,\xi)$ for every x in a and ξ in Ξ_-. Thus (35) leads to the operator equation:

$$N^+_+(a) = N^-_-(a)r_-(a) \quad , \tag{37}$$

where $r_-(a)$ is $r_-(Y)$ for $Y = a$, and $N^+_+(a)$ is the upward surface radiance distribution over a. Thus, to every x in a and ξ in Ξ_+, the value of $N^+_+(a)$ at (x,ξ) is $N^+_+(x,\xi)$ as given in (35). Similarly, $N^-_-(a)$ is the downward field radiance distribution over a. Equation (36) gives rise to the analogous operator equation:

$$N^+_-(a) = N^-_-(a)t_-(a) \tag{37a}$$

There are two more operator equations in addition to (37) and (37a), which round out the family of reflectance and transmittance operations over surface a:

$$N_+^+(a) = N_+^-(a)t_+(a) \tag{38}$$

$$N_-^+(a) = N_+^-(a)r_+(a) \tag{39}$$

which are the respective operator condensations of:

$$N_+^+(x,\xi) = \int_{\Xi_+} N_+^-(x,\xi')t_+(x;\xi';\xi)\, d\Omega(\xi') \tag{40}$$

and

$$N_-^+(x,\xi) = \int_{\Xi_+} N_+^-(x,\xi')r_+(x;\xi';\xi)\, d\Omega(\xi') \tag{41}$$

When working with the operators $r_\pm(a)$, $t_\pm(a)$ on one surface only, such as plane surface a, it is clear that the "+" and "-" superscripts on the radiance symbol are redundant (since field radiances are always operated on to yield surface radiances) and therefore may be dropped. However, when working with two or more interacting surfaces, the "+" and "-" superscripts often must be retained to avoid ambiguity during certain manipulations, as we shall see in later examples.

Radiometric Norm

We go on now to define the concept of a *radiometric norm* of radiance distributions over plane surfaces and of reflectance and transmittance operators. This concept is of central importance in both theoretical and practical computations of radiance distributions resulting from reflections or transmissions over surfaces. For every plane surface of finite area, let us write:

$$\text{"}|N_\pm(a)|\text{"} \quad \text{for} \quad \frac{1}{A(a)} \int_a \int_{\Xi_\pm} N(x,\xi)\, d\Omega(\xi)\, dA(x) \tag{42}$$

where "A(a)" denotes the area of surface a. If A(a) is infinite, then we write:

$$\text{"}|N_\pm(a)|\text{"} \quad \text{for} \quad \lim_{a' \to a} |N_\pm(a')| \tag{42a}$$

where a' is one of a family of subsets of a of finite area which equals a in the limit. We call $|N_\pm(a)|$ the *radiometric norm* of $N_\pm(a)$ over a. It is reminiscent of (and related to) the radiometric norm used in Example 15 of Sec. 2.11. The extension of (42) to curved surfaces is immediate. $N_\pm(a)$ in (42) can be either a field or surface radiance distribution for each direction (+) or (-). We observe that the radiometric

norm is linear in the sense that:

$$|cN_{\pm}(a) + c'N'_{\pm}(a)| = c|N_{\pm}(a)| + c'|N'_{\pm}(a)| \tag{43}$$

for every pair $N_{\pm}(a)$, $N'_{\pm}(a)$ of radiance distributions and nonnegative real numbers c and c'. Next, we write:

$$\text{"}\gamma_{\pm}(x;\xi')\text{"} \quad \text{for} \quad \int_{\Xi_{\pm}} r_{\pm}(x;\xi';\xi)\, d\Omega(\xi) \tag{44}$$

and

$$\text{"}\delta_{\pm}(x;\xi')\text{"} \quad \text{for} \quad \int_{\Xi_{\pm}} t_{\pm}(x;\xi';\xi)\, d\Omega(\xi) \tag{45}$$

For a given (field or surface) radiance distribution $N_{\pm}(a)$ with non zero radiometric norm over a, we write:

$$\text{"}\gamma_{\pm}(x,N)\text{"} \quad \text{for} \quad \frac{\int_{\Xi_{\pm}} N(x,\xi')\gamma_{\pm}(x,\xi')\, d\Omega(\xi')}{\int_{\Xi_{\pm}} N(x,\xi')\, d\Omega(\xi')} \tag{46}$$

and

$$\text{"}\delta_{\pm}(x,N)\text{"} \quad \text{for} \quad \frac{\int_{\Xi_{\pm}} N(x,\xi')\delta_{\pm}(x,\xi')\, d\Omega(\xi')}{\int_{\Xi_{\pm}} N(x,\xi')\, d\Omega(\xi')} \tag{47}$$

Finally, we write:

$$\text{"}\gamma_{\pm}(a,N)\text{"} \quad \text{for} \quad \frac{\int_a \gamma_{\pm}(x,N)\left[\int_{\Xi_{\pm}} N(x,\xi')\, d\Omega(\xi')\right] dA(x)}{\int_a \int_{\Xi_{\pm}} N(x,\xi')\, d\Omega(\xi')\, dA(x)} \tag{48}$$

$$\text{"}\delta_{\pm}(a,N)\text{"} \quad \text{for} \quad \frac{\int_a \delta_{\pm}(x,N)\left[\int_{\Xi_{\pm}} N(x,\xi')\, d\Omega(\xi')\right] dA(x)}{\int_a \int_{\Xi_{\pm}} N(x,\xi')\, d\Omega(\xi')\, dA(x)} \tag{49}$$

The motivation behind this seemingly bizarre set of definitions will now become clear. Let $N_-(a)$ be an incident downward radiance distribution on surface a, and let $N_+(a)$ be its reflected radiance distribution. Then by (35):

$$N_+(x,\xi) = \int_{\Xi_-} N_-(x,\xi')r_-(x;\xi';\xi)\, d\Omega(\xi') \quad .$$

We now wish to relate the radiometric norms of $N_+(a)$ and $N_-(a)$. The preceding definitions have been formulated with precisely this task in mind. We begin by integrating each side of this equation over Ξ_+. The result is:

$$\int_{\Xi_+} N_+(x,\xi)\, d\Omega(\xi) = \int_{\Xi_-} N_-(x,\xi')\, \gamma_-(x;\xi')\, d\Omega(\xi')$$

Then we integrate this result over a and divide by A(a). The new result is:

$$|N_+(a)| = \frac{1}{A(a)} \int_a \left[\int_{\Xi_-} N_-(x,\xi')\gamma_-(x;\xi')\, d\Omega(\xi') \right] dA(x)$$

by definition of the radiometric norm of $N_+(a)$. Next, by (46):

$$|N_+(a)| = \frac{1}{A(a)} \int_a \gamma_-(x,N_-) \left[\int_{\Xi_-} N_-(x,\xi')\, d\Omega(\xi') \right] dA(x)$$

and by (48):

$$|N_+(a)| = \frac{\gamma_-(a,N_-)}{A(a)} \int_a \int_{\Xi_-} N_-(x,\xi')\, d\Omega(\xi')\, dA(x)$$

$$= \gamma_-(a,N_-)|N_-(a)| \quad .$$

The last equality follows from the definition of the radiometric norm of $N_-(a)$. This is the desired relation between the norm of $N_-(a)$ and that of $N_+(a)$. This type of computation can be repeated three times. The collected results are as follows:

$$\left|N_{\pm}(a)\right| = \gamma_{\mp}\left(a,N_{\mp}\right)\left|N_{\mp}(a)\right| \qquad (50)$$

$$\left|N_{\pm}(a)\right| = \delta_{\pm}\left(a,N_{\pm}\right)\left|N_{\pm}(a)\right| \qquad (51)$$

In these equations the norms of incident radiance distributions are on the right, and the norms of response radiance distributions are on the left. To denote this fact

explicitly and to avoid possible ambiguities, superscripts "+" and "-" may be appropriately appended to "N". The numbers $\gamma_{\pm}(a,N)$, $\delta_{\pm}(a,N)$ are called the *special radiometric norms* of $r_{\pm}(a)$, $t_{\pm}(a)$ associated with N over a, respectively. These numbers are dependent on the incident radiance distributions as can be seen by inspection of (46)-(49). However, it is not difficult to show that, for every incident radiance distribution N over a and every plane surface a:

$$0 \leq \gamma_{\pm}(a,N) \leq 1 \tag{52}$$

$$0 \leq \delta_{\pm}(a,N) \leq 1 \quad . \tag{53}$$

These inequalities follow from the energy conservation principle as enunciated in Sec. 3.1. From this we conclude at once that the radiometric norms of response functions cannot exceed those of the incident functions. The proofs of (52) and (53) first bring out the facts that $0 \leq \gamma_{\pm}(x;\xi') \leq 1$ and $0 \leq \delta_{\pm}(x;\xi') \leq 1$ for every x in a and every ξ' in $\Xi_{\pm}$. An empirical form of these inequalities was established in Sec. 3.1 directly from the energy conservation principle and may be used as a basis for the present proofs. From these latter inequalities and (46) we see that $0 \leq \gamma_{\pm}(x,N) \leq 1$ and $0 \leq \delta_{\pm}(x,N) \leq 1$ for every x in a and every N. Inequalities (52) and (53) now follow for every a and N from (48) and (49). Finally, we can free the numbers $\gamma_{\pm}(a,N)$ and $\delta_{\pm}(a,N)$ from dependence on the incident radiance distributions N by taking their maxima (or suprema) as N varies over all possible incident radiance distributions in a. Thus let us write:

$$\text{"}\gamma_{\pm}(a)\text{"} \quad \text{for} \quad \max_N \gamma_{\pm}(a,N) \tag{54}$$

$$\text{"}\delta_{\pm}(a)\text{"} \quad \text{for} \quad \max_N \delta_{\pm}(a,N) \quad . \tag{55}$$

From this and (50)-(53) we have:

$$\left|N_{\pm}(a)\right| \leq \gamma_{\mp}(a)\left|N_{\mp}(a)\right| \tag{56}$$

$$\left|N_{\pm}(a)\right| \leq \delta_{\mp}(a)\left|N_{\pm}(a)\right| \tag{57}$$

$$0 \leq \gamma_{\mp}(a) \leq 1 \tag{58}$$

$$0 \leq \delta_{\mp}(a) \leq 1 \tag{59}$$

The response radiance distributions are on the left in (56), (57), and the incident radiance distributions are on the right. We call $\gamma_{\pm}(a)$ and $\delta_{\pm}(a)$ the *general radiometric norms* of $r_{\pm}(a)$ and $t_{\pm}(a)$ for surface a. The properties (56) and (57) will play an important role in the discussion of the existence of solutions of the radiative transfer formulations below. In particular the following properties will turn out to be sufficient conditions for the existence of many

solutions. We shall say that $r_{\pm}(a)$ and $t_{\pm}(a)$ are *norm contracting* if and only if:

$$0 < \gamma_{\pm}(a) < 1 \tag{60}$$

$$0 < \delta_{\pm}(a) < 1 \tag{61}$$

Iterated Operators

We turn next to a systematic description of certain integration details arising in interreflection calculations. Familiarity with these details will help the reader to attain a working understanding of how to translate into numerical form the results of the algebraic manipulations of the reflectance and transmittance operators.

We have seen in the preceding examples how the four operators $r_{\pm}(a)$, $t_{\pm}(a)$ associated with a surface a serve to describe the reflection or transmission of incident radiance distributions on a. These response radiance distributions can subsequently go on to interact with another surface b, or conversely, the incident radiance distributions on a may have come to a after being reflected or transmitted in some surface b. Hence there arises the possibility of considering an operation like $r_+(a)$ followed by an operation like $r_-(b)$; or $t_-(a)$ followed by $t_-(b)$; or $r_-(b)$ by $t_+(a)$, and so on. These combined operations are called *iterations*. We now systematically consider all such iteration possibilities of two operators when a and b are two parallel plane surfaces. There are eight such possibilities. They are schematically depicted in parts (a)-(d) of Fig. 3.9. Actually, the four possibilities in (a) and (b) of Fig. 3.9 are exhaustive of the basic possible combinations. Turn them upside down to get types in parts (c) and (d) of the figure. However, in a terrestrially-based coordinate system, generally one that is fixed independent of the planes a and b, it is convenient to explicitly and independently list also the possibilities (c) and (d) of Fig. 3.9.

As a specific example of the iteration of the interaction operators $r_{\pm}(a)$, $t_{\pm}(a)$, let us consider the iteration of $r_+(a)$ and $r_-(b)$ in that order. That is, an upward field radiance distribution $N_+^-(a)$ is incident on plane a and is reflected in plane a. The reflected flux goes on to become an incident radiance distribution $N_-^-(b)$ on b and which is reflected in turn in plane b. This is shown in part (b) of Fig. 3.9. If $N_+^+(b)$ is the resultant upward surface radiance distribution over b, then the value of $N_+^+(b)$ at y_j on b in the direction ξ_j in Ξ_+ is given by:

$$N_+^+(y_j,\xi_j) = \int_{\Xi_-} N_-^-(y_j;\xi_j')\, r_-(y_j;\xi_j';\xi_j)\, d\Omega(\xi_j') \tag{62}$$

Now by the radiance invariance law and with the help of Fig. 3.10:

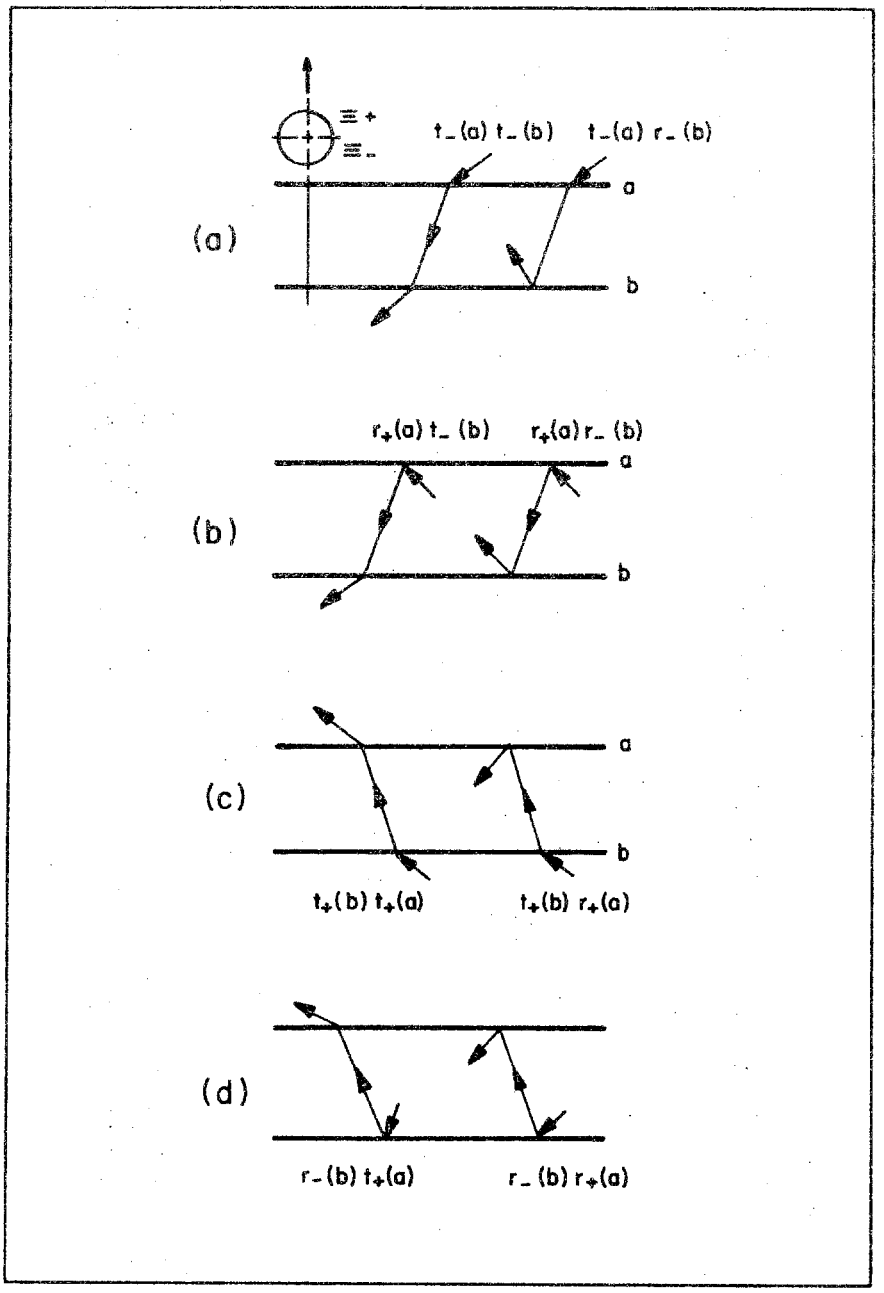


FIG. 3.9 Bookkeeping procedure in operator iterations: the eight possible types of interactions between two planes.

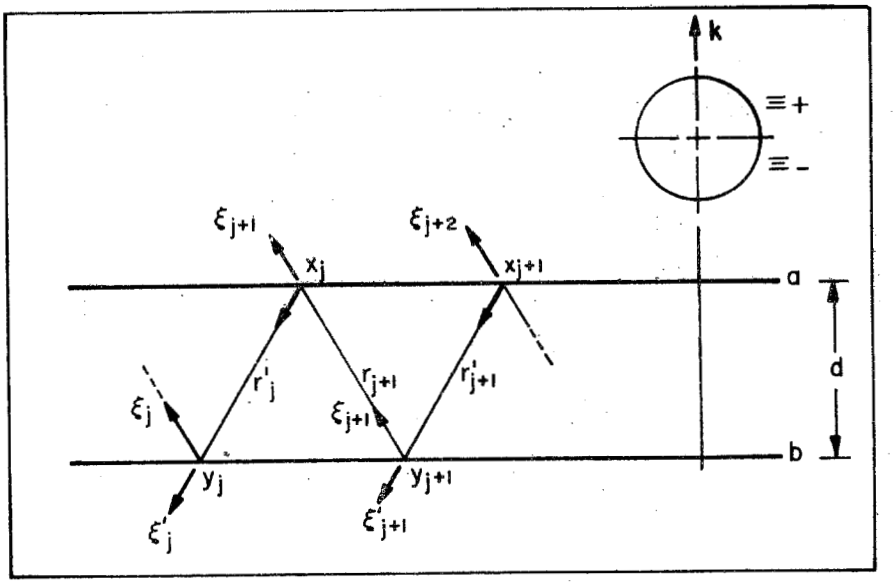


FIG. 3.10 Further detail of Diagram (b) in Fig. 3.9.

$$N_-^-(y_j,\xi'_j) = N_-^+(x_j,\xi'_j) \tag{63}$$

where:

$$y_j = x_j + r'_j\xi'_j$$

for every ξ'_j in Ξ_- at every y_j in b. But the value $N_-^+(x_j,\xi'_j)$ is the value of $N_-^+(a)$ at x_j and for ξ'_j, and this, by our present agreement, is given by:

$$N_-^+(x_j,\xi'_j) = \int_{\Xi_+} N_+^-(x_j,\xi_{j+1})r_+(x_j;\xi_{j+1};\xi'_j)\,d\Omega(\xi_{j+1}) \tag{64}$$

Combining (62), (63) and (64) we have:

$$N_+^+(y_j,\xi_j) =$$

$$= \int_{\Xi_-}\left[\int_{\Xi_+} N_+^-(x_j,\xi_{j+1})r_+(x_j;\xi_{j+1};\xi'_j)\,d\Omega(\xi_{j+1})\right] r_-(y_j;\xi'_j;\xi_j)\,d\Omega(\xi'_j) \tag{65}$$

which holds for every y_j in b and ξ_j in Ξ_+. In brief:

$$N_+^+(b) = N_+^-(a)r_+(a)r_-(b) \quad , \tag{66}$$

where we have written the left and right sides of (66) as abbreviations of the corresponding sides of (65). The variables x_j, y_j, ξ_j, ξ'_j, etc., are appropriately subscripted by integers j with $j = 0,1,2,\ldots$ simply in order to help the eye keep track of manipulations of variables in iterations of more than two operators.

To summarize, we have agreed to write:

"$r_+(a)r_-(b)$" for

$$\int_{\Xi_-}\left[\int_{\Xi_+}[\ ]r_+(x_j;\xi_{j+1};\xi'_j)\,d\Omega(\xi_{j+1})\right]r_-(y_j;\xi'_j;\xi_j)\,d\Omega(\xi'_j) \tag{67}$$

where x_j is in a, and $y_j = x_j + r_j\xi'_j$ and r'_j is the distance between x_j and y_j, i.e.,

$$r'_j = d/|\xi'_j \cdot k| \tag{68}$$

and where d is the distance between a and b, and k is the unit upward normal to a and b. Equation (66) may be conveniently interpreted as follows: if $N_+^-(a)$ is a member of a set $A_1(a)$ of incident functions on a and $B_1(a)$ is a set of response functions containing $N_-^+(a)$, we view $r_+(a)$ as an operator which maps elements of $A_1(a)$ into $B_1(a)$. Similarly, $r_-(b)$ maps elements of a set $A_1(b)$ of incident functions on b into a set $B_1(b)$ of response functions on b. Therefore, the iterated operator $r_+(a)r_-(b)$ maps elements of $A_1(a)$ into $B_1(b)$.

As another example suppose $N_-^+(b)$ is the transmitted radiance distribution of the surface b in response to $N_-^-(b)$. If the space between surfaces a and b is a vacuum, then by the radiance invariance law, $N_-^-(b)$ is equal to $N_-^+(a)$, which in turn is the transmitted response to $N_-^-(a)$. See Fig. 3.11, which is a version of part (a) of Fig. 3.9. Then for every y_j in b and ξ_j in Ξ_-, we have:

$$N_-^+(y_j,\xi_j) = \int_{\Xi_-} N_-^-(y_j,\xi'_j)t_-(y_j;\xi'_j;\xi_j)\,d\Omega(\xi'_j) \tag{69}$$

By the radiance invariance law:

$$N_-^-(y_j,\xi'_j) = N_-^+(x_j,\xi'_j) \quad ,$$

where:

$$y_j = x_j + r'_j\xi'_j$$

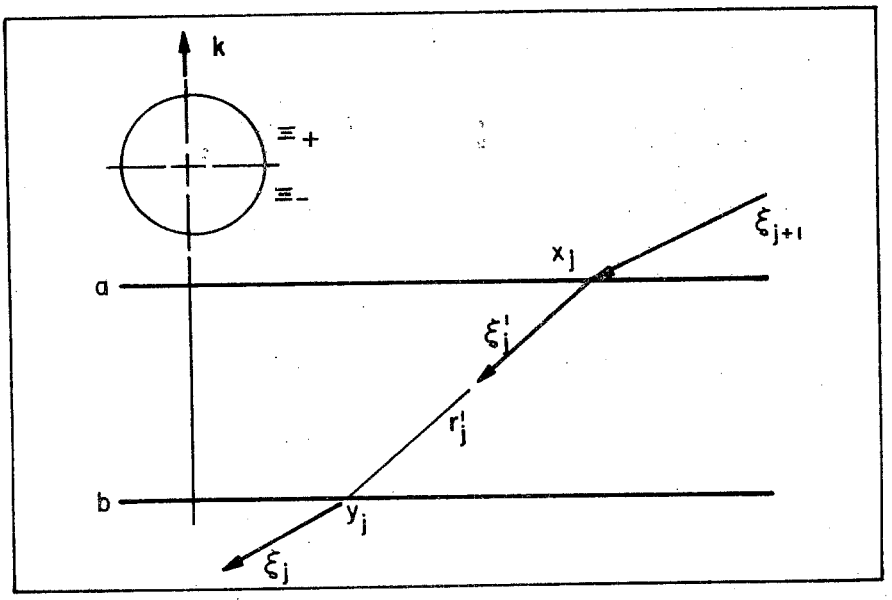


FIG. 3.11 Further detail of Diagram (a) in Fig. 3.9.

for every ξ'_j in Ξ_- at every y_j in b. But $N_-^+(x_j,\xi'_j)$ is the value of $N_-^+(a)$ at x_j and for ξ'_j and this, by our present agreement, is given by:

$$N_-^+(x_j,\xi'_j) = \int_{\Xi_-} N_-^-(x_j,\xi_{j+1})t_-(x_j;\xi_{j+1};\xi'_j)\, d\Omega(\xi_{j+1}) \quad (70)$$

Combining (69) and (70) we have:

$$N_-^+(y_j,\xi_j) =$$

$$\int_{\Xi_-}\left[\int_{\Xi_-} N_-^-(x_j;\xi_{j+1};\xi'_j)t_-(x_j;\xi_{j+1};\xi'_j)d\Omega(\xi_{j+1})\right]t_-(y_j;\xi'_j;\xi_j)d\Omega(\xi'_j) \quad (71)$$

which holds for every y_j in b and ξ_j in Ξ_-. In brief:

$$N_-^+(b) = N_-^-(a)t_-(a)t_-(b) \quad (72)$$

where we have written the left and right sides of (72) as abbreviations of the corresponding sides of (71). Thus we agree to write:

"$t_-(a)t_-(b)$" for

$$\int_{\Xi_-}\left[\int_{\Xi_-}[\ ]t_-(x_j;\xi_{j+1};\xi'_j)\,d\Omega(\xi_{j+1})\right]t_-(y_j;\xi'_j;\xi_j)\,d\Omega(\xi'_j) \tag{73}$$

It is clear that the iteration $t_-(a)t_-(b)$ maps elements in an incident set of surface a into the response set of surface b. Similar definitions can now be made of the remaining six types of iterations depicted in (a)-(d) of Fig. 3.9. These are left to the reader as exercises.

The radiometric norms of response functions under doubly iterated operators are related to the radiometric norms of the original incident function as follows. From (66) to which we apply (56) twice:

$$\begin{aligned}|N_+(b)| &= |N_+(a)r_+(a)r_-(b)| \\ &\le \gamma_-(b)|N_+(a)r_+(a)| \\ &\le \gamma_+(a)\gamma_-(b)|N_+(a)|\ .\end{aligned}$$

Similarly from (72) to which we apply (57) twice:

$$|N_-(b)| \le \delta_-(a)\delta_-(b)|N_-(a)|\ .$$

Similar inequalities hold for the remaining six iteration possibilities.

Operator Algebras and Radiative Transfer

We close this discussion of Example 5 by observing that iterations of the operators $r_+(a)$ and $t_\pm(a)$ can be continued indefinitely as long as each iteration is meaningful. As an example of a meaningful continued iteration consider iterating $r_+(a)r_-(b)$ with itself. Thus let us write:

$$"(r_+(a)r_-(b))^1" \quad \text{for} \quad r_+(a)r_-(b) \tag{74}$$

and for every $j \ge 1$ we write:

$$"(r_+(a)r_-(b))^{j+1}" \quad \text{for} \quad (r_+(a)r_-(b))^j(r_+(a)r_-(b)). \tag{75}$$

Thus $(r_+(a)r_-(b))^j$ is an iterated integral, iterated 2j times using the reflectance functions $r_+(a)$ and $r_-(b)$. An example of a meaningless iteration is $t_+(a)r_-(b)$, since after a transmits flux upward, the flux does not go directly to b for reflection. Finally, we note that interaction operators can be added together. Thus "$r_-(b) + r_-(b)r_+(a)r_-(b)$" denotes the operator which acts on downward radiance distributions like

$N^{-}(b)$ to give upward radiance distributions like $N^{-}(b)r_{-}(b) + N^{-}(b)r_{-}(b)r_{+}(a)r_{-}(b)$, etc. Hence we can speak of *sums* and *products* (i.e., iterations) of interaction operators and so handle the operators in many instances as if they were numbers, and as if they obeyed the same laws as numbers.

The properties of iterated and added operators arising in radiative transfer theory, when all assembled in a systematic list, turn out to be rather interesting from an algebraic point of view. Suppose that r, s, and t are any of the interaction operators defined above whose iterations rs, rt, st or sums r+s, r+t, s+t, are meaningful, then we can readily establish the following properties under suitable regularity conditions on the *reflectance*, or *transmittance* operators underlying the symbols. Let $\mathcal{a}$ be the set of all such operators. Then to every pair r,s of operators in $\mathcal{a}$ there corresponds a sum r+s in $\mathcal{a}$ of r and s such that

(i)	$r + s = s + r$	(commutativity)
(ii)	$r + (s+t) = (r+s) + t$	(associativity)
(iii)	$0 + r = r$	(identity)
(iv)	$r + (-r) = 0$	(inverse)

Furthermore, to every real number α an operator r in $\mathcal{a}$ there corresponds an operator αr, the *product* of α and r. (For example, if r is a reflectance operator, then αr is just the integral operator formed in the usual manner after multiplying $r_{\pm}(\cdot;\xi';\xi)$ in (10) of Sec. 3.3 by α.) The product αr has the general properties:

(v)	$\alpha(r+s) = \alpha r + \alpha s$	(operator distributivity)
(vi)	$(\alpha+\beta)r = \alpha r + \beta r$	(scalar distributivity)
(vii)	$(\alpha\beta)r = \alpha(\beta r)$	(scalar associativity)
(viii)	$0r = 0,\ 1r = r$	(zero, identity)

In the preceding equations α,β are real numbers, and 0,1 on the left sides in (viii) are the usual zero and unit real numbers. The 0 on the right in (viii) is the zero integral operator obtained by using $r_{+} = 0$ or $r_{-} = 0$.

If the set $\mathcal{a}$ of operators satisfies all eight of the preceding conditions, then $\mathcal{a}$ is a *vector space* of operators. Now a vector space is a most useful object to work with because of its highly intuitive concepts and because of the rich body of computational theorems that exists for it. This gives one incentive to arrange matters so that, in suitably enlarged domains, the operator sums r+s and products rs are meaningful for every r, s in $\mathcal{a}$. We shall not pursue such matters in the present work, as it is too potentially vast a subject to compress into one section of a study devoted primarily to the immediate mathematical-physical foundations of the subject. However, before leaving this area of ideas,

several more interesting algebraic aspects of the reflectance and transmittance operators will be brought out for future reference.

As observed above, we can assign meaning to the product rs of two reflectance or transmittance operators. Consider once again the collection $\mathcal{A}$ of all such operators associated with some optical medium. Then we have, in addition to properties (i)-(viii), the following:

$$\text{(ix)} \quad r(st) = (rs)t$$

$$\text{(x)} \quad r(s+t) = rs + rt$$

$$\text{(xi)} \quad (r+s)t = rt + st$$

When $\mathcal{A}$ satisfies (i)-(xi) we call it an *operator ring*. If, moreover, we have

$$\text{(xii)} \quad \alpha(rs) = (\alpha r)s = r(\alpha s)$$

for every real number α and pair of operators r, s, in $\mathcal{A}$, then $\mathcal{A}$ is an *operator algebra*. All of these twelve properties are readily shown to hold for the surface reflectance and transmittance operators of the form $r_{\pm}(a)$, $t_{\pm}(a)$, provided their products are definable. Unfortunately the commutative property, rs = sr, does not generally hold (except when certain reciprocity conditions are in force).

Going still further, let us write "$||r||$" for the *norm* of r, obtained as shown in (54), (55), depending on the nature of r. Then it is possible to show that, under suitable conditions:

$$\text{(xiii)} \quad 0 \leq ||r||, \text{ and } ||r|| = 0 \text{ if and only if } r = 0$$

$$\text{(xiv)} \quad ||r+s|| \leq ||r|| + ||s||$$

$$\text{(xv)} \quad ||\alpha s|| = |\alpha| \, ||s||$$

$$\text{(xvi)} \quad ||rs|| \leq ||r|| \, ||s||$$

$$\text{(xvii)} \quad \text{If } \mathcal{A} \text{ has an identity element I, then } ||I|| = 1$$

The first of these, namely (xiii), was essentially proved in (58), (59). Condition (xiv) is the triangle inequality, condition (xv) is trivial, and (xvi) is readily established using definitions of the kind (67), (73). The identity element I of $\mathcal{A}$ can be a suitably defined Dirac delta function. A set $\mathcal{A}$ of operators satisfying (i)-(xvii) is called a *normed operator algebra*. Unfortunately, for each r we do not generally have an s such that rs = I, so that operator division and hence the finding of inverses is generally not possible (however transmittance operators generally have inverses).

It appears that all of classical radiative transfer theory for surfaces and solids can be mathematically cast in terms of the theory of normed operator algebras. While this

The radiance distribution $N_+(b)$ is therefore of uniform directional structure and of magnitude:

$$N_+(b) = \frac{t_-r_-}{\pi} H_-^o(a) \sum_{j=0}^{\infty} (r_+r_-)^j = \frac{H_-^o(a)}{\pi} \frac{t_-r_-}{[1 - r_+r_-]} \qquad (86)$$

From (82):

$$\begin{aligned} N_-(a) &= N_-^o(a)t_-^o(a) + N_+(b)r_+(a) \\ &= \frac{t_-}{\pi} H_-^o(a) + \frac{1}{\pi} \frac{r_+r_-t_-H_-^o(a)}{[1 - r_+r_-]} \\ &= \frac{H_-^o(a)}{\pi} \left[t_- + \frac{r_+r_-t_-}{(1 - r_+r_-)} \right] \\ &= \frac{H_-^o(a)}{\pi} \frac{t_-}{[1 - r_+r_-]} \\ &= \frac{N_+(b)}{r_-} \end{aligned} \qquad (87)$$

This completes the general calculation of $N_+(b)$ and $N_-(a)$. By adopting sufficiently severe assumptions, the calculation has been reduced to a finite number of arithmetic operations. Nevertheless, these results constitute useful rules of thumb for estimating the order of magnitudes of interreflected radiance distributions between two extensive parallel surfaces.

As a specific example of the use of (86) and (87), let $r_+ = r_- = t_- = 1/2$. Hence

$$N_+(b) = \frac{1}{\pi} \frac{(1/4)H_-^o(a)}{1 - (1/4)} = \frac{1}{3\pi} H_-^o(a)$$

$$N_-(a) = \frac{N_+(b)}{r_-} = \frac{2}{3\pi} H_-^o(a)$$

Truncation Error Estimates

The matter of truncating the series expansion of $[I - r_+(a)r_-(b)]^{-1}$ and estimating the resultant truncation error will be taken up next. The technique we use for the present special case is indicative of what can be done in the general truncation processes.

Suppose that planes a and b in Fig. 3.12 have arbitrary reflectance and transmittance operators subject only to the condition that at least one of $r_+(a)$ or $r_-(b)$ is norm contracting. For simplicity in manipulation let us suppose that

$N_+^o(b) = 0$, and $N_+^o(a) = 0$, with $N_-^o(a)$ arbitrary. Then from (84):

$$N_+(b) = N_-^o(a)t_-^o(a)r_-(b)[1 - r_+(a)r_-(b)]^{-1}$$

$$= N_-^o(a)t_-^o(a)r_-(b)\sum_{j=0}^{\infty}(r_+(a)r_-(b))^j \quad .$$

Let us truncate the infinite series at term k and denote by "$N_+(b;k)$" the associated radiance distribution. Thus:

$$N_+(b;k) = N_-^o(a)t_-^o(a)r_-(b)\sum_{j=0}^{k}(r_+(a)r_-(b))^j \quad .$$

From this, we have

$$N_+(b) - N_+(b;k) = N_-^o(a)t_-^o(a)r_-(b)\sum_{j=k+1}^{\infty}(r_+(a)r_-(b))^j \quad .$$

We are interested in the radiometric norm of the difference $N_+(b) - N_+(b;k)$, for this difference is evidently the component of the actual radiance distribution omitted by the truncation process. To facilitate the estimate of the norm, let us write *ad hoc*:

$$\text{"}N^o\text{"} \quad \text{for} \quad N_-^o(a)t_-^o(a)r_-(b) \quad .$$

Then

$$N_+(b) - N_+(b;k) = \left|N^o\sum_{j=k+1}^{\infty}(r_+(a)r_-(b))^j\right|$$

$$= \left|\sum_{j=k+1}^{\infty}N^o(r_+(a)r_-(b))^j\right|$$

$$= \sum_{j=k+1}^{\infty}\left|N^o(r_+(a)r_-(b))^j\right|$$

The last equation follows by repeated use of the linearity of the radiometric norm (42) (with $c = c' = 1$). The next step will be facilitated by examining a typical term of the preceding sum of norms. Thus observe that:

$$\left|N^{o}(r_{+}(a)r_{-}(b))^{j}\right| = \left|[N^{o}(r_{+}(a)r_{-}(b))^{j-1}](r_{+}(a)r_{-}(b))\right|$$

$$= \left|[N^{o}(r_{+}(a)r_{-}(b))^{j-1}r_{+}(a)]r_{-}(b)\right|$$

$$\leq \gamma_{-}(b)\left|N^{o}(r_{+}(a)r_{-}(b))^{j-1}r_{+}(a)\right|$$

$$\leq \gamma_{+}(a)\gamma_{-}(b)\left|N^{o}(r_{+}(a)r_{-}(b))^{j-1}\right|$$

The last two inequalities follow from the definition of the general radiometric norms given in Example 6. The pattern of reductions arising in the estimate of $|N_{+}(b) - N_{+}(b;k)|$ is now clear. We have:

$$|N_{+}(b) - N_{+}(b;k)| \leq \sum_{j=k+1}^{\infty} (\gamma_{+}(a)\gamma_{-}(b))^{j}|N^{o}|$$

$$= \frac{|N^{o}|[\gamma_{+}(a)\gamma_{-}(b)]^{k+1}}{[1 - \gamma_{+}(a)\gamma_{-}(b)]} .$$

In addition:

$$|N^{o}| \leq |N_{-}^{o}(a)|\delta_{-}(a)\gamma_{-}(b) .$$

Since $|N_{-}^{o}(a)| < |N_{+}(b)|$ in many reflection problems*, we find:

$$\frac{|N_{+}(b) - N_{+}(b;k)|}{|N_{+}(b)|} < \frac{\delta_{-}(a)\gamma_{-}(b)(\gamma_{+}(a)\gamma_{-}(b))^{k+1}}{[1 - \gamma_{+}(a)\gamma_{-}(b)]} . \quad (88)$$

This gives an estimate of the relative error arising from truncation of the natural mode of solution of $N_{+}(b)$ at the kth term.

As an example of the use of (88) suppose $\gamma_{+}(a) = \gamma_{-}(b) = 1/2$ and $\delta_{-}(a) = 1/2$. Then:

$$\frac{|N_{+}(b) - N_{+}(b;k)|}{|N_{+}(b)|} < \frac{4}{3}\left(\frac{1}{4}\right)^{k+2} = \frac{1}{3}\left(\frac{1}{4}\right)^{k+1}$$

By terminating the series of iterated integral operators at $k = 2$, we have

*Otherwise, simply divide through by $|N_{-}^{o}(a)|$ instead of $|N_{+}(b)|$.

$$\frac{|N_+(b) - N_+(b;2)|}{|N_+(b)|} < \frac{1}{3}\left(\frac{1}{4}\right)^3 = \frac{1}{192} = .0052 \quad .$$

Thus the resultant relative error of truncation in this case is on the order of 1/2 percent.

By using the calculation in Example 6 for the response radiances of Lambert surfaces a and b with r_+, r_- and t_- in that example now replaced by $\gamma_+(a)$, $\gamma_-(b)$, $\gamma_-(a)$, an estimate of $|N_+(b)|$ can be made. From this and the preceding inequality an estimate of the absolute error of truncation can be made. Thus, as an illustration, we use the result (84) of Example 6 to find:

$$|N_+(b)| = \frac{1}{3\pi} H_-^o(a) \quad .$$

Hence:

$$|N_+(b) - N_+(b;k)| < 5.5 \times 10^{-4} H_-^o(a) \quad .$$

It follows that if $H_-^o(a)$ is on the order of 10^3 watts/m^2 then:

$$|N_+(b) - N_+(b;k)| < .55 \text{ watts}/(\text{m}^2 \times \text{steradian}) \quad .$$

This estimate of $H_-^o(a)$ is a reasonable one for natural light fields as may be seen by an inspection of Table I of Sec. 2.4.

One can occasionally profitably reverse the preceding error estimate calculation as follows. We agree on an error $\varepsilon > 0$ at the outset and then solve for the k which will yield that ε. Thus, from (88), we set:

$$\frac{\delta_-(a)\gamma_-(b)\left(\gamma_+(a)\gamma_-(b)\right)^{k+1}}{\left[1 - \gamma_+(a)\gamma_-(b)\right]} = \varepsilon$$

whence we find k by means of the relation:

$$k+1 = \frac{\ln\left[\dfrac{\varepsilon\left(1 - \gamma_+(a)\gamma_-(b)\right)}{\delta_-(a)\delta_-(b)}\right]}{\ln\left(\gamma_+(a)\gamma_-(b)\right)}$$

This formula is associated with the particular geometric arrangements of the present example. It is a relatively simple matter to extend this result to other formulas in connection with related problems, one of which will be discussed next.

Quantum-Terminable Calculations

In closing Example 7 we remarked that the preceding method of determining the value of k, which goes with a particular ε, may be extended to certain interesting extreme cases. For example, suppose that the average number n of photons of a given frequency ν incident per second per unit area per unit solid angle on a surface falls below some number n_0, say $n_0 = 10^{-2}$, or $n_0 = 10^{-3}$, etc. Suppose that this magnitude of n_0 is so small that it is operationally meaningless to theorize about or experiment with the radiance N_0 produced by n_0. That is, N_0 is not measurable using available radiance meters because it is below their threshold of sensitivity. Suppose $\varepsilon = N_0/N$, where N is some fiducial magnitude for radiance--say that of the order of magnitude of the sun's maximum spectral radiance. This value of ε will then determine a corresponding finite value of k, say $k(\varepsilon)$, after the manner illustrated for the special case above. This value $k(\varepsilon)$ in turn can evidently be used in defining a terminable response radiance calculation. For example:

$$N_+(b;k(\varepsilon)) = N_-^o(a)t_-^o(a)r_-(b)\sum_{j=0}^{k(\varepsilon)}(r_+(a)r_-(b))^j$$

would define a terminable calculation for $N_+(b;k(\varepsilon))$. This would in turn give rise to a terminable calculation for $N_-(a;k(\varepsilon))$.

Such calculations, which are terminable by introducing quantum concepts in the way just indicated, are called *quantum-terminable calculations* and provide a basis for a strong physical argument in favor of the study of terminable calculations in radiative transfer theory. Terminations therefore need not be arbitrary; but can be based on real physical limitations of the apparatus on which rest the phenomenological foundations of the discipline. A systematic study of quantum-terminable calculations appears to hold certain interesting theoretical challenges (for example, can a consistent finite algebra of operators be developed on the basis of quantum-terminable calculations?). This study, however, is beyond the scope of the present work and is left for the interested reader to pursue.

Example 8: Two Interacting Finite Plane Surfaces

In the present example we return to the setting of the preliminary example in Sec. 3.1 and reformulate the problem of that section using now theoretical radiances and the method of the interaction principle. Fig. 3.13 reconstructs the essential features of the setting of Fig. 3.2 in anticipation of the use of the appropriate forms of the integral operators $r_\pm(Y)$ and $t_\pm(Y)$. The unit *outward* normals $\mathbf{k}_1$ and $\mathbf{k}_2$ for the two plane surfaces S_1 and S_2 fix the outward $\Xi_+(S_i)$ and

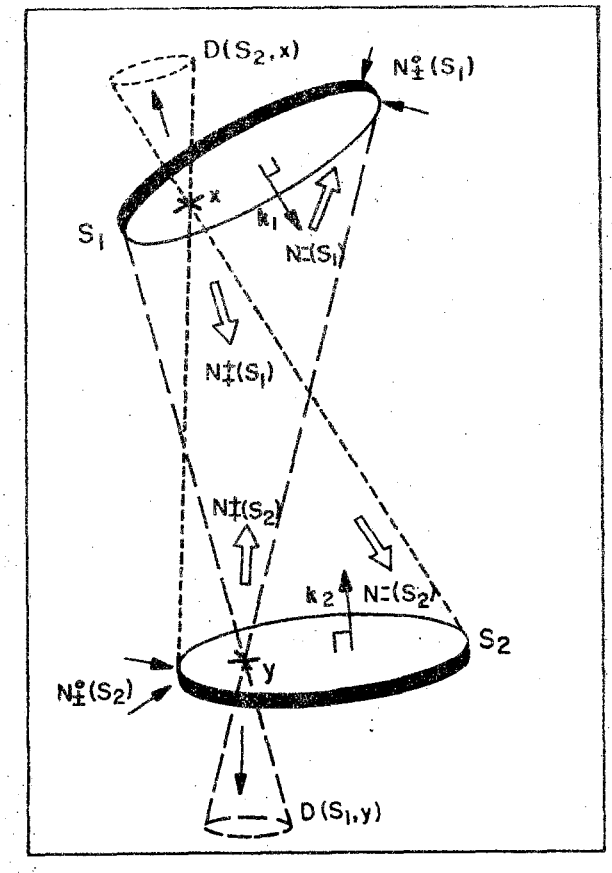


FIG. 3.13 Two interacting finite plane surfaces.

inward $\Xi_-(S_i)$ hemispheres on S_i, i = 1 or 2. Thus $\Xi_+(S_i)$ consists of all unit vectors ξ such that $\xi\cdot\mathbf{k}_i > 0$ and $\Xi_-(S_i)$ consists of all unit vectors ξ such that $\xi\cdot\mathbf{k}_i < 0$. This convention of fixing outward and inward hemispheres of *interacting* surfaces is to be distinguished from the corresponding convention for *collecting* surfaces used in Sec. 2.4. For collecting surfaces it is sometimes more convenient to refer the directions of incident flux to a unit *inward* normal. For a surface which is explicitly considered to interact with another, the outward unit normal is occasionally a more convenient reference direction to use. We do not intend, however, to permanently fix such conventions. Rather we shall choose between the conventions as a given situation favors one or the other. With the direction coordinate frames anchored to S_1 and S_2 in the above manner we now require that for every y in S_2 the set $D(S_1,y)$ of all directions from points of S_1 to y to lie in $\Xi_-(S_2)$. Conversely, we require for every x in S_1

the set $D(S_2,x)$ of all directions from points of S_2 to x to lie in $\Xi_-(S_1)$. See Fig. 3.13. These conditions amount to the simple requirement that each surface lie above the other's horizon. This is not an essential restriction; it serves only to shorten the number of cases considered in the analysis below. We require that the reflectance and transmittance functions of S_1 and S_2 be known and that the space between S_1 and S_2 be a vacuum. The two surfaces are irradiated by incident radiance distributions $N_\pm^o(S_i)$, i = 1 or 2. It is then required to formulate and solve the interreflection problem associated with S_1 and S_2 under these conditions. In particular we require the response (surface) radiance distributions $N_\pm(S_i)$ of S_i, i = 1 or 2. Thus $N_+(S_1)$, e.g., is a function which assigns to each x in S_1 and ξ in $\Xi_+(S_1)$ a surface radiance $N_+(x,\xi)$. As usual when the need arises to distinguish between field and surface radiances for S_1, the appropriate superscripts "-" or "+" respectively, will be appended to "N". Thus "$N_\pm^-(S_i)$" will denote field (incident) radiance distributions over S_i, i = 1 or 2, and "$N_\pm^+(S_i)$" will denote the surface (response) radiance distributions of S_i.

We isolate surface S_1 and enumerate the sets of incident radiometric functions on S_1:

A_1: All field radiance distributions like $N_-^o(S_1)$

A_2: All field radiance distributions like $N_+^o(S_1)$

A_3: All field radiance distributions like $N_-^-(S_1)$

The set of response radiometric functions for S_1 are:

B_1: All surface radiance distributions like $N_+^+(S_1)$

B_2: All surface radiance distributions like $N_-^+(S_1)$

Thus, in the case of surface S_1, m = 3, n = 2, and the six interaction operators s_{ij} supplied by the interaction principle are in the form of reflectance and transmittance integral operators as follows:

$$s_{11} \text{ -- } r_-^o(S_1)$$

$$s_{12} \text{ -- } t_-^o(S_1)$$

$$s_{21} \text{ -- } t_+^o(S_1)$$

$$s_{22} \text{ -- } r_+^o(S_1)$$

$$s_{31} \text{ -- } r_-(S_1)$$

$$s_{32} \text{ -- } t_-(S_1)$$

The six operators $r_-^o(S_1),\ldots,t_-(S_1)$, are instances of definitions (10) and (11) of Sec. 3.3. They are handled using the techniques illustrated in Example 5. Then, according to the interaction principle the response radiance distributions are given by:

$$N_+^+(S_1) = N_-^0(S_1)r_-^0(S_1) + N_+^0(S_1)t_+^0(S_1) + N_-^-(S_1)r_-(S_1) \quad (89)$$

$$N_-^+(S_1) = N_-^0(S_1)t_-^0(S_1) + N_+^0(S_1)r_+^0(S_1) + N_-^-(S_1)t_-(S_1). \quad (90)$$

By repeating this process of application of the interaction principle to surface S_2, we arrive at the analogous pair of statements:

$$N_+^+(S_2) = N_-^0(S_2)r_-^0(S_2) + N_+^0(S_2)t_+^0(S_2) + N_-^-(S_2)r_-(S_2) \quad (91)$$

$$N_-^+(S_2) = N_-^0(S_2)t_-^0(S_2) + N_+^0(S_2)r_+^0(S_2) + N_-^-(S_2)t_-(S_2). \quad (92)$$

An inspection of (89) and (91) shows that these equations are identical in structure; similarly for (90) and (92). The present choice of coordinate frames has rendered the formulation completely symmetrical with respect to S_1 and S_2. It is of interest to note that the domain of integration of the operator $r_-(S_1)$, e.g., may be limited at each x of S_1 to $D(S_2,x)$, and that of $r_-(S_2)$, may be limited to $D(S_1,y)$ at each y of S_2. Similar observations hold for $t_-(S_i)$, $i = 1$ or 2.

The solution procedure of the system (89)-(92) will not be exhibited; it is similar in all essential respects to that for the system (76)-(79). Those who wish to solve (91)-(92) explicitly should observe that the present counterparts to (80), (81) are given by the symmetric pair of auxiliary equations:

$$N_+^+(S_2) = N_-^-(S_1) \quad (93)$$

$$N_+^+(S_1) = N_-^-(S_2) \quad (94)$$

where the domain of the distributions are suitably restricted. Thus, e.g., (93) is understood to state that

$$N_+^+(y,\xi) = N_-^-(x,\xi) \quad (95)$$

for every x in S_1 and y in S_2 such that

$$\xi = (x-y)/|x-y| \quad .$$

By allowing S_1 and S_2 to be mutual point sources as in the preliminary example of Sec. 3.1, and by setting $N_+^0(S_i) = 0$, $i = 1$ or 2, the reader may easily show that (89) can be reduced to (2) of Sec. 3.1 and that (91) can be reduced to (3) of Sec. 3.1. In this reduction, observe that D_{12} in Fig. 3.2 is now replaced by $-D(S_2,x)$, and D_{21} by $-D(S_1,y)$. Of even more interest is the fact that the present formulation contains as a limiting case all the preceding examples on infinite parallel planes (set S_1 and S_2 parallel, and let them become arbitrarily large).

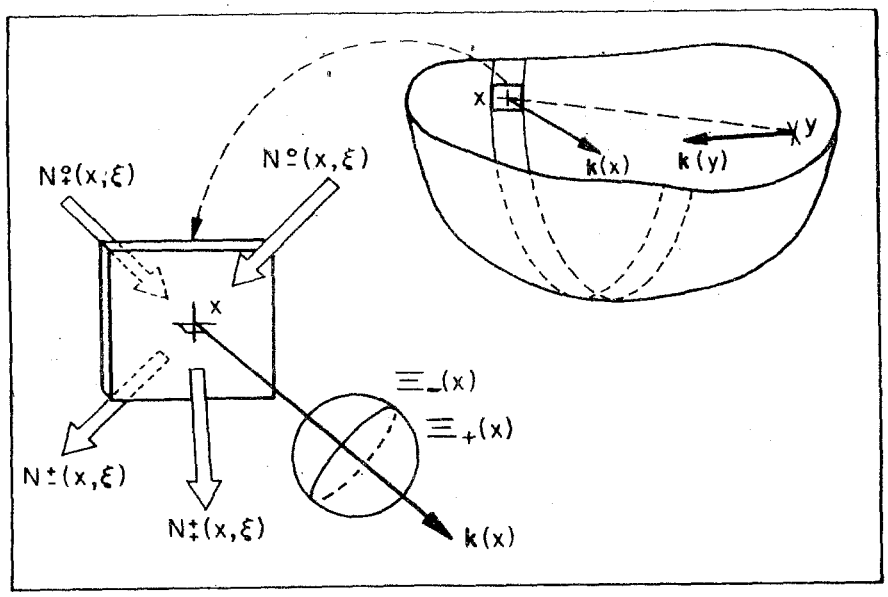


FIG. 3.14 Geometric conventions for radiometry on open concave surfaces.

3.5 Applications to Curved Surfaces

The distinguishing feature of curved surfaces for radiometry in general and the interaction principle in particular is the fact that such surfaces, unlike plane surfaces, may interact radiometrically with themselves. In this section we illustrate the application of the interaction principle to curved surfaces with this feature of self-interaction particularly in mind.

Example 1: Open Concave Surfaces

As a first illustration, consider a smooth open concave surface S in an optical medium X which is otherwise a vacuum. S is of finite extent and, as depicted in Fig. 3.14, has the general appearance of a dish or bowl. Each point x on S is visible to every other point y on S. At each point x of S we erect a unit outward normal $\mathbf{k}(x)$ which automatically determines the *outward* hemisphere: $\Xi_+(x)$; and *inward* hemisphere: $\Xi_-(x)$, at x. Instead of going into complete analytical specifications of the sense of "outward", we let Fig. 3.14 help fix the sense which is intended: the angles between $\mathbf{k}(x)$ and the directions to every other y in S from x are less than 90°. Here "outward" direction at x, as usual, means "away from S" in the immediate vicinity of x along some specified direction. By traveling in an outward direction from a plane surface, one is carried ever farther from the plane.

In the case of a curved surface such as S however, by traveling far enough along some outward directions from x, one can eventually reach S again at a point y and make contact while traveling along an inward direction at y. This elementary observation is a key observation needed in the formulation of the present interreflection problem. We let S be irradiated at each point by steady inward and outward incident radiance distributions $N_+^o(S)$ which are conveniently thought of as originating at places other than points on S. Thus the value of $N_+^o(S)$ at x and in the direction ξ in $\Xi_+(x)$ is $N_+^o(x,\xi)$. For example, if S is a portion of the sea surface at an instant of time, then $N_-^o(S)$ is the radiance distribution over that part of the sky visible at the point of interest, and $N_+^o(S)$ is the radiance distribution of that part of the underwater scene visible at the point of interest. These sources initiate and sustain an interreflection process on S where it is now possible for the immediate neighborhoods of every pair of points x and y of S to interact radiometrically. Returning to Fig. 3.14, let "$N_+^+(S)$" denote the resultant response radiance distribution over S. Thus for every x in S and ξ in $\Xi_+(x)$, $N_+^+(x,\xi)$ is the resultant surface radiance of S at x in the direction ξ. As usual, the superscript "+" denotes surface radiance, and the subscript denotes that ξ is in $\Xi_+(x)$. Furthermore let "$N_-^-(S)$" denote the resultant field radiance distribution over S. Thus for every x in S and ξ in $\Xi_-(x)$, $N_-^-(x,\xi)$ is the resultant field radiance at x in the direction ξ. In summary, then, $N_-^-(S)$ plays the role of an incident radiance on S, and $N_+^+(S)$ that of a response radiance of S.

The connection between $N_+^+(S)$ and $N_-^-(S)$ is readily established using the radiance invariance law. We have for every distinct pair x,y of points in S:

$$N_+^+(y,\xi) = N_-^-(x,\xi) \tag{1}$$

whenever x and y are two points whose common line lies in a vacuum, and:

$$\xi = (x-y)/|x-y| \quad .$$

The reader will find it of interest to compare (95) of Sec. 3.4 and (1), and dwell on the points of similarity between the formulations of Example 8 of Sec. 3.4 and those of the present example. In particular he may ask: which of the two problems is more general (in the usual sense that the more general problem yields as a special case the less general problem)?

Equation (1) can be stated in a more useful manner by first letting "D(S,x)" denote the set of all directions ξ from points y in S to the fixed point x. Thus D(S,x) is analogous to the sets $D(S_2,x)$ and $D(S_1,y)$ of Example 8 of Sec. 3.4. Then (1) holds at x in S for every ξ in D(S,x), where $y = x-r\xi$, and $r = |x-y|$. Observe that D(S,x) is part of $\Xi_-(x)$ for every x.

We are now ready to use the interaction principle to formulate the present problem. We isolate S and then enumerate the sets of all incident radiometric functions on S.

A_1: all field radiance distributions like $N_-^o(S)$

A_2: all field radiance distributions like $N_+^o(S)$

A_3: all field radiance distributions like $N_-^-(S)$

Enumerating all the sets of response radiometric functions for S:

B_1: all surface radiance distributions like $N_+^+(S)$

B_2: all surface radiance distributions like $N_-^+(S)$

In the present case m = 3, n = 2, and the interaction principle yields the following six interaction operators s_{ij}:

$$s_{11} \text{ -- } r_-^o(S)$$

$$s_{12} \text{ -- } t_-^o(S)$$

$$s_{21} \text{ -- } t_+^o(S)$$

$$s_{22} \text{ -- } r_+^o(S)$$

$$s_{31} \text{ -- } r_-(S)$$

$$s_{32} \text{ -- } t_-(S)$$

The six operators $r_-^o(S),\ldots, t_-(S)$ are instances of definitions (10), (11) of Sec. 3.3. Then, according to the interaction principle, $N_\pm^+(S)$ are given by:

$$N_+^+(S) = N_-^o(S)r_-^o(S) + N_+^o(S)t_+^o(S) + N_-^-(S)r_-(S) \qquad (2)$$

$$N_-^+(S) = N_-^o(S)t_-^o(S) + N_+^o(S)r_+^o(S) + N_-^-(S)t_-(S) \qquad (3)$$

This pair of interaction equations and the auxialiary equation:

$$N_+^+(S) = N_-^-(S) \qquad (4)$$

form an autonomous system of equations. The latter equation is understood in the sense of (1). The order of solution of the equations is dictated by (2): it must be solved first. Thus using (4) in (2) we have:

$$\boxed{N_+^+(S) = A_+(S) + N_+^+(S)r_-^!(S)} \qquad (5)$$

where we have written:

"$A_+(S)$" for $N_-^o(S)r_-^o(S) + N_+^o(S)t_+^o(S)$,

and where we have written:

$$\text{"}r'_{-}(S)\text{"} \quad \text{for} \quad \int_{D(S,x)} [\ ]'r_{-}(x;\xi';\xi)\, d\Omega(\xi') \tag{6}$$

The prime (') on the square bracket denotes retardation of the argument of the radiance function on which $r'_{-}(S)$ operates. Thus, if $N^{\ddagger}(S)$ is to be evaluated at x and for ξ in $\Xi_{+}(x)$, then (5) becomes:

$$N^{+}_{+}(x,\xi) = A_{+}(x,\xi) + \int_{D(S,x)} N^{+}_{+}(x-r'\xi',\xi')r_{-}(x;\xi';\xi)\, d\Omega(\xi')$$

where "$A_{+}(x,\xi)$" denotes the value of $A_{+}(S)$ at x and ξ. It is clear that $r'_{-}(S)$ can operate on any element in the response set B_1. It is particularly to be noted that $A_{+}(S)$ is an element of B_1. Thus, solving (5), we have:

$$N^{+}_{+}(S) = A_{+}(S)[I - r'_{-}(S)]^{-1} \tag{7}$$

and the inverse of $[I - r'_{-}(S)]$ exists provided $r'_{-}(S)$ is norm contracting (cf. (60) of Sec. 3.4).

The prime on the operator $r'_{-}(S)$ is adequate to communicate the idea of a retarded argument in (6) and (7) to the general reader whose insight into our intentions fortunately can lighten our expository task. However, if (7) is to be programmed for evaluation on an automatic computer, then another--a more mechanical--expedient must be devised to communicate the idea of retarded arguments of a function. For example we could define a mapping $t_0(S)$ which assigns to every x in S and ξ in $\Xi_{-}(x)$ the point $x - r\xi$ where r is the distance from x to S along the direction $-\xi$. Knowing the analytic description of S, it is in principle possible to compute this r for each x and ξ in $\Xi_{-}(x)$ and hence to construct $t_0(S)$. Then "$N^{\ddagger}_{+}(S)t_0(S)$" will denote the function which assigns to every x in S and ξ in $\Xi_{-}(x)$ the radiance $N^{-}_{-}(x,\xi)$ $(= N^{\ddagger}_{+}(x-r\xi,\xi))$ where $x - r\xi$ is on S and ξ is in $\Xi_{+}(x-r\xi)$. With this definition of $t_0(S)$, we can rewrite (4) as:

$$N^{-}_{-}(S) = N^{+}_{+}(S)t_0(S) \tag{8}$$

without the need of further qualifications as was necessary in qualifying (4) by (1). Then using (8) in (2), the more detailed version of (5) is:

$$N^{+}_{+}(S) = A_{+}(S) + N^{+}_{+}(S)t_0(S)r_{-}(S)$$

from which the more detailed version of (7) follows:

$$\boxed{N^{+}_{+}(S) = A_{+}(S)[I - t_0(S)r_{-}(S)]^{-1}} \ . \tag{9}$$

It is easy to see that if $r_-(S)$ is norm contracting, then so is $t_o(S)r_-(S)$, where we have written:

$$\text{"}t_o(S)r_-(S)\text{"} \quad \text{for} \quad \int_{D(S,x)} [\cdot t_o(S)]r_-(x;\xi';\xi)\, d\Omega(\xi') \qquad (10)$$

Here any function on which $t^o(S)r_-(S)$ operates automatically has its argument x,ξ (ξ in $\Xi_-(x)$) first retarded to $x-r\xi$, and ξ respectively (ξ now considered in $\Xi_+(x-r\xi)$). With this definition of $t_o(S)$, (9) and (7) are equivalent ways of indicating the computation of the response function $N_+^+(S)$. The computation of $N_-^+(S)$ can be performed using (8), (9) and (3).

In closing we note that one can also view the object $t_o(S)$ as a mapping from response set B into incident set A. This interpretation is based on (8). Such mappings occur naturally in the strictly mechanical formulations of the auxiliary equations arising from step (vi) in the interaction method (cf. Example 2, Sec. 3.4).

Example 2: Closed Concave Surfaces; the Integrating Sphere

In the present example we allow the rim of the surface S in Fig. 3.14 to diminish in diameter while leaving the area of S greater than some fixed constant. Thus S becomes a closed concave surface (as seen from within). It is the purpose of this example to point out that the formulations of Example 1 remain unchanged as the open concave surface becomes a closed concave surface. Indeed, as a review of Example 1 would show there is no essential use made at all of the openness of S as depicted in Fig. 3.14. The only important change to note is that $D(S,x)$ is now exactly $\Xi_-(x)$ for every x in a closed concave surface S. Hence (9) holds also for closed concave surfaces. We shall now illustrate (9) for the most useful case of a closed concave surface: the integrating sphere.

Figure 3.15 depicts a spherical surface S of diameter d enclosing a vacuous region. Incident source radiance is restricted to a general part a of S. For simplicity we let the incident source radiance be represented by $N_+^o(S)$ over part a of S. Hence we will write "$N_+^o(a)$" for $N_+^o(S)$ and set $N_-^o(S) = 0$ in $A_+(S)$ of Example 1. $N_+^o(a)$ is of arbitrary directional structure but is independent of location over a with respect to the local direction frame, determined at each point y by $\mathbf{k}(y)$. Then (9) becomes:

$$N_+^+(S) = N_+^o(a)t_+^o(a)\left[I - t_o(S)r_-(S)\right]^{-1} \quad . \qquad (11)$$

We next adopt the classical assumption that the inside surface of S is a lambert reflector. In addition we assume a is a lambert transmitter. That is, we are assuming (cf. (17) of Sec. 3.3):

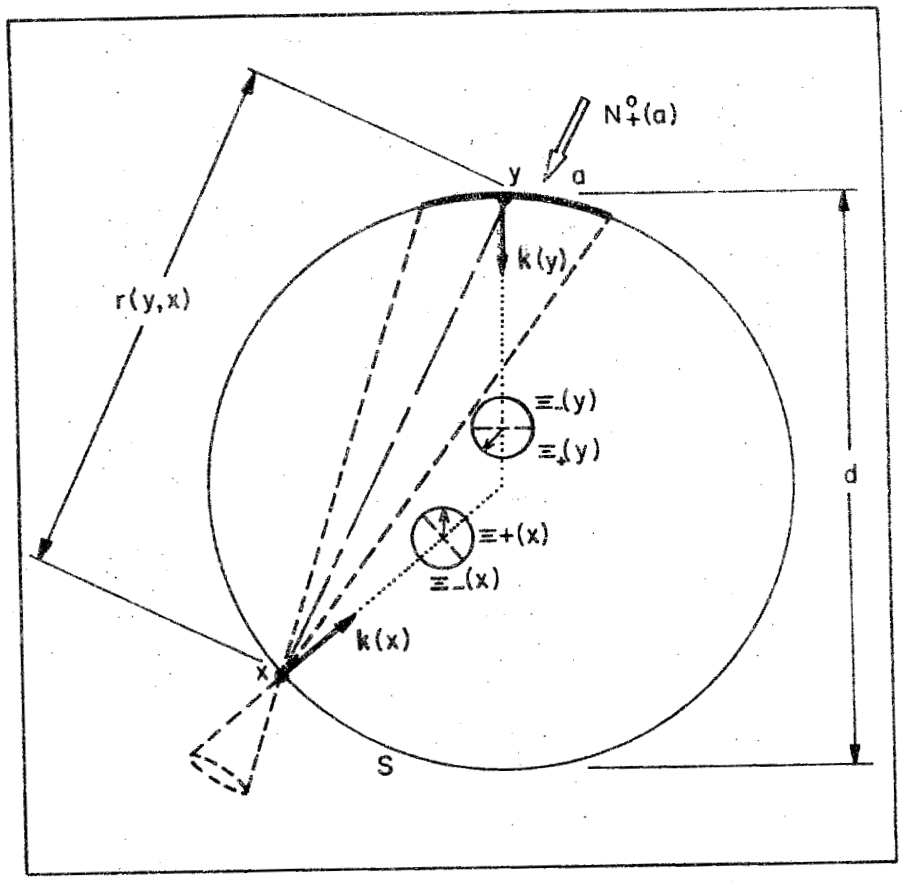


FIG. 3.15 Illustrating the radiometric self-interaction of a closed concave surface. The case of the integrating sphere.

$$r_-(x;\xi';\xi) = \frac{r_-}{\pi}|\xi'\cdot\mathbf{k}(x)|$$

for every x in S, ξ' in $\Xi_-(x)$, and ξ in $\Xi_+(x)$; and with $0 \leq r_- < 1$. Further we assume:

$$t_+(x;\xi';\xi) = \frac{t_+}{\pi}\,\xi'\cdot\mathbf{k}(x)$$

for every x in a, ξ' in $\Xi_+(x)$, and ξ in $\Xi_+(x)$. Then:

$$N^o_+(a)t^o_+(a) = \frac{t_+}{\pi}\,H^o_+f_a$$

where f_a is a function on S such that $f_a(x) = 1$ if x is in a and $f_a(x) = 0$ if x is not in a. Further, H^o_+ is the constant

irradiance on a produced by $N_+^o(a)$. Hence $N_+^o(a)t_+^o(a)$ is a uniform radiance distribution over a and is a member of set B_1, (see Example 1) which we shall denote by "N^o".

We now write (11) in the form:

$$N_+^+(s) = N^o\left[I + \sum_{j=1}^{\infty} (t_o(S)r_-(S))^j\right]$$

$$= N^o + \sum_{j=1}^{\infty} N^o(t_o(S)r_-(S))^j \qquad (12)$$

Consider the first term:

$$N^o t_o(S) r_-(S)$$

of the infinite series. We write "N^1" for $N^o t_o(S)r_-(S)$. Then by (10):

$$N^1(x,\xi) = \int_{D(S,x)} N^o(x-r'\xi',\xi')r_-(x;\xi';\xi)\, d\Omega(\xi')$$

for every x in S and ξ in $\Xi_+(x)$. Since the incident source N^o vanishes outside part a of S and is of constant magnitude within a, the domain of integration shrinks from D(S,x) to D(a,x), and

$$N^1(x,\xi) = \frac{t_+}{\pi} H_+^o \int_{D(a,x)} r_-(x;\xi';\xi)\, d\Omega(\xi')$$

$$= \frac{t_+}{\pi} H_+^o \frac{r_-}{\pi} \int_{D(a,x)} |\xi'\cdot\mathbf{k}(x)|\, d\Omega(\xi') \quad .$$

Now by means of an observation following (22) of Sec. 2.11, this integral is readily evaluated:

$$\int_{D(a,x)} |\xi'\cdot\mathbf{k}(x)|\, d\Omega(\xi') = \frac{A(a)}{d^2} \quad .$$

Hence

$$N^1(x,\xi) = \frac{t_+}{\pi} H_+^o\left[r_- \frac{A(a)}{\pi d^2}\right] \quad ,$$

for every x in S and ξ in $\Xi_+(x)$. This result was grouped in the indicated manner to show the effect of the operator $t_o(S)r_-(S)$ on N^o. The effect is to multiply the value of N^o by $[r_- A(a)/\pi d^2] = r_- A(a)/A(S)$.

The uniform surface radiance distribution N^1 is now acted on by $t_o(S)r_-(S)$ to yield the second term of the series.

Observe that N^1 is constant-valued over S, whereas N^0 is constant-valued over a and zero over the remainder of S outside a. Thus the second term of the series is:

$$N^1 t_0(S) r_-(S) = \left(N^0 t_0(S) r_-(S)\right)\left(t_0(S) r_-(S)\right)$$

$$= N^0 \left(t_0(S) r_-(S)\right)^2 \quad .$$

Let us write "N^2" for $N^1 t_0(S) r_-(S)$. Then:

$$N^2(x,\xi) = \int_{D(S,x)} N^1(x-r'\xi',\xi') r(x;\xi';\xi)\, d\Omega(\xi')$$

$$= \frac{t_+}{\pi} H_+^0 \left[r_- \frac{A(a)}{A(S)} \right] \frac{r_-}{\pi} \int_{D(S,x)} |\xi' \cdot \mathbf{k}(x)|\, d\Omega(\xi') \quad .$$

With this second iterate of $t_0(S) r_-(S)$, the pattern begins to form. We first note that:

$$\int_{D(S,x)} |\xi' \cdot \mathbf{k}(x) d|\, d\Omega(\xi') = \frac{A(S)}{d^2} = \pi \quad ,$$

for every x in S. Hence:

$$N^2(x,\xi) = \frac{t_+}{\pi} H_+^0 \left[r_- \frac{A(a)}{A(S)} \right] r_-$$

$$= N^1(x,\xi) r_- \quad .$$

Thus if we write: for $j = 1,2,\ldots,$:

$$\text{"}N^j\text{"} \quad \text{for} \quad N^{j-1} t_0(S) r_-(S) \quad ,$$

then for $j = 2,3,\ldots,$:

$$N^j(x,\xi) = \int_{D(S,x)} N^{j-1}(x-r'\xi',\xi) r_-(x;\xi';\xi)\, d\Omega(\xi')$$

$$= N^{j-1}(x,\xi) \int_{D(S,x)} r_-(x;\xi';\xi)\, d\Omega(\xi')$$

$$= N^{j-1}(x,\xi)\, r_- \quad .$$

Since

$$N^2(x,\xi) = N^1(x,\xi) r_- \quad ,$$

for every x in S and ξ in $\Xi_+(x)$, we then have:

$$N^j(x,\xi) = N^1(x,\xi)(r_-)^{j-1}$$

for $j = 2,3,\ldots$.

Hence:
$$\sum_{j=1}^{\infty} N^0(t_0(S)r_-(S))^j = N^1 + \sum_{j=2}^{\infty} N^j$$

$$= N^1 + N^1 \sum_{j=2}^{\infty} (r_-)^{j-1}$$

$$= N^1\left(\sum_{j=0}^{\infty} r_-^{\,j}\right)$$

$$= \frac{N^1}{[1-r_-]} \quad .$$

It follows from (12) that

$$N_+^+(S) = N^0 + \frac{1}{\pi}\frac{\left[r_-t_+H_+^0A(a)\right]/A(S)}{[1-r_-]} \quad .$$

Hence using the explicit values of N^0, we have for every x in S and ξ in $\Xi_+(x)$:

$$N(x,\xi) = \frac{t_+H_+^0}{\pi}\left[f_a(x) + \frac{r_-A(a)/A(S)}{[1-r_-]}\right] \qquad (13)$$

where as noted before $f_a(x) = 1$ if x is in a, and $f_a(x) = 0$ if x is not in a. This formula shows that $N(x,\xi)$ is of uniform directional structure over $\Xi_+(x)$ at each x in S, and is independent of x over a and over the part S-a of S outside of a. However, the radiance distributions over a exceed those in S-a by an amount $t_+H_+^0/\pi$, which is precisely the transmitted radiance through the "window" a of the integrating sphere S. (Observe that no essential use has been made of the sphericity of the surface S. Hence we should expect to extend (13), with only minor changes, to the case of an arbitrary closed surface with lambert properties.)

Example 3: Open and Closed Convex Surfaces

The need to illustrate in great detail the interaction principle for the case of open and closed convex surfaces is obviated by the observation that concavity of surfaces is a

relative property, that is, a property relative to the vantage point of the observer. Thus the surface S in Fig. 3.14 is concave relative to an observation point inside the space enclosed by S--i.e., within the bowl of S. On the other hand it appears convex when viewed from below S in the Figure. The interaction equations automatically adjust, without alteration of their general forms, to these two points of view and equations (2), (3), (4) hold also for the convexity interpretation. The only changes in (2)-(4) that might occur are those associated with a reversal of direction of $\mathbf{k}(x)$. *Accordingly, if the user deems to introduce this change, so as to study convex surfaces, then all subscripts "+" and "-" in (2)-(4) and their logical descendants may be interchanged: every occurrence of subscript "+" may be replaced by "-", and conversely.*

Example 4: General Two-Sided Surfaces

In this example we ascend one more rung with respect to the generality of the type of surface considered: we shall apply the interaction principle to general self-interacting one-piece, two-sided surfaces which may be either locally concave or convex, i.e., have alternating hollows or hills. The surfaces may be closed in the sense of enclosing a volume, or open. We assume their reflectance and transmittance properties are known at each point and that they are embedded in a vacuum. As a concrete illustration that may be kept in mind during the following discussion, the instantaneous configuration of a dynamic wind-blown air-water surface will serve well. An application of this example of the air-water surface is made in Sec. 12.10.

Parts (a)-(g) of Figure 3.16 depict some particular instances falling within the scope of the present discussion. An examination of these general surfaces reveals two features which were not present in the cases of concave or convex surfaces considered above. First, for some point x of S there may be points y of S such that x and y are not mutually visible. That is, on the straight line between x and y there lies at least one other point of S. Second, for some points x of S there may be points y of S such that x and y are mutually visible but are on opposite sides of S.

The interaction principle is immediately applicable to general surfaces such as those depicted in Fig. 3.16, and which have the two additional features just described. In order to display the application so that it may be useful in practice and be subject to mechanical manipulation, it is desirable to introduce some preliminary geometric concepts. First we assign, as usual, a unit outward normal $\mathbf{k}(x)$ to S at each x. This fixes $\Xi_+(x)$ at each x and arbitrarily determines the outside and the inside of S. Further, we let "D(S,z)" denote in general the set of all directions from points y of S to a point z. D(S,x) consists of directions ξ in either $\Xi_+(x)$ or $\Xi_-(x)$. Further if ξ is extended to meet S at y, then ξ may be also in either $\Xi_+(y)$ or $\Xi_-(y)$. See, e.g., parts (a) and (e) of Fig. 3.16. It will be necessary to distinguish between such members of D(S,x) which are in $\Xi_+(y)$ or $\Xi_-(y)$.

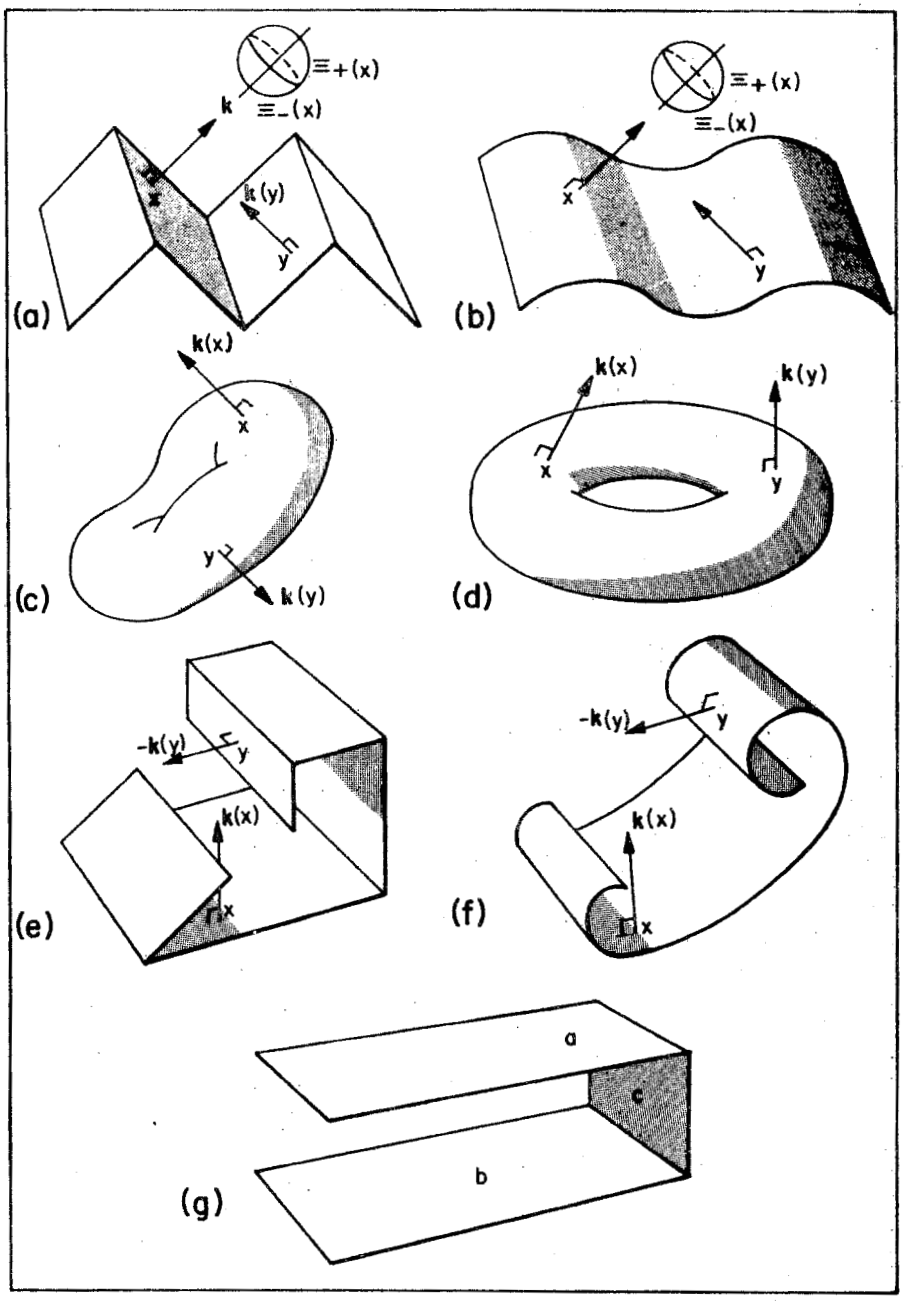


FIG. 3.16 Radiometric self-interaction of two-sided surfaces.

We shall do this in the following way. To begin, for every x in S and ξ in $D(S,x)$ we let "$r_m(x,\xi)$" or simply "r_m" denote the least non negative r such that $x-r\xi$ is a point of S. The geometrical significance of r_m is clear: if at x in S we go along a straight line in direction $-\xi$, then eventually we may reach S, and since some surfaces are corrugated (as in parts (a) or (b) of Fig. 3.16) we may reach S again and again if we continue to travel along the straight line in the direction $-\xi$. The distance r_m is the distance to the first of such meetings with S. Next we let "$t_m(S)$" denote the function which assigns to each x in S and ξ in $D(S,x)$ the point $x-r_m\xi$ on S and the direction ξ. That is, $t_m(x,\xi) = (x-r_m\xi,\xi)$ for every x in S and ξ in $D(S,x)$. Hence $t_m(S)$ is a mapping which is an immediate generalization of the mapping $t_o(S)$ introduced in Example 1. Finally we construct two functions $\chi_+(S)$ and $\chi_-(S)$ with the following properties: for every x in S and ξ in $D(S,x)$, $\chi_\pm(x,\xi) = 1$ if ξ is in $\Xi_\pm(x)$ and $\chi_\pm(x,\xi) = 0$ if ξ is not in $\Xi_\pm(x)$. Finally, for every x in S and ξ not in $D(S,x)$, $\chi_\pm(x,\xi) = 0$. With these geometric preliminaries completed, we can now go on directly to the application of the interaction principle.

Let $N^o_\pm(S)$ be steady incident source radiance distributions on S. These incident distributions generally initiate and sustain a self-interreflection process over S. Let $N^+_\pm(S)$ and $N^-_\pm(S)$ be the resultant surface and field radiance distributions over S.

We first isolate S and then enumerate the incident radiometric functions on S:

A_1: all field radiance distributions like $N^o_-(S)$

A_2: all field radiance distributions like $N^o_+(S)$

A_3: all field radiance distributions like $N^-_-(S)$

A_4: all field radiance distributions like $N^-_+(S)$.

The sets of response functions of S are:

B_1: all surface radiance distributions like $N^+_+(S)$

B_2: all surface radiance distributions like $N^+_-(S)$.

In the present case $m = 4$, $n = 2$, and the interaction principle yields the following eight interaction operators s_{ij}:

$$s_{11} \text{ -- } r^o_-(S)$$

$$s_{12} \text{ -- } t^o_-(S)$$

$$s_{21} \text{ -- } t^o_+(S)$$

$$s_{22} \text{ -- } r^o_-(S)$$

$$s_{31} \text{ -- } r_-(S)$$

$$s_{32} \text{ -- } t_-(S)$$

$$s_{41} \text{ -- } t_+(S)$$

$$s_{42} \text{ -- } r_+(S)$$

The preceding eight operators $r_-^o(S),\ldots,r_+(S)$ are instances of definitions (10), (11) of Sec. 3.3. Then, according to the interaction principle, $N_\pm^+(S)$ are given by:

$$N_+^+(S) = N_-^o(S)r_-^o(S) + N_+^o(S)t_+^o(S) + N_-^-(S)r_-(S) + N_+^-(S)t_+(S) \tag{14}$$

$$N_-^+(S) = N_-^o(S)t_-^o(S) + N_+^o(S)r_+^o(S) + N_-^-(S)t_-(S) + N_+^-(S)r_+(S). \tag{15}$$

From the radiance invariance-law and the definitions of the geometric functions $X_\pm(S)$, $t_m(S)$, we have the two auxiliary equations:

$$N_-^-(S) = N_+^+(S)\cdot X_+(S)t_m(S) + N_-^+(S)\cdot X_-(S)t_m(S) \tag{16}$$

$$N_+^-(S) = N_+^+(S)\cdot X_+(S)t_m(S) + N_-^+(S)\cdot X_-(S)t_m(S) \ . \tag{17}$$

These auxiliary equations together with (14) and (15) constitute an autonomous system of integral equations governing the surface radiance distribution $(N_+^+(S),N_-^+(S))$ over an arbitrary two-sided surface S. The dots denote multiplication of functions, and the multiplication is done after the operation $t_m(S)$ is applied to $X_\pm(S)$ and $N_\pm^+(S)$.

As an illustration of the use of (16), let x be a point of S and let ξ be in $D(S,x)$. Then the value of $t_m(S)$ at (x,ξ) is $(x-r_m\xi,\xi)$ and this is used in the argument of $X_-(S)$. It follows that one of the two values $X_+(x-r_m\xi,\xi)$ or $X_-(x-r_m\xi,\xi)$ must be 0, and the other 1. Say the former is 0 and the latter is 1. Then the value of $N_-^-(x,\xi)$ of $N_-^-(S)$ at x, is:

$$N_-^-(x,\xi) = N_-^+(x-r_m\xi,\xi) \ .$$

Hence the downward field radiance $N_-^-(x,\xi)$ comes from the downward surface radiance at $x-r_m\xi$. Thus x and $x-r_m\xi$ are on opposite sides of S, and so S must be curled like that in (e) or (f) of Fig. 3.16. This illustration shows that $X_-(S)t_m(S)$ is to be interpreted as the *composition* of the functions $X_-(S)$ and $t_m(S)$.

The integral operations in (14) and (15) are all generally alike. For the purposes of illustration we take $N_\mp^-(S)t_+(S)$ as typical. Then, according to (11) of Sec. 3.3 with $Y = S_1$, the value of $N_\mp^-(S)t_+(S)$ at x in S and ξ in $\Xi_+(x)$ is:

$$\int_{\Xi_+(x)} N_+^-(x,\xi')t_+(x;\xi';\xi)\ d\Omega(\xi')$$

It is clear that we may replace $\Xi_+(x)$ by $\Xi_+(x) \cap D(S,x)$, i.e., by that part of $D(S,x)$ in $\Xi_+(x)$. Furthermore, if we use (17) to replace $N_+^-(x,\xi')$ in the integrand, the integral becomes:

$$\int_{\Xi_+(x) \cap D(S,x)} \Big[N_+^+(x-r_m\xi',\xi') \cdot X_+(x-r_m\xi',\xi') + N_-^+(x-r_m\xi',\xi') \cdot X_-(x-r_m\xi',\xi') \Big] t_+(x;\xi';\xi)\, d\Omega(\xi')$$

The reader should now examine the set (14)-(17) with the purpose in mind of noting that the set contains as special cases the convex and concave examples above. Plane surfaces are also covered; for then $D(S,x)$ has zero solid angle measure for every x on S and the last two terms in (14), (15) vanish by virtue of (16), (17) and the definitions of $X_\pm$. The preceding illustration bears this out in part. Furthermore, by invoking a certain amount of geometric-radiometric trickery, the set (14)-(17) can also yield, in the limit, the cases of a set of finite or infinite parallel planes. For example, S may have the three part configuration as in part (g) of Fig. 3.16, with parts a and b parallel planes and part c having zero reflectance and unit transmittance. This would yield the case of parallel finite planes. These observations will make plausible the assertion that the system (14)-(17) actually constitutes the interaction equations for a general collection of two-sided surfaces S, where S is either in one piece or several distinct pieces, and of concave, convex, or mixed curvature. It is not intended, however, that the set (14)-(17) itself always be reduced to each case as it arises. We have exhibited the preceding interaction equations mainly to show the scope of the interaction method and the mechanical ease with which it can formulate radiometric interaction problems. It is desirable, rather, especially for students of the subject, that each specific instance of an interaction problem be derived anew from the principle and that simplifications be made and auxiliary equations invoked which are motivated by the particular features of the individual case.

Example 5: General One-Sided Surfaces

In this example we apply the interaction method to the formulation of the interreflection light field over one-sided surfaces. Before going into the details it is of interest to observe that one-sided surfaces are mathematical objects which arose originally in critical studies of the classical surface integration theorems of Stokes and Gauss. One-sided surfaces were used principally as counterexamples to show the limitations of the classical forms of these theorems. It is because of this predominantly negative role played by one-sided surfaces in the early training of physics and mathematics students, and because of the spectacular and intuitively bizarre claims made for these surfaces, that a student eventually carries away with him the general impression that

one-sided surfaces are conceptual beasts which are inferior to their more applicable two-sided cousins, and are better left alone. This impression is, for the most part, defensible since the classical surface integration theorems in the usual physical applications of mathematics pertain only and implicitly to two-sided surfaces. The implicitness of the two-sided condition is eventually forgotten, and the ingrained avoidance of one-sided surfaces prevents their use in the application of the usual theorems. However, with a little additional effort the one-sided surfaces can occasionally be brought into physical discussions and their physical properties compared--usually with deeper resultant insight--with the corresponding properties of two-sided surfaces. In this example we shall perform this service for the radiative transfer context. We shall briefly consider the interaction principle applied to the most notorious of one-sided surfaces, the Möbius Strip. What we shall find out in this application will be typical of the radiometric properties of one-sided surfaces in general, and no more dramatic than the simple but useful insight that it takes exactly one half the number of equations to formulate the radiometric interaction equations for one-sided surfaces as it does for two-sided surfaces. Hence the four general equations of Example 4 will be reduced to two for the most general one-sided surface.

To help fix ideas we shall adopt as the prototype of one-sided surfaces the Möbius strip depicted in Fig. 3.17. The Möbius strip S is shown in plan view in part (b) of Fig. 3.17 and its mode of generation is shown in perspective in (a) of Fig. 3.17. To generate S, one can imagine first of all a circle C_o of radius r_o in a plane. Then a line segment L of length 2a is placed so that its center is on C_o and so that the line segment, extended, goes through the center of C_o. If L is moved around C_o with this orientation maintained, and remaining in the plane of C_o, L will sweep out a circular annulus of radii r_o+a and r_o-a. To generate the Möbius strip S itself, instead of keeping L in the plane of C_o, now, keeping L perpendicular to C_o, let L rotate with its center always on C_o, and at a uniform angular speed so that as L makes one circuit of C_o, it will turn 180° about its point of contact with C_o. The equations for this Mobius strip are given in parametric form using cylindrical coordinates as shown in (b) of Fig. 3.17:

$$r = r_o + \rho \cos(\phi/2)$$

$$\phi = \phi$$

$$z = \rho \sin(\phi/2)$$

$$-a \leq \rho \leq a, \quad 0 < a < r_o, \quad 0 \leq \phi \leq 4\pi \qquad (18)$$

where ρ and ϕ are parameters for the surface and r, ϕ, z are the usual cylindrical coordinate variables. The surface (the set of points in space) is then completely specified once we give the magnitudes a and r_o.

The parameters ρ and ϕ can be used to locate a point on S just as latitude and longitude are used to locate a point

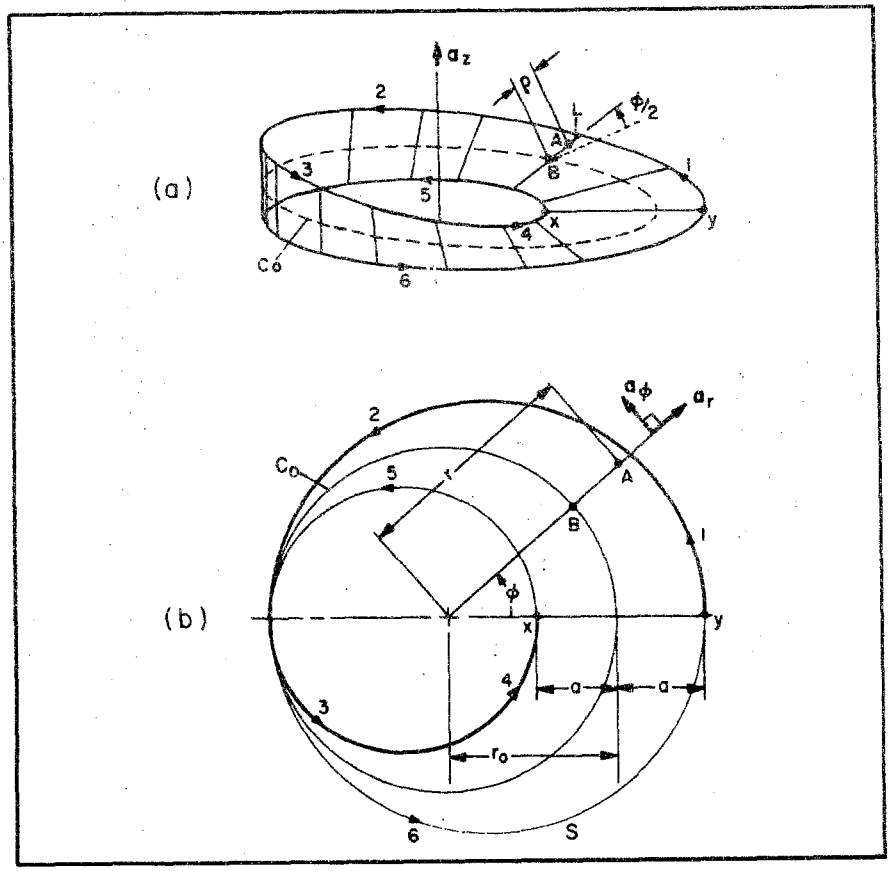


FIG. 3.17 Radiometric self-interaction of one-sided surfaces. The case of the Möbius strip.

on the surface of the earth. For example, given $\rho = a, \phi = 90°$; one finds the corresponding point on S by walking on S along C_o in the direction of increasing ϕ until a point on C_o is reached which is on the 90° radial line through the origin. Then, still standing on S and still facing in the direction of increasing ϕ, one's extended right arm points in the direction of increasing ρ; and one's extended left arm points in the direction of decreasing ρ. In the present case one goes a units of distance along S to the right to get to $(a, 90°)$. The coordinates at this point in the cylindrical reference frame are obtained by setting $\phi = 90°$ and $\rho = a$ in the system (18) of equations:

$$r = r_o + a\sqrt{2}/2$$

$$\phi = 90°$$

$$z = a\sqrt{2}/2 \quad .$$

Now, from a radiometric point of view any realization of the surface S will serve as a perfectly good example of a passive reflector and transmitter of radiant flux. By placing a realization of S say, in the form of a translucent matte white plastic strip out in sunlight, one can observe the play of light over its surface and see the interreflection effects where flux from one part augments the natural external flux falling on another part. Hence, the one-sidedness of S does not at all interfere with the inherent interactions of S with light. The one-sidedness of S enters the picture when an unsuspecting human observer wishes to unambiguously fix the unit outward $\mathbf{k}(x)$ at some point x of S preparatory to constructing the sets $\Xi_+(x)$ and $\Xi_-(x)$ of outward and inward directions relative to S at x. The observer soon sees that a certain property of "two-sided" surfaces is now lacking--a property which is so deeply ingrained in our geometric intuition about surfaces that it is virtually taken as universal and as possessed by every surface. This property of two-sided surfaces may be phrased as follows: *if one travels a closed curve* C *on a two-sided surface* S, *constructing* $\Xi_+(x)$ *at each* x *of* C *from knowledge of the surface coordinates* ρ *and* ϕ *according to some fixed rule, then when the journey along* C *is brought to a close, the last set* $\Xi_+(x)$ *constructed coincides with the set* $\Xi_+(x)$ *constructed at the outset of the journey.* This property is not possessed by the Möbius Strip S. Thus suppose we agree that, as we stand at any x on S, the direction of $\mathbf{k}(x)$ will be from one's foot to one's head. If one walks from a point x_0 around a small closed path in the neighborhood of a point x, then as one's feet return to x_0 his head will be in the same spot as when he started out. However, if he walks from x_0 around a larger circuit--say all around circle C_0, then when he returns to x_0, he will be relatively upside down from his original position. By making the circuit around C_0 once more, he will regain his original orientation. Hence for the present surface S, he can return, after traversing certain closed paths, to his starting point, but with an orientation opposite to that with which he started out. A surface S is *two-sided* (or *orientable*) if the preceding italicized statement holds for *every* C on S; otherwise a surface is *one-sided* (or *non orientable*).

The salient effect of one-sided surfaces on radiative transfer theorizing is that such surfaces obviate the necessity of considering *both* the upper $\Xi_+(x)$ and the lower hemispheres $\Xi_-(x)$ of the unit sphere Ξ. It is found that the upper hemisphere $\Xi_+(x)$, e.g., alone suffices to describe the radiometric interactions of such a surface. This may be seen as follows. First, by means of some elementary vector analysis, we see that a unit "outward" normal $\mathbf{k}(x)$ at point x on the Möbius Strip is given by the equation:

$$\mathbf{k}(x) = \frac{1}{1+\frac{\rho^2}{2}}\left[-a_r(x)\,\sin\left(\frac{\phi}{2}\right) - \frac{\rho}{2}\,a_\phi(x) + a_z(x)\,\cos\left(\frac{\phi}{2}\right)\right] \qquad (19)$$

in which we have written:

$$\text{"x"} \quad \text{for} \quad (r,\phi,z) \quad ,$$

and r, ϕ, and z are as given in (18) for every choice of parameters ρ and ϕ. Further, the unit vectors $\mathbf{a}_r(x)$, $\mathbf{a}_\phi(x)$, $\mathbf{a}_z(x)$ of the cylindrical coordinate system point along the directions of increasing r, ϕ, and z in that system, as usual. Hence $\mathbf{k}(x)$ is uniquely determined by each choice of ρ and ϕ, $-a \leq \rho \leq a$, $0 \leq \phi \leq 2\pi$. We can denote this fact also by writing:

$$\text{"}\mathbf{k}(\rho,\phi)\text{"} \quad \text{for} \quad \mathbf{k}(x) \quad .$$

Now the heart of the one-sidedness of S resides in the following fact: for every ρ,ϕ, by (18), (ρ,ϕ) and $(-\rho,\phi+2\pi)$ determine the same point x on S and:

$$\mathbf{k}(\rho,\phi) = -\mathbf{k}(-\rho,\phi+2\pi) \quad . \tag{20}$$

Next for each $\mathbf{k}(x)$ we define $\Xi_\pm(x)$ as usual, where x corresponds to (ρ,ϕ). To point up this dependence of $\Xi_\pm(x)$ on (ρ,ϕ), we shall write "$\Xi_\pm(\rho,\phi)$" for $\Xi_\pm(x)$. Then we see that:

$$\boxed{\Xi_-(\rho,\phi) = \Xi_+(-\rho,\phi+2\pi)} \quad . \tag{21}$$

This is what we wished to show: the lower hemisphere corresponding to (ρ,ϕ) is equal to the upper hemisphere corresponding to $(-\rho,\phi+2\pi)$. Hence a downward direction at (ρ,ϕ) may be thought of as an upward direction at $(-\rho,\phi+2\pi)$.

It should be recalled that the numbers ρ and ϕ are coordinates one uses while maneuvering about the surface S. If one starts at (ρ,ϕ) where $\rho = 0$ and $\phi = 0$ and walks to (ρ,ϕ) in the manner explained above, and then starts all over again and walks to $(-\rho,\phi+2\pi)$, his feet will come to the same point in space but the final positions of his head will be diametrically opposed. We shall call the number pairs (ρ,ϕ), $(-\rho,\phi+2\pi)$ *conjugate coordinates*. If a $\rho\phi$-coordinate system were painted on a realization of the Möbius Strip S, then the conjugate coordinates (ρ,ϕ) and $(-\rho,\phi+2\pi)$ could be thought of as specifying two "different" points on S. These points could be imagined different in the sense that if S were sufficiently large and opaque then two people standing on points with conjugate coordinates would be hidden from each other. However, from a construction point of view, i.e., viewing S as an assembly of points, the point located by (ρ,ϕ), once placed into position, obviated the need of placing that corresponding to $(-\rho,\phi+2\pi)$.

We can now apply these accumulated findings to the problem of formulating the radiometric interaction equations for a general one-sided surface S. If (ρ,ϕ) are the coordinates of a point on S, with ρ and ϕ variables drawn from some suitable range of numbers such as that suggested above, we let "$\mathbf{k}(\rho,\phi)$" denote the unit outward normal to S at the point with coordinates (ρ,ϕ). To each point on S with coordinates

(ρ,ϕ) there exist conjugate coordinates (ρ',ϕ') such that (ρ,ϕ) and (ρ',ϕ') determine the same point on S but:

$$\mathbf{k}(\rho,\phi) = -\mathbf{k}(\rho',\phi') \quad . \tag{22}$$

Let "Φ" denote the function which assigns to each coordinate (ρ,ϕ) its conjugate coordinate (ρ',ϕ'):

$$(\rho',\phi') = \Phi(\rho,\phi) \quad .$$

In the case of the Möbius Strip above, $\rho' = -\rho$ and $\phi' = \phi+2\pi$. Then (22) can also be written:

$$-\mathbf{k}(\rho,\phi) = \mathbf{k}\big(\Phi(\rho,\phi)\big) \quad . \tag{23}$$

Further

$$\Xi_-(\rho,\phi) = \Xi_+\big(\Phi(\rho,\phi)\big) \tag{24}$$

which may be abbreviated to

$$\Xi_- = \Xi_+\Phi \quad . \tag{25}$$

Still further, if $N_-^+(\rho,\phi,\xi)$ is the downward surface radiance of S at (ρ,ϕ) in direction ξ in $\Xi_-(\rho,\phi)$, then the radiance may be represented as the upward surface radiance $N_+^+(\rho',\phi',\xi)$ where (ρ,ϕ) and (ρ',ϕ') are conjugate coordinates on S. Hence:

$$\begin{aligned} N_-^+(\rho,\phi,\xi) &= N_+^+(\rho',\phi',\xi) \\ &= N_+^+\big(\Phi(\rho,\phi),\xi\big) \end{aligned}$$

or more briefly:

$$N_-^+(S) = N_+^+(S)\Phi \quad , \tag{26}$$

for the one-sided surface S. Analogously,

$$N_+^-(S) = N_-^-(S)\Phi \quad . \tag{27}$$

Hence on a one-sided surface S, *downward surface* radiance and *upward field* radiances can be transformed away according to (26) and (27) into respective surface and field radiances of the *opposite* polarity. It is therefore sufficient at a point x on S to speak only of *upward surface* radiance $N_+^+(S)$ and *downward field* radiance $N_-^-(S)$. In particular, incident source radiance distributions therefore are limited to downward field radiances $N_-^o(S)$.

With these geometrical preliminaries in mind, we are now ready to apply the interaction principle to a general one-sided surface S' with conjugate mapping Φ. We first isolate S' and then enumerate the incident radiometric functions on S':

A_1: all field radiance distributions like $N_-^o(S')$

A_2: all field radiance distributions like $N_-(S')$

The sets of response functions of S' are:

B_1: all surface radiance distributions like $N_+(S')$

In the present case m = 2, n = 1, and the interaction principle yields the following two interaction operators:

$$s_{11} \text{ -- } r_-^o(S')$$

$$s_{12} \text{ -- } r_-(S')$$

The two operators $r_-^o(S')$ and $r_-(S')$ are *not* instances of definitions (10), (11) of Sec. 3.3, as has been the case all along in applications to two-sided surfaces. We shall see the specific form of these operators in a moment. For the present we go on to write the interaction equation for a one-sided surface S as:

$$N_+(S') = N_-^o(S')r_-^o(S') + N_-(S')\, r_-(S') \tag{28}$$

The auxiliary equation in the present case is :

$$N_-(S') = N_+(S')t_m(S') \tag{29}$$

where $t_m(S')$ has exactly the same task in the present one-sided case as it did in the two-sided case of Example 4. Thus (28) and (29) constitute the general radiative transfer equations for a one-sided surface. The formal solution of (28) and (29) is:

$$\boxed{N_+(S') = N_-^o(S')r_-(S')\left[I - t_m(S')r_-(S')\right]^{-1}} \tag{30}$$

The similarity of this solution with (11) above for the integrating sphere is particularly to be noted.

We next establish the respective connections between the operators $r_-^o(S')$, $r_-(S')$, and $t_m(S')$ for a one-sided surface, with their correspondents for a two-sided surface. This can most easily be done by making the present one-sided surface S' a two-sided surface S without changing its radiometric properties. Thus we make any radial cut in S', such as from x to y in Fig. 3.17. Then the equations of system (18) hold but we limit ϕ to the range $0 \leq \phi \leq 2\pi$. The resultant surface S is two-sided in the sense defined above. Clearly, its local and global radiometric properties are the same as those of S'. The salient difference between S and S' is that in the case of S there are no pairs of conjugate coordinates assigned to each point and so there is no mapping which conveniently rids us of $\Xi_-(x)$, etc. It follows that (14)-(17) now apply to S.

Comparing (28) and (14), we have the following functional relations between the reflectance and transmittance operators for S' and S:

$$r_-^o(S') = r_-^o(S) + \Phi t_+^o(S) \qquad (31)$$

$$r_-(S') = r_-(S) + \Phi t_+(S) \quad . \qquad (32)$$

Comparing (16) and (29) we have:

$$t_m(S') = \chi_+(S)t_m(S) + \chi_-(S)\Phi t_m(S)$$

$$= [\chi_+(S) + \chi_-(S)\Phi]t_m(S) \quad .$$

Since we have not changed the definition of $t_m(S)$ in going to the one-sided context, we require:

$$t_m(S') = t_m(S) \quad .$$

Thus in the one-sided context we should have

$$\chi_+(S) + \chi_-(S)\Phi = 1$$

which is indeed the case, by the definitions of $\chi_\pm(S)$ and Φ.

One final set of remarks may be in order concerning the role of Möbius Strips and general one-sided surfaces in radiative transfer matters. As noted in the introductory comments to this example, one-sided surfaces arose in the search for the domains of validity of the classical surface integral formulas. Surface integrals in radiative transfer theory arise, for example, in irradiance calculations as we saw in Examples 8-10 of Sec. 2.11. Hence caution must be exercised in using Stokes' Theorem, for example, in transferring from line to surface integrals when working with one-sided surfaces. Surface integrals also arise in calculations of net flux across surfaces S. Thus,

$$\int_S \mathbf{H}(x)\cdot\mathbf{k}(x)\, dA(x)$$

normally gives the net radiant flux across a two-sided surface S, where $\mathbf{H}(x)$ is the vector irradiance at x. This integral can have positive, negative, or zero values. Now as an indication of a radiometric pathology arising on a one-sided surface, let S be the Möbius Strip defined by (18) and in Fig. 3.17. If "x" denotes the point (r,ϕ,z) then:

$$\int_{\phi=0}^{4\pi}\int_{\rho=-a}^{a} \mathbf{H}(x)\cdot\mathbf{k}(x)\, dA(x) = 0$$

for every irradiance field $\mathbf{H}$ over S. In other words a Möbius Strip S, lets every light field "slip through its fingers"

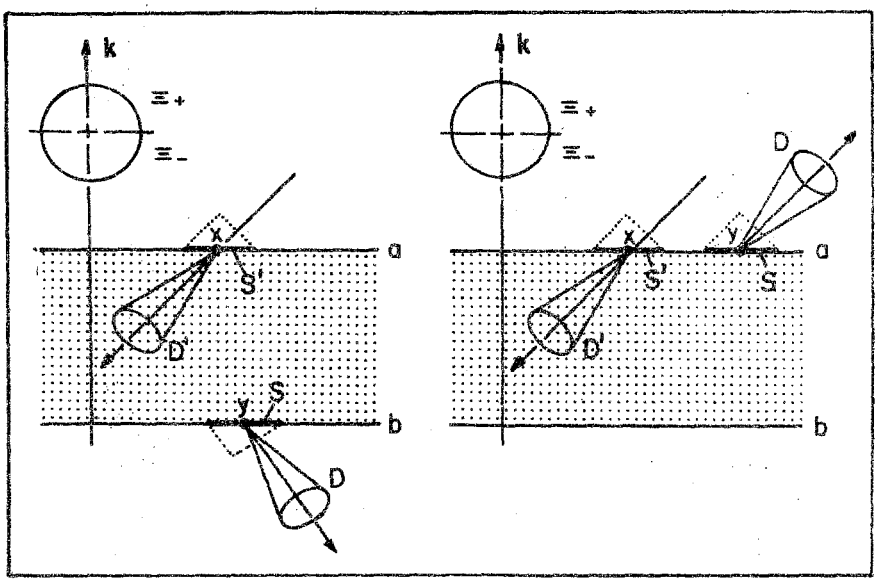


FIG. 3.18 Setting for empirical reflectances and transmittances of plane-parallel media.

when integrated over in the manner usual for two-sided surfaces. However, by slicing S in a manner designed to render it two-sided, as done above, the same integral considered above and taken over the same point set can have non zero net radiant flux $\bar{P}$ across its extent and such that $\bar{P}$ has valid physical significance.

3.6 Reflectance and Transmittance Operators for Plane-Parallel Media

In this section we continue the sequence of constructions, begun in Sec. 3.3, of the basic concepts used in radiative transfer theory. We shall use the interaction principle to develop the reflectance and transmittance operators for plane-parallel media which will subsequently be used (Sec. 3.7) in the formulation of interaction equations for such media. The development of such operators will be carried on within the space-level interpretation of the interaction principle (cf. Sec. 3.2). This is in contrast to the constructions leading to the r and t functions of Sec.3.3; they were carried out in the surface-level interpretation of the principle.

Geometrical Conventions

The geometrical conventions for plane-parallel media are depicted in Fig. 3.18. First of all a *plane-parallel optical medium* is a subset of euclidean space X consisting of

all points between and including two infinite parallel planes a and b. The plane-parallel medium is imagined to contain matter which can scatter, absorb or emit radiant flux. In a terrestrially-based reference frame (Sec. 2.4) planes a and b are customarily parallel to and coextensive with the xy coordinate plane, as shown in Fig. 3.18. The sets of downward Ξ_- and upward Ξ_+ directions are available for use in describing flows of radiant flux within the medium, and we shall use once more the narrow conical circular type of cones D' and D of Sec. 3.3 to establish the empirical reflectances and transmittances of plane-parallel media.

The Empirical Reflectances and Transmittances

Let planes a and b define a plane-parallel medium X, as in Fig. 3.18. The medium is irradiated in the vicinity of point x on one of its boundary planes. No other irradiation falls on X and no sources of radiant flux are within X. Fig. 3.18 shows x on plane a. The point x can also be on plane b. The empirical radiance of the incident flux at x is N(S',D') where S' is the projection on plane a and along the axis of D', of the collecting surface of a radiance meter, and D' is the small solid angle of directions of the incident flux. D' lies wholly within Ξ_- if x is on a, or wholly within Ξ_+ if x lies on b. For notational convenience we have used the regions S' and S lying in the boundary planes, instead of their projected counterparts on the planes normal to the axis of the radiance meter (shown dotted in Fig. 3.18). With this irradiation fixed, a radiance meter is directed at various points y on the boundary of X. Let N(S',D';S,D) be the resultant measured surface radiance of X emergent through direction set D in the vicinity of point y on the boundary of X. Let S be the small patch of surface on the boundary of X about y defined as the projection on plane a or b along the axis of D, of the collecting surface of the radiance meter. (Recall the convention for measuring surface radiance in Sec. 2.6.) The cone D lies completely in Ξ_+ or Ξ_- as the case may be. For a fixed plane-parallel medium X and every such pair S',D' of incident variables and every such pair S,D of response variables, let us write:

$$\text{"S(X;S',D';S,D)"} \quad \text{for} \quad \frac{N(S',D';S,D)}{N(S',D')A(S')\Omega(D')} \quad . \quad (1)$$

In this way we can generate a table of values S(X;S',D';S,D) of the function $S(\cdot;\cdot,\cdot;\cdot,\cdot)$, and this table of values would be representative of the reflecting or transmitting properties of the medium X. As it stands the value S(X;S',D';S,D) does not tell whether the scattered radiant flux has been reflected or transmitted by X. To help distinguish between these processes we keep in mind on which hemisphere of Ξ, the direction sets D' and D lie. If D' and D lie in the same hemisphere, then the value is a transmittance; if D' and D lie in opposite hemispheres, then the value is a reflectance. Thus part (a) of Fig. 3.18 depicts a transmittance arrangement, and part (b) depicts a reflectance arrangement.

By systematically considering the four general possibilities we arrive at the following set of definitions for a plane-parallel medium X defined by planes a and b. We write:

$$\text{"}R_+(X;S',D';S,D)\text{"} \quad \text{for} \quad S(X;S',D';S,D) \quad \text{if} \quad S' \subset b,\ D' \subset \Xi_+ \ \text{and}\ S \subset b,\ D \subset \Xi_- \tag{2}$$

$$\text{"}R_-(X;S',D';S,D)\text{"} \quad \text{for} \quad S(X;S',D';S,D) \quad \text{if} \quad S' \subset a,\ D' \subset \Xi_- \ \text{and}\ S \subset a,\ D \subset \Xi_+ \tag{3}$$

$$\text{"}T_+(X;S',D';S,D)\text{"} \quad \text{for} \quad S(X;S',D';S,D) \quad \text{if} \quad S' \subset b,\ D' \subset \Xi_+ \ \text{and}\ S \subset a,\ D \subset \Xi_+ \tag{4}$$

$$\text{"}T_-(X;S',D';S,D)\text{"} \quad \text{for} \quad S(X;S',D';S,D) \quad \text{if} \quad S' \subset a,\ D' \subset \Xi_- \ \text{and}\ S \subset b,\ D \subset \Xi_- \tag{5}$$

The preceding four definitions are designed to run parallel to (2)-(5) of Sec. 3.3. As in the earlier case, these definitions could have been based directly on the interaction principle with the same operators resulting. (The interested reader should consult the discussion following (5) of Sec. 3.3.) However, in this introductory discussion, we wish to keep the intuitive and operational aspects of the concepts foremost.

Once a table of reflectances and transmittances is made for a given plane-parallel medium, the table can be used to compute the responses to incident flux. Suppose, e.g., that X is irradiated over S' in a by radiance N(S',D') where the reflected radiance N(S,D) over S in a is sought. Then

$$N(S,D) = N(S',D')R_-(X;S',D';S,D)\Omega(D')A(S') \quad .$$

The essential properties of the empirical radiance N(S',D';S,D), which is the response to N(S',D') of the plane-parallel medium X, is its S' and D' additivities and its S' and D' continuities. These properties are analogous to (i) and (ii) of Sec. 3.1 enunciated for surfaces. These properties are sufficiently important to the theory of plane-parallel media (and media in general) that they will be stated below: In each case D,D' are circular conical solid angles with central directions ξ,ξ', respectively, etc.

(i) *(D'-additivity).* If a and b are the plane boundaries of a plane-parallel medium X and either a or b is irradiated in turn by radiances $N(S',D_1')$ and $N(S',D_2')$ with $N(S',D_1';S,D)$ and $N(S',D_2';S,D)$ as the respective observed response radiances, then $N(S',D_1';S,D) + N(S',D_2';S,D)$ is the observed radiance of S (on a or b) under simultaneous irradiation.

Furthermore:

(ii) *(D'-continuity).* Let the geometric setting be as defined in (i). If $\Omega(D') = 0$ and $\xi \neq \xi'$ then $N(S',D';S,D) = 0$.

A similar pair of statements can be made about S'-additivity and S'continuity by varying the areas over which the flux is incident, keeping the solid angles fixed. The reason behind all this attention to additivity and continuity rests in the fact that from these properties--which are simply intuitive manifestations of the linearity of radiative processes in the domain of radiative transfer theory--we form an empirical justification of the mathematical model of radiometry by means of which we can rigorously deduce the existence of the classical integral operators for the reflectances and transmittances of plane-parallel slabs of scattering material. The details of the derivation of the appropriate integral operators for the plane-parallel setting are quite close to those for the surface operators discussed in Sec. 3.3. In order to avoid excessive repetition, it will be shown in Sec. 3.16 how all these integral operators for surfaces, slabs and general media can be deduced from the interaction principle.

The Theoretical Reflectances and Transmittances

We now use the definitions (2)-(5) as a basis for the definition of the theoretical reflectances and transmittances of a plane-parallel medium. As was pointed out in the preceding paragraph, the existences of theoretical reflectance and transmittance functions and their associated integral operators all follow mechanically from the interaction principle and will be discussed later. However, it is instructive to show how these functions come about by means of every day limit operations applied to the empirical response functions. This we now do. In (2)-(5), we let $S' \to \{x'\}$, $S \to \{x\}$, $D' \to \{\xi'\}$, $D \to \{\xi\}$ in any order desired.

We choose first to write:

$$\text{"}S(X;S',D';x,\xi)\text{"} \quad \text{for} \quad \lim_{\substack{S\to\{x\} \\ D\to\{\xi\}}} S(X;S',D';S,D) . \qquad (6)$$

This limit exists by virtue of the S and D additivity and continuity properties of radiant flux established in Sec. 2.3. Next we write:

$$\text{"}S(X;x',\xi';x,\xi)\text{"} \quad \text{for} \quad \lim_{\substack{S'\to\{x'\}\\ D'\to\{\xi'\}}} S(X;x',\xi';x,\xi) \tag{7}$$

This limit exists by virtue of the S' and D' additivity and properties of the response radiance just cited above. One may envision the physical significance of the magnitude of limit $S(X;x',\xi';x,\xi)$ as follows: the magnitude represents the radiance of X at x in the direction ξ induced by a unit incident radiance on X at x' in the direction ξ'. The establishment of $S(X;x',\xi';x,\xi)$ in the two stages of limit operations (6) and (7) was prompted by a desire to facilitate the reader's study of related matters in Sec. 3.14 and in reference [251], in particular Chapter III of that reference.

We come now to the definition of the integral operators associated with plane-parallel media. These operators are the three-dimensional counterparts to $r_-(Y)$ and $t_-(Y)$ defined in (10), (11) of Sec. 3.3, and the theoretical counterparts to R_- and T_- in (2)-(5) above. In the present case we identify the plane-parallel medium by the two bounding planes. Whenever possible, we choose to use their depths also as identifying names, thus "a" denotes a plane at depth a, etc. However, occasionally, to avoid confusion, the full name "X_y" will be used for the plane at depth y, $a \leq y \leq b$. In what follows we shall, unless explicitly stated otherwise, always assume that $a \leq b$. We write:

$$\text{"}R(a,b)\text{"} \quad \text{for} \quad \int_{\Xi_-}\int_{X_a}[\,]S(X;x',\xi';x,\xi)\,dA(x')\,d\Omega(\xi') \tag{8}$$

$$\text{"}T(a,b)\text{"} \quad \text{for} \quad \int_{\Xi_-}\int_{X_a}[\,]S(X;x',\xi';x,\xi)\,dA(x')\,d\Omega(\xi') \tag{9}$$

if $x \epsilon a$, $\xi \epsilon \Xi_+$ for (8); and $x \epsilon b$, $\xi \epsilon \Xi_-$ for (9);

$$\text{"}R(b,a)\text{"} \quad \text{for} \quad \int_{\Xi_+}\int_{X_b}[\,]S(X;x',\xi';x,\xi)\,dA(x')\,d\Omega(\xi') \tag{10}$$

$$\text{"}T(b,a)\text{"} \quad \text{for} \quad \int_{\Xi_+}\int_{X_b}[\,]S(X;x',\xi';x,\xi)\,dA(x')\,d\Omega(\xi') \tag{11}$$

if $x \epsilon b$, $\xi \epsilon \Xi_-$ for (10); and $x \epsilon a$, $\xi \epsilon \Xi_+$ for (11).

By using "a" and "b" as explained above, we obviate the need for the signatures "+" and "-" on "R" or "T" as in (6)-(9) of Sec. 3.3 or (2)-(5) above. The operators R(a,b) and R(b,a) are called the *standard reflectance operators* associated with the plane-parallel medium determined by a,b. The operators T(a,b) and T(b,a) are called the *standard transmittance operators* associated with the plane-parallel medium determined by a,b.

The remaining matters requiring mention in this section

can be presented very much in the manner that the corresponding surface concepts were presented in Sec. 3.3 following (10), (11) of Sec. 3.3. In particular, to denote an incident radiance distribution on the upper plane a of a plane-parallel medium X we write "$N_-(a)$". Thus, $N_-(a)$ is a function which assigns to x' in a and ξ' in Ξ_- the radiance $N_-(x',\xi')$. The reflected radiance distribution initiated by $N_-(a)$ has its values given by:

$$\int_{\Xi_-}\int_{X_a} N_-(x',\xi')S(X;x',\xi';x,\xi)\, dA(x')\, d\Omega(\xi') \qquad (12)$$

(where x is in a and ξ is in Ξ_+), and which we denote simply as: "$N_-(a)R(a,b)$". In a similar way we define $N_-(a)T(a,b)$ as the transmitted radiance distribution emerging from level b and initiated by the downward incident distribution $N_-(a)$. Further, $N_+(b)R(b,a)$ and $N_+(b)T(b,a)$ are the reflected and transmitted radiance distributions initiated by the upward incident distribution $N_+(b)$ on the lower boundary b of X. When necessary (i.e., if ambiguity is to be avoided) we can append the signatures "+" and "-" as superscripts to "N" to denote surface and field radiances, respectively. However, it has been found by experience that the *surface* radiance concept by itself is for the most part adequate to describe without ambiguity theoretical discussions involving plane-parallel media, particularly those centering on the principles of invariance. The main functional relations for the operators (8)-(11) will be developed in Chapter 7.

Variations of the Basic Theme

The operators defined in (8)-(11) may be used as a basis for defining still further operators for radiometric concepts other than radiance. We illustrate this observation for the case of radiant emittance and irradiance.

Let X be a plane-parallel medium defined by planes a and b. Let $N_-(a)$ be an incident downward radiance distribution over boundary a. This gives rise to an irradiance function $H_-(a)$ over a. Thus the value of $H_-(a)$ at each x in a is given by:

$$\int_{\Xi_-} N_-(x,\xi')\xi'\cdot(-\mathbf{k})\, d\Omega(\xi') \quad .$$

The medium X responds to this incident flux with a radiance distribution $N_+(x,\cdot)$ at each point x of a, as given in (12). The associated value of radiant emittance $W_+(a)$ of a at x is:

$$\int_{\Xi_+} N_+(x,\xi)\xi\cdot\mathbf{k}\, d\Omega(\xi)$$

Therefore to every plane-parallel medium defined by two parallel planes a and b we can associate a reflectance for irradiance, namely:

$$\frac{1}{H_-(a)} \int_{\Xi_+} \left[\int_{\Xi_-} \int_{X_a} N_-(x;\xi') S(X;x;\xi';x,\xi)\, dA(x')\, d\Omega(\xi') \right] \xi \cdot k\, d\Omega(\xi) \tag{13}$$

We will rarely mix radiance and irradiance calculations in one discussion so that it will usually be possible to economize on notation and write "R(a,b)" for the reflectance (13). Therefore the numerical product $H_-(a)R(a,b)$ denotes the radiant emittance of X induced by $H_-(a)$ on X, where $H_-(a)$ in turn is associated with $N_-(a)$. Hence in the irradiance context "R(a,b)" will denote a pure number; in the radiance context "R(a,b)" will denote an integral operator. We can make three more definitions of reflectance and transmittance for $H_+(b)$ and $H_-(a)$: R(b,a), T(b,a), and T(a,b), respectively. The full details of all these definitions and the discussion of the properties of the R and T quantities are reserved for Chapter 8.

We could now go on and consider further variations on the theme such as the present plane-parallel counterpart to (18) of Sec. 3.3. However, the point that potential variations are possible seems well made by now, and we shall therefore go on to consider the applications of the interaction principle to plane-parallel media.

3.7 Applications to Plane-Parallel Media

The application of the interaction principle to plane-parallel media, which is the main theme of this section, will perhaps be most interesting from the following two points of view. First, the interaction equations for interacting plane-parallel media will be seen to be identical in form to those for interacting planes illustrated in Sec. 3.4. This point of similarity of the two types of interaction equations for ostensibly dissimilar radiative transfer contexts should encourage a closer examination of the classical modes of solution of the problems associated with these media with the purpose in mind of obtaining a unified means of solution for both types of settings. The natural mode of solution, as we shall see, is one candidate for such a unification.

The second interesting observation that can be made about the illustrations below concerns the ontogenetical foundations of radiative transfer theory, that is, the basic concepts on which the theory rests. Advanced students of radiative transfer theory know that the classical framework of the subject can be made to rest on the equation of transfer for radiance and on the principles of invariance. For some time there was a question as to the primacy of one or the other of these concepts; which was more fundamental: the equation of transfer or the principles of invariance? This question is naturally of interest to those who are concerned with the logical connections between these two tap roots of the subject. The principles of invariance have been developed and made a powerful tool of radiative transfer theory by Chandrasekhar

and his followers. The systematic use of the principles of invariance by Chandrasekhar's school has led to brilliant solutions of longstanding intractable problems in radiative transfer theory, problems which were insuperable using the standard approach to them by means of the equation of transfer alone. In view of this fact, one is led naturally to speculate on whether the principles of invariance incorporate a deeper and logically independent insight into radiative transfer phenomena than does their predecessor, the classical equation of transfer. This speculation was examined in an earlier work (Ref. [251]) with the purpose of resolving the question of the logical status of the principles of invariance within the theory of radiative transfer. It was found that the principles of invariance were logically deducible from the equation of transfer. Moreover, it was found that, by suitably enlarging the domain of valid applicability of the principles of invariance--that is, by capturing their abstract essence in a sufficiently general physical setting--the equation of transfer was in turn logically forthcoming from these more comprehensive principles. The net result was the establishment of the logical equivalence of the principles of invariance and the equation of transfer. By continuing the abstraction process of the principles of invariance still further, the interaction principle eventually was attained.

With this brief historical sketch in mind, the following illustrations may take on a deeper significance than would a mere enumeration of examples of the application of the interaction principle. Thus, the illustrations are intended to summarize a unification and an extension, by means of the interaction principle, of the classical framework of radiative transfer theory on plane-parallel media. In subsequent sections, the illustrations will be extended to cover ever wider applications of the interaction principle.

Example 1. Irradiances on Plane-Parallel Media

We consider a general plane-parallel medium X bounded by two distinct transparent parallel planes, a and b, as in Fig. 3.19. The medium consists of scattering-absorbing material with no reflecting surfaces on the boundary or on parallel planes within X, and is irradiated by external sources only at its upper and lower boundaries. These sources are radiance distributions $N_-(a)$ on a and $N_+(b)$ on b of fixed directional structure and are independent of location on a and b. That is, the directional structure (but not the size) of $N_-(a)$ is fixed and is independent of position in a. Similarly for b. Furthermore, we assume the medium to be stratified, i.e., its optical properties and light field are independent of position on each intermediate plane at depth y within X, $a \leq y \leq b$. Let $H(y,+)$ and $H(y,-)$ be the resultant irradiances at depth y (for notation, see Sec. 2.4). For brevity and uniformity of exposition throughout this chapter, we write "$H_+(y)$" for the irradiance $H(y,+)$, and "$H_-(y)$" for the irradiance $H(y,-)$. Similarly "$W_+(y)$", "$W_-(y)$" will denote the radiant emittance of plane y in the upward (+) and downward (-) directions. The incident radiance distribution $N_-(a)$ induces

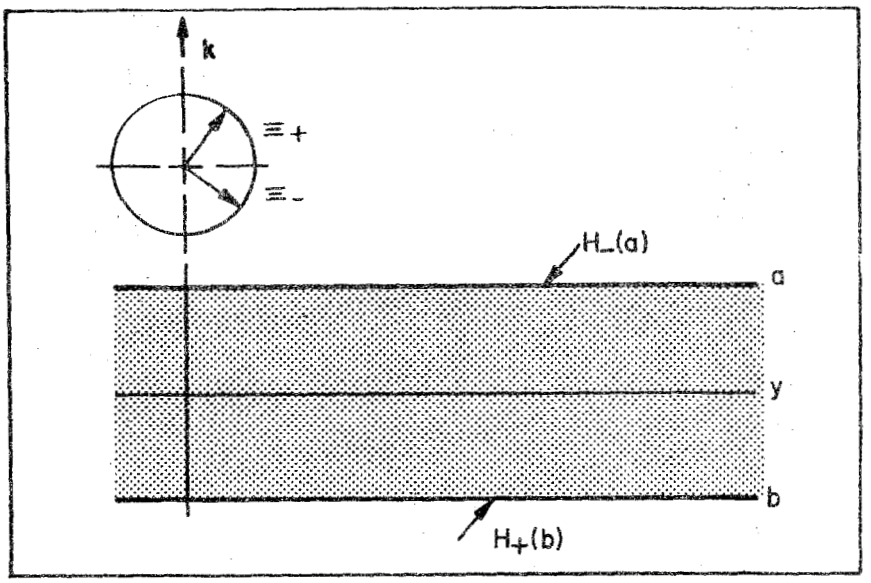


FIG. 3.19 The basic interaction setting for irradiances on plane-parallel media.

an irradiance which we denote by "$H_-(a)$". Similarly, $N_+(b)$ induces an irradiance $H_+(b)$. Our main purpose in this example is to show how the interaction principle can be used to systematically formulate the radiative transfer problem on X for irradiance: *given* $H_-(a)$, $H_+(b)$; *required*, the irradiances $H_\pm(y)$ at every depth y, $a \leq y \leq b$.

The application of the interaction principle is facilitated by having a convenient designation of the present optical medium X and its various plane-parallel subsets. Let us write "X(x,z)" for a plane-parallel medium bounded by planes x and z, $a \leq x \leq z \leq b$. We use, for convenience, "x", etc., now to denote both the plane and its depth in the medium. The present medium X is of the form X(a,b). As a first application of the interaction principle, we consider the subset X(a,y) of X(a,b), $a \leq y \leq b$. The enumeration of the sets of incident radiometric functions on X(a,y) in the present case is:

A_1: all irradiances like $H_-(a)$

A_2: all irradiances like $H_+(y)$

The enumeration of the sets of response functions of X(a,y) is:

B_1: all radiant emittances like $W_+(a)$

B_2: all radiant emittances like $W_-(y)$

Thus, in the case of medium X(a,y), m = 2, n = 2, and the interaction principle supplies four interaction operators s_{ij} (in the present case these are numbers) of the form:

$$s_{11} \text{ -- } R(a,y)$$

$$s_{12} \text{ -- } T(a,y)$$

$$s_{21} \text{ -- } T(y,a)$$

$$s_{22} \text{ -- } R(y,a)$$

The four numbers R(a,y), T(a,y), T(y,a), and R(y,a) can be represented, if need be, in terms of the S-function of Sec. 3.6 (see paragraph on *Variations of the Basic Theme*), and the discussion of that section shows how they can come by an alternate route from the interaction principle. For practical numerical work, one may integrate the Riccati equations obeyed by R and T as shown in Ref. [251] and in Chapters 7, 8 below. Hence these four numbers depend in a known manner on the depth y in X(a,b), we have then four functions R(a,·), T(a,·), T(·,a), R(·,a) associated with X(a,b) which take on specific values for each choice of subset X(a,y) of X(a,b). For the purposes of the present example, we assume specific knowledge of these four functions. In later discussions, throughout Chapter 8, we will show how the functions can be obtained.

According to the interaction principle, $W_+(a)$ and $W_-(y)$ are given by:

$$W_+(a) = H_-(a)R(a,y) + H_+(y)T(y,a) \qquad (1)$$

$$W_-(y) = H_-(a)T(a,y) + H_+(y)R(y,a) \qquad (2)$$

Hence if all six quantities on the right side are known, $W_+(a)$ and $W_-(y)$ are determinable. Of these six, five are known as given properties of X(a,y) or as given radiometric data. The remaining quantity, namely $H_+(y)$ is not generally known. This indicates that we should apply the interaction principle once again, now to the subset X(y,b) of X(a,b); $a \leq y \leq b$.

Isolating X(y,b), we then enumerate the sets of all incident radiometric functions on X(y,b):

A_1: all irradiances like $H_-(y)$

A_2: all irradiances like $H_+(b)$.

The enumeration of the response functions of X(y,b) is:

B_1: all radiant emittances like $W_+(y)$

B_2: all radiant emittances like $W_-(b)$.

The associated four interaction operators s_{ij} are:

$$s_{11} \text{ -- } R(y,b)$$

$$s_{12} \text{ -- } T(y,b)$$

$$s_{21} \text{ -- } T(b,y)$$

$$s_{22} \text{ -- } R(b,y)$$

The interaction principle then states that:

$$W_+(y) = H_+(b)T(b,y) + H_-(y)R(y,b) \tag{3}$$

$$W_-(b) = H_+(b)R(b,y) + H_-(y)T(y,b) \tag{4}$$

The auxiliary equations for the set (1)-(4) are:

$$W_+(y) = H_+(y) \tag{5}$$

$$W_-(y) = H_-(y) \tag{6}$$

which follow from the equality of field and surface radiance at a given point and in a given direction and the fact that no parallel planes within X(a,b) are reflecting planes (see statement (b) following (13) of Sec. 3.4, and see (32) of Sec. 2.5). With these two auxiliary equations, Eqs. (2) and (3) become autonomous:

$$\boxed{H_-(y) = H_-(a)T(a,y) + H_+(y)R(y,a)} \tag{7}$$

$$\boxed{H_+(y) = H_+(b)T(b,y) + H_-(y)R(y,b)} \tag{8}$$

These are the principles of invariance for irradiance in plane-parallel media. They will play an important role in the studies of Chapter 8. In essence, (7), (8) are two equations for $H_\pm(y)$ with solutions:

$$H_+(y) = \frac{H_+(b)T(b,y) + H_-(a)T(a,y)R(y,b)}{1 - R(y,a)R(y,b)} \tag{9}$$

$$H_-(y) = \frac{H_-(a)T(a,y) + H_+(b)T(b,y)R(y,a)}{1 - R(y,a)R(y,b)} \,. \tag{10}$$

From $H_\pm(y)$ we can determine $W_+(a)$ and $W_-(b)$ using (1) and (4).

We conclude this example with several general observations. First of all, the complete solution of $H_\pm(y)$ and $W_+(a)$, $W_-(b)$ is contingent on knowledge of the eight reflectances and transmittances associated with the subsets X(a,y) and X(y,b) of X(a,b). Methods of finding these numbers will be discussed in Chapter 8. Even without referring ahead to these methods, the following properties of these numbers can be brought to light. A reflectance such as R(a,y) is dependent not only on the material comprising X(a,y) but also the

directional structure* of $N_-(a)$ irradiating $X(a,y)$. This fact may be checked by studying (13) of Sec. 3.6. It turns out that $R(a,y)$ is not an inherent optical property of $X(a,y)$ but rather only an apparent optical property (cf., Sec. 9.1). It is for this reason that we hypothesized incident radiance distributions on $X(a,b)$ of *fixed* directional structure. The matter of R and T factors for plane-parallel media shall be discussed in detail in Chapter 8.

Secondly, we can see that the magnitudes of $R(a,y)$ and $R(y,b)$ must not exceed 1 if there is to be a determinate solution of $H_\pm(y)$. By appealing to the energy conservation law of general physics, suitably tailored to the radiative transfer context (Sec. 3.1), we can show that:

$$0 \leq T(a,b) \leq 1 \tag{11}$$

$$0 \leq R(a,b) \leq 1 \tag{12}$$

for every plane-parallel optical medium $X(a,b)$.

We can go further than (11), (12) and state that:

$$T(a,a) = 1 \tag{13}$$

$$R(a,a) = 0 \tag{14}$$

for every degenerate optical medium $X(a,a)$. This state of affairs is expected since the transmittance of a transparent plane should be 1 and its reflectance should be 0. These limiting values follow directly from (1), (2) and the auxiliary equations (5), (6), after setting $a = b$ and noting that (1) and (2) hold for every $H_-(a)$ and $H_+(b)$.

Example 2. Radiances in Plane-Parallel Media

As a second illustration of the interaction principle we consider once again the plane-parallel medium $X(a,b)$ of Example 1 but now with attention directed toward incident and response *radiance* distributions. The incident external radiance distributions on $X(a,b)$ and $N_-(a)$ and $N_+(b)$ are arbitrary. The stratification assumption is now dropped. We require the determination of radiance distributions $N_\pm(y)$ over any intermediate plane y in $X(a,b)$, as shown in Fig. 3.20.

We begin by partitioning $X(a,b)$ into two parts: $X(a,y)$ and $X(y,b)$, $a \leq y \leq b$. Isolating $X(a,y)$, and enumerating the sets of incident radiance distributions, we have:

---

*This is the case for the *irradiance* context and generally the radiometric quantities derived from radiance by various integral operations. The analogous operators $R(a,y)$, etc., considered below for the *radiance* context, are independent of the light field and are therefore inherent optical properties of $X(a,y)$.

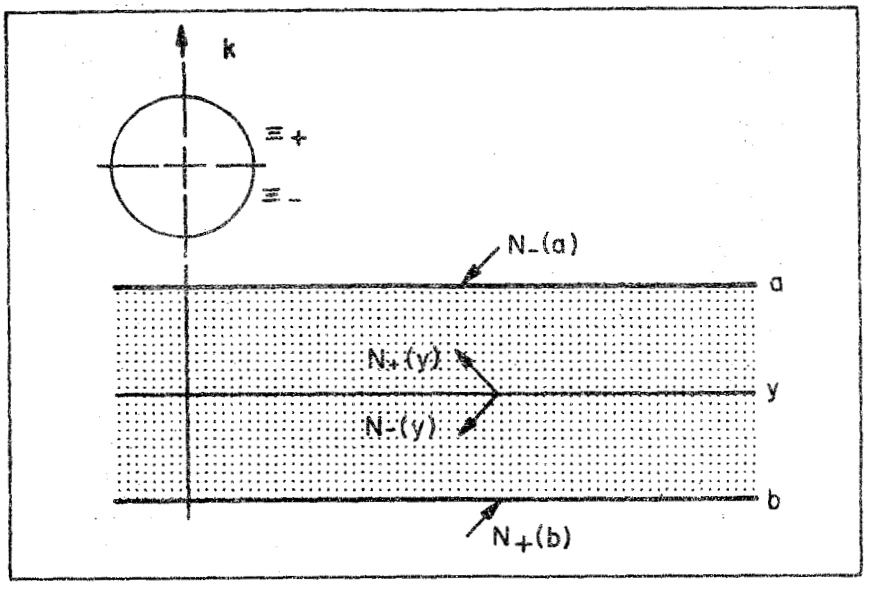


FIG. 3.20 The basic interaction setting for radiance distributions on plane-parallel media.

A_1: all field radiance distributions like $N_-^-(a)$

A_2: all field radiance distributions like $N_+^-(y)$

Enumerating the sets of response radiance distributions, we have:

B_1: all surface radiance distributions like $N_+^+(a)$

B_2: all surface radiance distributions like $N_-^+(y)$.

The four interaction operators s_{ij} are:

s_{11} -- R(a,y)

s_{12} -- T(a,y)

s_{21} -- T(y,a)

s_{22} -- R(y,a)

The four operators above are integral operators as defined in (8)-(11) of Sec. 3.6. Hence these operators ultimately come from the interaction principle in the sense that the existence of the kernel function $S(X;\cdot,\cdot;\cdot,\cdot)$ is guaranteed by the interaction principle for every optical medium X (see Sec. 3.16).

The interaction principle then states that:

$$N_+^+(a) = N_-^-(a)R(a,y) + N_+^-(y)T(y,a) \qquad (15)$$

$$N_-^+(y) = N_-^-(a)T(a,y) + N_+^-(y)R(y,a) \quad . \qquad (16)$$

By repeating this process of application of the interaction principle to medium X(y,b), we arrive at the analogous pair of statements:

$$N_+^+(y) = N_+^-(b)T(b,y) + N_-^-(y)R(y,b) \qquad (17)$$

$$N_-^+(b) = N_+^-(b)R(b,y) + N_-^-(y)T(y,b) \quad . \qquad (18)$$

When we append the following two auxiliary equations:

$$N_+^+(y) = N_+^-(y) \qquad (19)$$

$$N_-^+(y) = N_-^-(y) \qquad (20)$$

the set of equations (15)-(20) becomes autonomous. These auxiliary equations are specific instances of (32) of Sec. 2.5, and are a succinct way of ruling out the presence of any internal reflecting interfaces in the medium being analyzed. Thus, we see that, besides making the system (15)-(20) autonomous, the auxiliary equations allow us to write the system (15)-(18) in terms of surface radiance only, without possibility of ambiguity. It follows that $N_+^-(y)$ in (15) and (16) is equal to $N_+^+(y)$. $N_-^-(y)$ in (17) and in (18) is equal to $N_-^+(y)$. The incident radiance distributions $N_-^-(a)$ and $N_+^-(b)$ are from external sources and are immediately convertible to surface radiances using (32) of Sec. 2.5. Hence the surface radiance signature "+" may be dropped from the superscript position on "N". The set (15)-(18) thus becomes:

$$N_+(a) = N_-(a)R(a,y) + N_+(y)T(y,a) \qquad (21)$$

$$N_-(y) = N_-(a)T(a,y) + N_+(y)R(y,a) \qquad (22)$$

$$N_+(y) = N_+(b)T(b,y) + N_-(y)R(y,b) \qquad (23)$$

$$N_-(b) = N_+(b)R(b,y) + N_-(y)T(b,y) \qquad (24)$$

The middle two equations are autonomous. Their solutions are:

$$N_+(y) = \left[N_+(b)T(b,y) + N_-(a)T(a,y)R(y,b)\right]\left[I - R(y,a)R(y,b)\right]^{-1} \qquad (25)$$

$$N_-(y) = \left[N_-(a)T(a,y) + N_+(b)T(b,y)R(y,a)\right]\left[I - R(y,b)R(y,a)\right]^{-1} \qquad (26)$$

From (21) and (24) we can determine $N_+(a)$ and $N_-(b)$. The term $[I - R(y,a)R(y,b)]^{-1}$ is understood to be the inverse of the integral operator $[I - R(y,a)R(y,b)]$, and the term $R(y,a)R(y,b)$ is the iteration of $R(y,a)$ with $R(y,b)$. That is, with the help of the definitions (8) and (10) of Sec. 3.6, we have written:

"R(y,a)R(y,b)" for

$$\int_{\Xi_-}\int_{X_y}\left\{\int_{\Xi_+}\int_{X_y}[\;]S(X(y,a);x'',\xi'';x',\xi')\;dA(x'')\;d\Omega(\xi'')\right\}\times$$

$$\times\; S(X(y,b);x',\xi';x,\xi)\;dA(x')\;d\Omega(\xi') \qquad (27)$$

in exact analogy to the iteration $r_+(a)r_-(b)$ for surface integral operators in (67) of Sec. 3.4. A similar definition is made for $R(y,b)R(y,a)$. Therefore, in order for the inverse of $[I - R(y,a)R(y,b)]$ to exist, at least one of $R(y,a)$ or $R(y,b)$ must be norm contracting (re. (60) of Sec. 3.4). This condition is invariably found to hold in every natural (real) optical medium encountered in atmospheric and hydrologic optics. The theoretical details of iteration of (27) are covered in Sec. 3.4. Therefore if the norm contraction condition holds,

$$[I - R(y,a)R(y,b)]^{-1} = \sum_{j=0}^{\infty}[R(y,a)R(y,b)]^j \quad . \qquad (28)$$

A similar equality is obtained by interchanging "a" and "b". Thus $N_\pm(y)$ is uniquely determinable via (25), (26) using the natural mode of solution. The practical truncation process discussed in Example 7 of Sec. 3.4 (cf. in particular (88) of Sec. 3.4) holds for the present setting also.

It may be of interest to observe that only the inverse operation (28) need be computed in order to find both $N_+(y)$ and $N_-(y)$ by means of (25) and (26) when, say, $N_-(a) = 0$. This observation is based on the identity:

$$B[I - AB]^{-1} = [I - BA]^{-1}B$$

which holds for every pair of operators A and B such that $[I - AB]^{-1}$ exists. By means of this, the operator combination

$$R(y,a)[I - R(y,b)R(y,a)]^{-1}$$

in (26) can be written:

$$[I - R(y,a)R(y,b)]^{-1}R(y,a) \quad .$$

If, on the other hand, $N_+(b) = 0$, then only $[I-R(y,b)R(y,a)]^{-1}$ need be evaluated, for similar reasons.

We conclude by observing that the solutions $N_\pm(y)$ are predicated on knowledge of the operators associated with $X(a,y)$ and $X(y,b)$, in particular knowledge of the function $S(X;\cdot,\cdot;\cdot,\cdot)$ when X is $X(a,y)$ and $X(y,b)$. We shall consider some means of arriving at this knowledge in Chapter 7.

Finally, from the same considerations leading to (13) and (14), or formally from (25) and (26), or from Ia, Ib of Sec. 23, of Ref. [251], it is easy to deduce that the integral operators T(a,a), R(a,a) obtained by setting y = a satisfy the conditions:

$$T(a,a) = I \tag{29}$$

$$R(a,a) = 0 \tag{30}$$

where "I" and "0" now denote the identity and zero operators, respectively, of the operator algebra of Sec. 3.4.

Example 3. The Classical Principles of Invariance

We pause in our illustrations of the interaction principle to show how the four classical principles of invariance emerge by applying the interaction principle to a suitably dissected plane-parallel medium X(a,b) without internal sources of radiant flux. Figure 3.21 exhibits the requisite partitioning of X(a,b). We consider an arbitrary subset X(x,z), which in turn is partitioned into X(x,y) and X(y,z), $a \leq x \leq y \leq z \leq b$. Thus the geometric setting for the principles of invariance requires consideration of a partitioned internal slab X(x,z) arbitrarily located within X(a,b). This partitioning is of sufficient generality to subsequently allow functional relations to be written down for the four operators R(a,b), R(b,a), T(a,b), T(b,a) associated with a general plane-parallel medium X(a,b),(see Sec. 7.1).

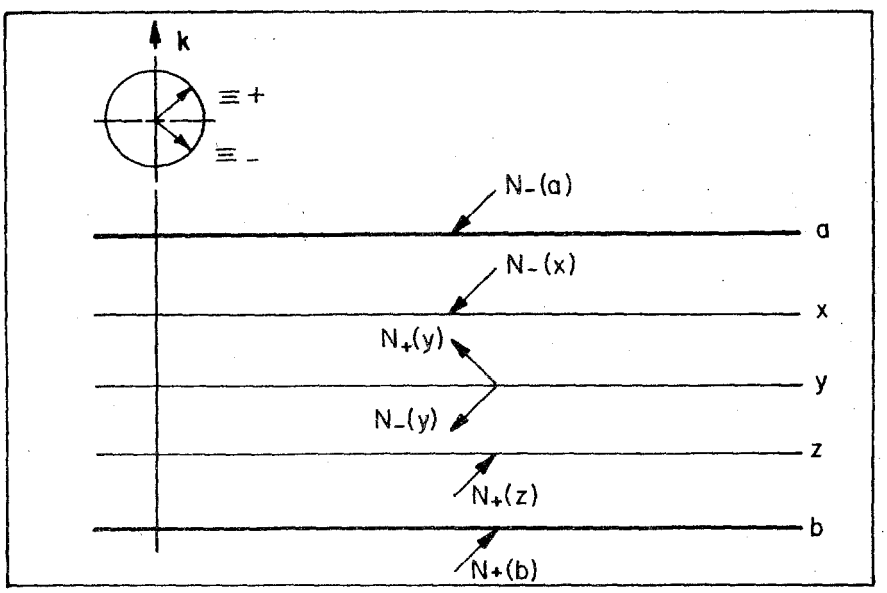


FIG. 3.21 The setting for the classical principles of invariance on plane-parallel media.

By repeating, line for line, the derivations of Example 2, now for X(x,z) (so that "a" is replaced by "x", and "b" by "z" and "y" is unchanged), we have the present counterparts to (23) and (22):

$$\text{I.}\quad N_+(y) = N_+(z)T(z,y) + N_-(y)R(y,z)$$

$$\text{II.}\quad N_-(y) = N_-(x)T(x,y) + N_+(y)R(y,x)$$

These are the two main principles of invariance I and II. Principles III and IV are obtained from them as follows. For III, use I twice: first let y = a, z = b; then let y = a, with z arbitrary:

$$\text{III.}\quad N_+(a) = N_+(b)T(b,a) + N_-(a)R(a,b)$$
$$= N_+(z)T(z,a) + N_-(a)R(a,z)$$

For IV, use II twice: first let y = b, x = a; then let y = b with x arbitrary:

$$\text{IV.}\quad N_-(b) = N_-(a)T(a,b) + N_+(b)R(b,a)$$
$$= N_-(x)T(x,b) + N_+(b)R(b,x)$$

Statements I-IV are the *principles of invariance* for a plane-parallel medium X(a,b). They are rules by which one can formulate the laws of radiative transfer on X(a,b), including the equation of transfer (cf. Sec. 25 and Sec. 126 of Ref. [251]).

The numbering of these principles is designed to facilitate their comparison with those in Sec. 50 of Ref. [43]. It should be noted that the present forms of the principles are written for generally inhomogeneous plane-parallel media so that *four* operators (rather than two as in [43]) are required for a complete determination of the light field in X(a,b). It should also be noted that the apparent simplicity and symmetry of I-IV above relative to their counterparts in [43] results from judicious use of operator concepts and also from leaving the light field in undecomposed form, i.e., the radiance distributions are not decomposed into reduced and diffuse flux (cf. Sec. 5.2). In this way the basic algebraic properties of the principles emerge and encourage formal manipulations such as those leading to (25), (26). (For the details of decomposition of light fields and their operators, see Sec. 7.1.) Of course, in the last analysis one must grapple with the realities of $S(X;\cdot,\cdot;\cdot,\cdot)$. However, one of the virtues of the interaction method resides precisely in its ability to defer such activity until the most propitious moment in a given analysis. In particular, one acquainted only

with the interaction principle and elementary algebra can analyze the most complex interaction problem to a stage typified by (25) and (26). From that stage onward, the natural mode of solution can be invoked for manual service, or service in automatic computer programs.

One further observation can be made which will facilitate the comparison of I-IV above with their classical counterparts in [43]. Suppose, as in [43], that $X(x,z)$ is homogeneous. Or, more generally, suppose that $X(x,z)$ and the light field within $X(x,z)$ are stratified, $a \leq x \leq z \leq b$. We further agree for the case where ξ' is in Ξ_- and ξ is in Ξ_+, to write:

$$\text{"}R(x,z;\xi';\xi)\text{"} \quad \text{for} \quad |k\cdot\xi| \int_{X_x} S(X;y',\xi';y,\xi)\, dA(y') \quad (31)$$

and for the case where ξ' is in Ξ_- and ξ is in Ξ_-, to write:

$$\text{"}T(x,z;\xi';\xi)\text{"} \quad \text{for} \quad |k\cdot\xi| \int_{X_x} S(X;y',\xi';y,\xi)\, dA(y') \quad (32)$$

where the domain of integration is over the upper plane boundary X_x of $X(=X(x,z))$ at depth x. The point y in (31) is in X_x; the point y in (32) is in X_z. Two more definitions can be made in a similar manner for upward reflectance and transmittance functions. However, if $X(x,z)$ is homogeneous, then it is easy to see that, under the present conditions, these functions are all non negative valued and depend spatially only on the difference $z-x$ of the depth parameters for $X(x,z)$. (The reason for this will be established in Sec. 7.1.) Hence to homogeneous $X(x,z)$ are associated two functions, the R and T functions defined above. The R-function in (31) is the present counterpart to the S-function in [43], and the T-function in (32) corresponds to the T-function in [43]. With the definitions (31) and (32) in mind we may represent the operators $R(a,b)$ and $T(a,b)$ (for downward incident flux) in (8), (9) of Sec. 3.6 as:

$$R(a,b) = \frac{1}{|k\cdot\xi|} \int_{\Xi_-} [\ ]R(a,b;\xi',\xi)\, d\Omega(\xi')$$

$$T(a,b) = \frac{1}{|k\cdot\xi|} \int_{\Xi_-} [\ ]T(a,b;\xi',\xi)\, d\Omega(\xi') \quad .$$

A similar pair of operators is associated with upward incident flux on $X(a,b)$.

Example 4: The Invariant Imbedding Relation

The solutions of the radiative transfer problem in plane-parallel media X(a,b) studied in Example 2 above were accomplished by using the interaction principle to set up the interaction equations for each of two subsets X(a,y) and X(y,b) of X(a,b). The resultant equations were then solved for the radiance distributions $N_{\pm}(y)$, $a \leq y \leq b$. In the present example we will apply the interaction principle directly to a subset X(x,z) of X(a,b) of Example 3 and ask it to give at once the interaction operators which yield $N_{\pm}(y)$ at some level y within X(x,z). The resultant operator equation is called the *invariant imbedding relation*, and employs the important concepts of the complete reflectance and transmittance operators.

We begin with the setting of Fig. 3.21. The subset X(x,z) is isolated, $a \leq x \leq z \leq b$. The sets of incident radiometric functions on X(x,z) are enumerated as:

A_1: all incident radiance distributions like $N_+(z)$

A_2: all incident radiance distributions like $N_-(x)$.

The sets of response functions on X(x,z) are those at level y, $x \leq y \leq z$:

B_1: all response radiance distributions like $N_+(y)$

B_2: all response radiance distributions like $N_-(y)$.

The interaction principle then asserts the existence of four interaction operators s_{ij}:

$$s_{11} \text{--} \mathcal{T}(z,y,x)$$

$$s_{12} \text{--} \mathcal{R}(z,y,x)$$

$$s_{21} \text{--} \mathcal{R}(x,y,z)$$

$$s_{22} \text{--} \mathcal{T}(x,y,z)$$

These four operators are not the simple integral operators of the kind in (8)-(11) of Sec. 3.6. Their structure will be considered shortly. For the present we go on to the assertion of the interaction equations in the present case. The interaction principle states that for $N_{\pm}(y)$:

$$N_+(y) = N_+(z)\mathcal{T}(z,y,x) + N_-(x)\mathcal{R}(x,y,z) \qquad (33)$$

$$N_-(y) = N_+(z)\mathcal{R}(z,y,x) + N_-(x)\mathcal{T}(x,y,z) \quad . \qquad (34)$$

This pair of equations can be written in matrix form. Let us first write:

$$\text{"}\mathcal{M}(x,y,z)\text{"} \quad \text{for} \quad \begin{bmatrix} \mathcal{T}(z,y,x) & \mathcal{R}(z,y,x) \\ \mathcal{R}(x,y,z) & \mathcal{T}(x,y,z) \end{bmatrix} \tag{35}$$

Then (33), (34) become:

$$\boxed{(N_+(y),N_-(y)) = (N_+(z),N_-(x))\,\mathcal{M}(x,y,z)} \tag{36}$$

Equation (36) is the *invariant imbedding relation*. Equation (36) is reminiscent of (10) in the preliminary example of Sec. 3.1 which resulted from considering the two surfaces S_1 and S_2 of that example as a radiometrically self-interacting entity. Such an interpolation may also be made in the case of (36); that is, we imagine $X(x,z)$ isolated with $N_+(z)$ and $N_-(x)$ as incident radiometric functions on $X(x,z)$. Then the internal radiance distribution $N(y)$, $(=(N_+(y),N_-(y))$ arises in response to this input by imagining the parts $X(x,y)$ and $X(y,z)$ to interact radiometrically. We denote the collection of all operators $\mathcal{M}(x,y,z)$ associated with $X(a,b)$, and with parameters x,y,z in the interval $[a,b]$, by "$\Gamma_3(a,b)$".

The four operators occurring in the invariant imbedding relation will now be related to the standard reflectance and transmittance operators (8)-(11) of Sec. 3.6. No loss in generality is engendered in setting $x = a$ and $z = b$ in (36), since $X(a,b)$ was initially arbitrary. The resultant equation is:

$$(N_+(y),N_-(y)) = (N_+(b),N_-(a))\,\mathcal{M}(a,y,b)\ . \tag{37}$$

In particular, we have:

$$N_+(y) = N_+(b)\,\mathcal{T}(b,y,a) + N_-(a)\,\mathcal{R}(a,y,b) \tag{38}$$

$$N_-(y) + N_+(b)\,\mathcal{R}(b,y,a) + N_-(a)\,\mathcal{T}(a,y,b)\ . \tag{39}$$

Now according to the interaction principle these $\mathcal{R}$ and $\mathcal{T}$ operators are unique, and yield the functions $N_\pm(y)$ corresponding to every member of the set of incident radiance distributions in A_1 and A_2. Thus, in particular by setting $N_+(b) = 0$ (the zero radiance distribution) in (38) and doing likewise in (25), we find that, since $N_-(a)$ is arbitrary:

$$\mathcal{R}(a,y,b) = T(a,y)R(y,b)\left[I - R(y,a)R(y,b)\right]^{-1}\ . \tag{40}$$

In a similar manner we find:

$$\mathcal{R}(b,y,a) = T(b,y)R(y,a)\left[I - R(y,b)R(y,a)\right]^{-1} \tag{41}$$

$$\mathcal{T}(a,y,b) = T(a,y)\left[I - R(y,b)R(y,a)\right]^{-1} \tag{42}$$

$$\mathcal{T}(b,y,a) = T(b,y)\left[I - R(y,a)R(y,b)\right]^{-1} \tag{43}$$

The operator $\mathscr{R}(a,y,b)$ is the *complete reflectance* operator and $\mathscr{T}(a,y,b)$ the *complete transmittance* operator. Equations (40)-(43) show how these new operators may be constructed from the standard operators of Sec. 3.6. The methods of Chapters 7 and 8 will show how $\mathscr{R}$ and $\mathscr{T}$ may be obtained directly by integration of the equation of transfer in two-flow form. We can now rewrite, if necessary, (40)-(43) for a general subset $X(x,z)$ of $X(a,b)$. From (29), (30) and (40)-(43) we deduce for every x,z, $a \leq x \leq z \leq b$; that:

$$\mathscr{T}(x,z,z) = T(x,z) \tag{44}$$

$$\mathscr{R}(x,x,z) = R(x,z) \tag{45}$$

$$\mathscr{T}(x,x,z) = I \tag{46}$$

$$\mathscr{R}(x,z,z) = 0 \tag{47}$$

The invariant imbedding relation (36) contains the four principles of invariance I-IV of Example 3 as special cases. Thus, let $x = y$, then (33), (44), (45) yield:

$$\text{I.} \quad N_+(y) = N_+(z)T(z,y) + N_-(y)R(y,z) \quad .$$

Further, let $z = y$, then (34), (44), (45) yield:

$$\text{II.} \quad N_-(y) = N_-(x)T(x,y) + N_+(y)R(y,x) \quad .$$

From these first two principles of invariance follow principles III and IV after the manner explained in Example 3.

Examples of the use of the invariant imbedding relation in extended computations in discrete space settings may be found in Sec. 70 of Ref. [251]. Further examples are given in Chapters 7 and 8 below.

We conclude this example with a few observations of historical interest. The invariant imbedding relation (36) was first given in Ref. [233] in an attempt to put into precise analytical form the verbal statement of the invariant imbedding principle of Bellman and Kalaba [13]. The latter principle was, in turn, an extension of the ideas of Ambarzumian [1], [2] and Chandrasekhar [42] centering around the classical forms of the principles of invariance. The work of Bellman and Kalaba was an important impetus to the eventual formulation of the invariant imbedding relation. This relation, in turn, motivated the *algebraic formulation* of the classical radiative transfer principles. This algebraic formulation evolved and eventually culminated in the interaction principle of Ref. [251], which is the foundation of the present work.

Example 5: Semigroup Properties of Transmitted and Reflected Radiant Flux

One of the most primitive of intuitions we have of light in natural optical media such as the atmosphere or the sea is that of its transmission from one point to another. In this example we show how the invariant imbedding relation yields a general analytical embodiment of this intuitive idea of transmitted light. In particular we shall derive the exact analytical expression of the following property of transmitted radiant flux: the amount of light transmitted over a path from point a to point b is equal to the amount first transmitted from a to an intermediate point y and thence transmitted from y to b. This type of property will eventually yield the volume attenuation function, one of the two main inherent optical properties used in the equation of transfer.

To derive the transmission property of radiant flux we return to the invariant imbedding relation (36) applied to an arbitrary level y in X(a,b) and with $N_+(b) = 0$ and $N_-(a)$ arbitrary:

$$\begin{aligned}\left(N_+(y),N_-(y)\right) &= \left(N_+(b),N_-(a)\right)\mathcal{M}(a,y,b)\\ &= \left(N_-(a)\,\mathcal{R}(a,y,b),N_-(a)\,\mathcal{T}(a,y,b)\right) \quad . \quad (48)\end{aligned}$$

Using (36) once again now for level b in X(y,b) with $N_+(b) = 0$:

$$\begin{aligned}\left(N_+(b),N_-(b)\right) &= \left(N_+(b),N_-(y)\right)\mathcal{M}(y,b,b)\\ &= \left(0,N_-(y)\right)\mathcal{M}(y,b,b)\\ &= \left(0,N_-(y)\right)\begin{bmatrix}\mathcal{T}(b,b,y) & \mathcal{R}(b,b,y)\\ \mathcal{R}(y,b,b) & \mathcal{T}(y,b,b)\end{bmatrix}\\ &= \left(0,N_-(y)\,\mathcal{T}(y,b,b)\right) \quad . \quad (49)\end{aligned}$$

From this, we have:

$$N_-(b) = N_-(y)\,\mathcal{T}(y,b,b) \quad .$$

From (48) we have:

$$N_-(y) = N_-(a)\,\mathcal{T}(a,y,b) \quad .$$

Hence

$$N_-(b) = N_-(a)\,\mathcal{T}(a,y,b)\,\mathcal{T}(y,b,b) \quad . \quad (50)$$

From (48) once more, now applied to level b in X(a,b):

$$N_-(b) = N_-(a)\,\mathcal{T}(a,b,b) \quad .$$

From this and (50), since $N_-(a)$ is arbitrary:

$$\mathcal{T}(a,b,b) = \mathcal{T}(a,y,b)\ \mathcal{T}(y,b,b) \tag{51}$$

which is the desired *semigroup property* of the complete transmittance operator. Observe that the last "b" in each transmittance operator is fixed and plays the role of a passive background parameter indicating the size of the medium in which the transfer takes place. When reading the equations, the attention of the reader should be directed to the first two parameters in each transmittance operator; then it will become clear that (51) indeed expresses in a very general form our basic intuition of transmitted radiant flux. Once the idea of the derivation is clear, the reader may derive a slight generalization of (51) wherein the subset $X(a,z)$ replaces $X(a,b)$ and y is an arbitrary level between a and z. The result is:

$$\boxed{\mathcal{T}(a,z,b) = \mathcal{T}(a,y,b)\ \mathcal{T}(y,z,b)} \quad . \tag{52}$$

Another semigroup relation similar to (51) but now for transmission from b to a is derivable from (36). This is left as still another exercise for the reader. A further relation is forthcoming from (36) which exhibits an interesting quasi-semigroup property for the complete reflectance operators:

$$\boxed{\mathcal{R}(a,z\ b) = \mathcal{T}(a,y,b)\ \mathcal{R}(y,z,b)} \tag{53}$$

for every level y in an arbitrary subspace $X(a,z)$ of $X(a,b)$.

The setting in which the semigroup properties for the complete operators $\mathcal{R}$ and $\mathcal{T}$ is best viewed is that of the generalized invariant imbedding relation which is considered in Examples 6 and 7 below. Furthermore, the full semigroup relations for members of the partial group $\Gamma_3(a,b)$ are developed (in the irradiance context) in Example 4 of Sec. 8.7. See in particular, (52)-(55) of Sec. 8.7.

Example 6: The Generalized Invariant Imbedding Relation

The generalized invariant imbedding relation, which we now consider, is the result of an attempt to increase the structural symmetry and comprehensiveness of (36) and such semigroup relations as (52) and (53). The setting for the present example is again that of Example 4: a general plane-parallel optical medium $X(a,b)$, with no internal sources of radiant flux, and irradiated only at its upper and lower boundaries X_a and X_b by $N_-(a)$ and $N_+(b)$, respectively. We shall work with radiance distributions $N_\pm(y)$ on arbitrary levels y in $X(a,b)$. Our goal in this example is the derivation of a generalized version of (36) which has greater analytic power and symmetry than (36). This will be bought,

however, at the expense of some intuitive value of the result. However, the sacrifice is soon lost sight of in the glare of the analytic and algebraic light shed by the operatorial degrees of freedom opened up by the extension.

We could begin the derivation of the generalized invariant imbedding relation by simply invoking the interaction principle with the appropriate choice of incident and response functions. However, the requisite choice of incident and response functions is not immediately intuitively clear, and a few motivational comments in this direction will now be made.

An examination of the invariant imbedding relation (36) shows that the matricial operator $\mathcal{M}(x,y,z)$ may be viewed as a single interaction operator mediating between the single incident radiometric function $(N_+(z),N_-(x))$ and the single response radiometric function $(N_+(y),N_-(y))$. This follows from the fact that the purview of the interaction principle need not be restricted to the run-of-the-mill kind of single radiometric functions. Indeed, its statement allows, for example, ordered sets of any finite (or infinite) number of radiance functions to play the role of a single incident radiometric quantity. Continuing to examine the invariant imbedding relation (36), we see that the planes x,y,z are customarily constrained to lie in the relations $x \leq y \leq z$ to each other. Furthermore, the response function $(N_+(y),N_-(y))$ is limited to a single plane y in $X(a,b)$.

Suppose we now relax the latter of the two conditions just cited. Then we would be supplied by the interaction principle with an operator s such that

$$(N_+(w),N_-(x)) = (N_+(u),N_-(v))s$$

where w and x are in the interval defined by $[v,u]$, $v \leq u$. We could go on to explore the properties of the four components s_{ij} of s in much the way we did those of $\mathcal{M}(x,y,z)$ in Examples 4 and 5. But we wish to go one step further and relax the remaining condition on the incident and response functions. We now do not even require w and x to lie in $[v,u]$, nor do we even require that $v \leq u$. The resultant operator arising from these relaxed conditions is the desired operator $\mathcal{M}(v,x;u,w)$ of the generalized invariant imbedding relation.

We now isolate $X(a,b)$ and for an arbitrary pair of depths u,v in $X(a,b)$ we enumerate the following sets of incident functions:

A_1: all incident radiance distributions like $N_+(u)$

A_2: all incident radiance distributions like $N_-(v)$.

Next we consider for an arbitrary pair of depths w,x in $X(a,b)$:

B_1: all response radiance distributions like $N_+(w)$

B_2: all response radiance distributions like $N_-(x)$.

Then the interaction principle asserts the existence of four unique interaction operators:

$$s_{11} \text{ -- } \mathcal{T}(u,w;v,x)$$

$$s_{12} \text{ -- } \mathcal{R}(u,x;v,w)$$

$$s_{21} \text{ -- } \mathcal{R}(v,w;u,x)$$

$$s_{22} \text{ -- } \mathcal{T}(v,x;u,w)$$

These operators depend generally on all four parameters, u,v, w,x. The choice of arrangement of the parameters in the symbols is guided by the resultant increased facility in establishing and reading the group properties of the operators presented below. For the present the individual parametric groupings simply may serve as long names for the operators. These operators are called the *extended reflectance* ($\mathcal{R}$) and *extended transmittance* ($\mathcal{T}$) operators. The interaction principle goes on to state that for $N_+(w)$, $N_-(x)$:

$$N_+(w) = N_+(u)\,\mathcal{T}(u,w;v,x) + N_-(v)\,\mathcal{R}(v,w;u,x) \qquad (54)$$

$$N_-(x) = N_+(u)\,\mathcal{R}(u,x;v,w) + N_-(v)\,\mathcal{T}(v,x;u,w)\ . \qquad (55)$$

This pair of equations can be written in matrix form. We first write:

$$\text{"}\mathcal{M}(v,x;u,w)\text{"} \quad \text{for} \quad \begin{bmatrix} \mathcal{T}(u,w;v,x) & \mathcal{R}(u,x;v,w) \\ \mathcal{R}(v,w;u,x) & \mathcal{T}(v,x;u,w) \end{bmatrix}. \qquad (56)$$

Then (54), (55) become:

$$\boxed{(N_+(w),N_-(x)) = (N_+(u),N_-(v))\,\mathcal{M}(v,x;u,w)}\ . \qquad (57)$$

Equation (57) is the *generalized invariant imbedding relation.* We denote by "$\Gamma_+(a,b)$" the collection of all operators $\mathcal{M}(v,x;u,w)$ with each variable u,v,w,x in the interval [a,b].

Before going on to deduce various consequences from (57) it may be well to add some explanatory comments on the structure of (57) in addition to those motivating its deduction from the interaction principle. The question that seems most likely to arise is this: if it was assumed that $N_-(a)$, $N_+(b)$ were incident radiometric quantities on X(a,b), and if X(a,b) was isolated just prior to the invocation of the interaction principle, why weren't the radiance distributions $N_-(a)$,$N_+(b)$ explicitly counted among the sets of incident quantities on X(a,b)? The answer to this question is that the choice of the sets of incident and response radiometric quantities on an isolated subset of an optical medium is quite arbitrary and subject only to the choice of the user of the interaction method. A re-reading of the interaction principle at this time may help make this answer clear. There is, in short, a precise logical basis for (57) in the statement of the principle, and (57) follows mechanically, so to speak from the principle under the present choice of the sets A and B.

The answer to the preceding question can be put into a more intuitive, less formal tone by means of the following observation: the light field within X(a,b) it is true, is generated *ab initio* and sustained by the hypothesized incident radiance distributions $N_-(a), N_+(b)$. Once generated and in the steady state, the light field within X(a,b) has a strong internal structural unity in the sense that the radiance distributions over any two separate planes of X(a,b) are closely and subtly interconnected one with the other; that is to say, the slightest change in the lighting over one plane is generally accompanied by a readjustment of the lighting in the other plane. Equation (57) is the formal expression of this intuitive insight into the internal unity of natural light fields. By our separating radiometric cause and effect in this extreme manner, the extended reflectance and transmittance operators of (57) have placed on them a relatively heavy burden to connect these distant radiometric causes and effects in one part of the light field with another. However, it is enough that the analytic connection--however complex or tenuous in reality--exists; for then a rich analytical harvest of results and techniques are available for use, especially those in the theory of continuous groups and semigroups, and which we shall state below and in subsequent examples.

The first deduction we wish to make from (57) is the invariant imbedding relation (30). The details of this deduction will add substance to the general comments above concerning the internal unity of the light field within X(a,b). It appears that for didactic purposes the deduction of the invariant imbedding relation from (57) is best made in reverse --that is we shall start from the invariant imbedding relation (36) and deduce (57). Then it will be observed that the path traversed from (36) to (57) is reversible. This mode of approach to (57) was the one actually followed in its discovery. Toward this end we use (36) to represent each of the radiance distributions occurring in (57):

$$N_+(w) = N_+(b)\,\mathcal{T}(b,w,a) + N_-(a)\,\mathcal{R}(a,w,b) \tag{58}$$

$$N_-(x) = N_-(a)\,\mathcal{T}(a,x,b) + N_+(b)\,\mathcal{R}(b,x,a) \tag{59}$$

$$N_+(u) = N_+(b)\,\mathcal{T}(b,u,a) + N_-(a)\,\mathcal{R}(a,u,b) \tag{60}$$

$$N_-(v) = N_-(a)\,\mathcal{T}(a,v,b) + N_+(b)\,\mathcal{R}(b,v,a) \quad . \tag{61}$$

Equations (54) and (58) are two ways of representing $N_+(w)$. Let us use (60) and (61) to replace $N_+(u)$ and $N_-(v)$ in (54) as follows:

$$\begin{aligned} N_+(w) = {} & N_-(a)\{\mathcal{T}(a,v,b)\,\mathcal{R}(v,w;u,x) + \mathcal{R}(a,u,b)\,\mathcal{T}(u,w;v,x)\} + \\ & + N_+(b)\{\mathcal{T}(b,u,a)\,\mathcal{T}(u,w;v,x) + \mathcal{R}(b,v,a)\,\mathcal{R}(v,w;u,x)\} \end{aligned} \tag{62}$$

Since $N_-(a)$ and $N_+(b)$ are arbitrary, (58) and (62) imply:

$$\mathcal{T}(a,v,b)\mathcal{R}(v,w;u,x) + \mathcal{R}(a,u,b)\mathcal{T}(u,w;v,x) = \mathcal{R}(a,w,b) \quad (63)$$

$$\mathcal{T}(b,u,a)\mathcal{T}(u,w;v,x) + \mathcal{R}(b,v,a)\mathcal{R}(v,w;u,x) = \mathcal{T}(b,w,a) \quad (64)$$

Next we observe that Equations (55) and (59) are two ways of representing $N_-(x)$. We use (60) and (61) to replace $N_+(u)$ and $N_-(v)$ in (55):

$$N_-(x) = N_-(a)\{\mathcal{R}(a,u,b)\mathcal{R}(u,x;v,w) + \mathcal{T}(a,v,b)\mathcal{T}(v,x;u,w)\}$$

$$+ N_+(b)\{\mathcal{T}(b,u,a)\mathcal{R}(u,x;v,w) + \mathcal{R}(b,v,a)\mathcal{T}(v,x;u,w)\} \quad .$$

Since $N_-(a)$ and $N_+(b)$ are arbitrary, this result and (59) imply:

$$\mathcal{R}(a,u,b)\mathcal{R}(u,x;v,w) + \mathcal{T}(a,v,b)\mathcal{T}(v,x;u,w) = \mathcal{T}(a,x,b) \quad (65)$$

$$\mathcal{T}(b,u,a)\mathcal{R}(u,x;v,w) + \mathcal{R}(b,v,a)\mathcal{T}(v,x;u,w) = \mathcal{R}(b,x,a) \quad . \quad (66)$$

The sets of equations (63),(64), and (65), (66) govern the extended reflectance and transmittance operators in terms of the complete reflectance and transmittance operators. We may view these equations in the present discussion as algebraic equations in the unknown extended operators with the known complete operators as "coefficients". This view is heuristic and will lead us correctly to results which can be established rigorously using advanced operator theory. Thus we are led to write (63) and (64) in matrix form:

$$\begin{bmatrix} \mathcal{T}(b,u,a) & \mathcal{R}(b,v,a) \\ \mathcal{R}(a,u,b) & \mathcal{T}(a,v,b) \end{bmatrix} \begin{pmatrix} \mathcal{T}(u,w;v,x) \\ \mathcal{R}(v,w;u,x) \end{pmatrix} = \begin{bmatrix} \mathcal{T}(b,w,a) \\ \mathcal{R}(a,w,b) \end{bmatrix} \quad . \quad (67)$$

Let us denote the 2×2 matrix of operators in (67) by "$\mathcal{H}(a,b;u,v)$". This matrix has an inverse provided $I - \mathcal{H}(a,b;u,v)$ is norm contracting, a condition which can be shown to generally hold in all natural optical media. Hence:

$$\mathcal{H}^{-1}(a,b;u,v) = [I - (I - \mathcal{H}(a,b;u,v))]^{-1} \quad ,$$

and by the norm contracting theorem (see, e.g., Sec. 40 of Ref. [251]):

$$\mathcal{H}^{-1}(a,b;u,v) = \sum_{j=0}^{\infty} (I - \mathcal{H}(a,b;u,v))^j \quad (68)$$

Therefore:

$$\boxed{\begin{pmatrix} \mathcal{T}(u,w;v,x) \\ \mathcal{R}(v,w;u,x) \end{pmatrix} = \mathcal{H}^{-1}(a,b;u,v) \begin{pmatrix} \mathcal{T}(b,w,a) \\ \mathcal{R}(a,w,b) \end{pmatrix}} \quad (69)$$

Next, we go on to write (65), and (66) in matrix form:

$$\begin{bmatrix} \mathcal{T}(a,v,b) & \mathcal{R}(a,u,b) \\ \mathcal{R}(b,v,a) & \mathcal{T}(b,u,a) \end{bmatrix} \begin{pmatrix} \mathcal{T}(v,x;u,w) \\ \mathcal{R}(u,x;v,w) \end{pmatrix} = \begin{pmatrix} \mathcal{T}(a,x,b) \\ \mathcal{R}(b,x,a) \end{pmatrix} \tag{70}$$

whence:

$$\boxed{\begin{pmatrix} \mathcal{T}(v,x;u,w) \\ \mathcal{R}(u,x;v,w) \end{pmatrix} = \eta^{-1}(b,a;v,u) \begin{pmatrix} \mathcal{T}(a,x,b) \\ \mathcal{R}(b,x,a) \end{pmatrix}} \tag{71}$$

Equations (69) and (71) express the four extended operators in terms of the complete operators in a manner analogous to that in (40)-(43) wherein the complete operators were represented in terms of the standard operators for subsets of X(a,b) of the form X(a,b), X(y,b). Hence we may conclude that: *in all natural optical media, the extended operators of the general invariant imbedding relation are ultimately representable in terms of the standard operators of the form (8)-(11) of Sec. 3.6.* An alternate proof of this conclusion, along with specific formulas establishing the asserted representations, is given in Sec. 7.4.

Some observations on the preceding results will now be made. One observation that is immediately forthcoming from (69) and (71) is the somewhat startling fact that $\mathcal{T}(u,w;v,x)$ and $\mathcal{R}(v,w;u,x)$ are independent of x, and $\mathcal{T}(v,x;u,w)$ and $\mathcal{R}(u,x;v,w)$ are independent of w. However, some reflection on the equations (54) and (55) and the choice of notation will show that there is no compelling reason why one *response* function should depend on another *response* function. Dependence of response functions on chosen incident functions must certainly be the case, but not necessarily on response functions. Therefore the right-end variables x and w in (69) and (71) are superfluous in the extended $\mathcal{R}$ and $\mathcal{T}$ operators in the sense just observed. The extended notation "$\mathcal{M}(v,x;u,w)$" with all four variables shown is still desirable for reasons which will become clear in the group theoretic discussions of Example 7 below. Hence the individual extended operators inherit an added loose variable which, like a human appendix, has meaning only when the entire domain of evolution of the operator $\mathcal{M}(v,x;u,w)$ and the radiometric activity over X(a,b) is considered. This freedom of choice of the right-end variables will be utilized in deriving special semigroup relations subsequently (cf., e.g., Ex. 4, Sec. 8.7).

The second observation is on the manner in which the extended operators reduce to the complete operators upon suitable confluence of the variables u,v,w,x. In this way we fulfill our obligation of showing that the invariant imbedding relation is a special case of (57). This can most readily be seen by returning to (63)-(66). For example, let w = x, a = v, and b = u in (63) with $a \leq w \leq b$. The result is:

$$\mathcal{T}(a,a,b)\mathcal{R}(a,w;b,w) + \mathcal{R}(a,b,b)\mathcal{T}(b,w;a,w) = \mathcal{R}(a,w,b)$$

From (44)-(47) we have:

$$\mathscr{R}(a,w;b,w) = \mathscr{Q}(a,w,b) \quad . \tag{72}$$

The three remaining relations are obtained similarly:

$$\mathscr{R}(b,w;a,w) = \mathscr{Q}(b,w,a) \tag{73}$$

$$\mathscr{T}(a,w;b,w) = \mathscr{T}(a,w,b) \tag{74}$$

$$\mathscr{T}(b,w;a,w) = \mathscr{T}(b,w,a) \quad . \tag{75}$$

Hence:

$$\mathscr{M}(a,w;b,w) = \mathscr{M}(a,w,b) = \begin{bmatrix} \mathscr{T}(b,w,a) & \mathscr{R}(b,w,a) \\ \mathscr{R}(a,w,b) & \mathscr{T}(a,w,b) \end{bmatrix} , \tag{76}$$

which is an instance of the operator (56). It is clear that by suitable choice of parameter values in (69) and (71) we can once again retrace our steps to the invariant imbedding relation. However, the route just taken is certainly equivalent and somewhat less arduous.

Finally, it should be observed that in general we need not restrict the parameters a,w,b in (76) or in (72)-(75) to have the orders $a \leq w \leq b$ or $b \leq w \leq a$. Thus it is possible to explore the properties of $\Gamma_3(a,b)$ under less restrictive conditions on the three parameters than given at the outset in (36). One such extension will be made in Sec. 7.4 as a matter of course. However, unless specifically noted otherwise, we shall work only with members of $\Gamma_3(a,b)$ whose parameters x,y,z are ordered either as $x \leq y \leq z$ or $z \leq y \leq x$. In the following example we shall introduce a new set (namely $\Gamma_2(a,b)$) of operators which will supply a powerful working tool free from any restrictions on the parameters of the operators.

A final word on the choice of notation for $\mathscr{M}(v,x;u,w)$ may be in order. A choice was made in (56) between the displayed order of variables and the alternative "$\mathscr{M}(u,w;v,x)$". The latter would look more natural in (57). However, our current choice works better in remembering the reductions (72)-(76), and was accordingly made with that in mind.

Example 7: Group-Theoretic Structure of Natural Light Fields

The interaction principle, via the generalized invariant imbedding relation of Example 6, leads to some interesting properties of natural light fields--group theoretic properties--which appear to offer not only some novel analytical means for the numerical determination of light fields in practice, but also some fundamental ways of formulating radiative transfer theory. We illustrate the basis of these new means in this example. The sense in which we use the term "group-theoretic" is best explained by going on directly to the derivations of these properties. Some further discussion will follow the derivations.

The group-theoretic structure will first be considered for a relatively easily visualized case--one which we shall return to in Chapter 7 (in particular Sec. 7.11) as a base for a new computational method of determining radiance distributions. For the present we are concerned only with the bare logical structure of that method. Our setting may be a general plane-parallel medium X(a,b)--that of Example 6 once again--or it may be a general one-parameter optical medium (ref.: Ex. 2, Sec. 3.9). We begin by setting w = x and u = v in (57). The result is:

$$\begin{aligned}\left(N_+(x'),N_-(x)\right) &= \left(N_+(u),N_-(u)\right)\mathcal{M}(u,x;u,x)\\ &= \left(N_+(u),N_-(u)\right)\mathcal{M}(u,x,u) \quad . \qquad (77)\end{aligned}$$

The latter equality follows formally from (76). For brevity, let us write:

$$\text{"N(y)"} \quad \text{for} \quad \left(N_+(y),N_-(y)\right)$$

for every level u in X(a,b) and *ad hoc*:

$$\text{"}\mathcal{M}(u,x)\text{"} \quad \text{for} \quad \mathcal{M}(u,x,u) \quad .$$

Then (77) becomes:

$$N(x) = N(u)\mathcal{M}(u,x) \qquad (78)$$

for every level x and u in X(a,b). Before going on, it must be pointed out that equation (78) has been obtained from (57) by a purely formal tactic--that is, the operator $\mathcal{M}(u,x,u)$ was obtained by a formal change of parameters in $\mathcal{M}(v,x;u,w)$, and it turns out that $\mathcal{M}(u,x,u)$ as we have agreed to use it (i.e., as a member of $\Gamma_3(a,b)$) is strictly not defined for $x \neq u$. This may be seen by recalling the usual domain of definition of the operator $\mathcal{M}(x,y,z)$ of the invariant imbedding relation in Example 4, wherein x,y,z are constrained to have the relations $x \leq y \leq z$. However, our goal at present is to draw certain necessary conclusions from (78), assuming it is possible to extend the domain of the invariant imbedding operator (35), and we shall now use the observations made at the close of Example 6 to justify this extension. The rewards for such a tactic are occasionally great and one interesting precedent for such tactics was in drawing certain necessary conclusions from the equation $x^2 + y = 0$, assuming it possible to extend the meaning of the equation to positive numbers y and values of x other than real numbers. The result, as is well known, was the theory of complex numbers. It is in such a heuristic spirit that we now proceed, using (78) as a premise.

Select any three levels x,y,z in X(a,b) and apply (78) to these levels in the following way:

$$\begin{aligned}N(y) &= N(x)\,\mathcal{M}(x,y)\\ N(z) &= N(y)\,\mathcal{M}(y,z)\\ N(z) &= N(x)\,\mathcal{M}(x,z)\end{aligned}$$

From the first two of these equations:

$$N(z) = \left[N(x)\,\mathcal{M}(x,y)\right]\mathcal{M}(y,z)$$
$$= N(x)\left[\mathcal{M}(x,y)\,\mathcal{M}(y,z)\right] \quad .$$

Comparing this with the third equation in the preceding group, and using the uniqueness of interaction operators, we have:

$$\boxed{\mathcal{M}(x,z) = \mathcal{M}(x,y)\,\mathcal{M}(y,z)} \quad . \qquad (79)$$

This is the *closure property* of the operators of the type $\mathcal{M}(u,v)$. Furthermore, uniqueness also yields the fact that: for every level w,x,y,z in $X(a,b)$:

$$\boxed{\left[\mathcal{M}(w,x)\,\mathcal{M}(x,y)\right]\mathcal{M}(y,z) = \mathcal{M}(w,x)\left[\mathcal{M}(x,y)\,\mathcal{M}(y,z)\right]} \qquad (80)$$

which is the *associativity* of the operators of the type $\mathcal{M}(u,v)$. In addition for every level x in $X(a,b)$:

$$\boxed{\mathcal{M}(x,x) = \begin{bmatrix} I_+ & 0 \\ 0 & I_- \end{bmatrix} = I} \quad . \qquad (81)$$

Equation (81) expresses the *identity property* and I is the *identity operator*, where $I_\pm$ are identity operators which act on upward (+) and downward (-) radiance distributions. Finally, for every level x,y in $X(a,b)$:

$$\boxed{\mathcal{M}(x,y)\,\mathcal{M}(y,x) = \mathcal{M}(x,x) = I} \quad . \qquad (82)$$

which is the *inverse property*. If we now denote by "$\Gamma_2(a,b)$" the set of all operators of the form $\mathcal{M}(x,y)$, $a \le x \le b$, $a \le y \le b$, we see that $\Gamma_2(a,b)$ forms a *partial group* in the sense that the closure, associativity, identity, and inverse properties hold. The product of elements $\mathcal{M}(w,x)$ and $\mathcal{M}(y,z)$ of $\Gamma_2(a,b)$ is defined whenever $x = y$ or $w = z$. (Mathematical readers will recognize $\Gamma_2(a,b)$ as an instance of a *local topological group* under suitable regularity conditions. See, e.g., [208]. Physicists will note the pertinence of the group-theoretic approach to the foundations of quantum mechanics. See [150] and problem IV of [251].)

The findings so far may be summarized as follows: by setting $w = x$ and $u = v$ in the generalized invariant imbedding relation we obtain an equation (78) which displays a formal extension $\mathcal{M}(u,x)$ of the invariant imbedding mapping $\mathcal{M}(x,y,z)$ and which shows that the set $\Gamma_2(a,b)$ of such *extended* operators has group structure. Its differential properties will

be explored in Sec. 7.5. These will lead to practical methods for numerical studies of $\Gamma_2(a,b)$.

The reader will now find it instructive to return to the generalized invariant imbedding relation and use the alternate definition:

$$\text{"}\mathcal{M}(u,x)\text{"} \quad \text{for} \quad \mathcal{M}(u,x;u,x)$$

for $\mathcal{M}(u,x)$ with no restrictions on u,z in the interval [a,b]. Alternatively, one may return to the interaction principle, using collections of radiance distributions like N(u) for the incident set A, and collections of radiance distributions like N(x) for the response set B. We choose to call the resultant interaction operator s_{11} so obtained, by the same name, "$\mathcal{M}(u,x)$"; there should be no confusion henceforth with the *ad hoc* definition of $\mathcal{M}(u,x)$ above, which has finished serving its introductory purposes. In either of these ways the reader can place on solid footing the heuristic procedure between (78) and (82). It is also readily seen using the relation (78) and the properties of the invariant imbedding relation that the new operators $\mathcal{M}(u,x)$ in $\Gamma_2(a,b)$ for X(a,b) have the following representation in terms of the invariant imbedding operators $\mathcal{M}(x,y,z)$:

$$\mathcal{M}(u,x) = \mathcal{M}^{-1}(a,u,b)\,\mathcal{M}(a,x,b) \tag{83}$$

for every u,x, $a \leq u \leq b$, $a \leq x \leq b$, provided $\mathcal{M}(a,u,b)$ or $I - \mathcal{M}(a,u,b)$ is norm contracting. (See (68) for the general case.) Equation (83) will be established as a matter of course in (40) of Sec. 7.4.

We now can make clear the opening statement of this example, to the effect that natural light fields have group-theoretic properties. From (78) we see that if we fix the parameter u, say let u = a, *then the radiance distribution* N(x) *at every level* x *in* X(a,b) *is associated with a unique interaction operator* $\mathcal{M}(a,x)$, *such that* $\mathcal{M}(a,x)$ *is a member of the partial group* $\Gamma_2(a,b)$. As a consequence of this, if N(x) and N(y) are radiance distributions at any two levels x and y in X(a,b), they are connected by the group products:

$$\mathcal{M}(x,a)\,\mathcal{M}(a,y) \quad (= \mathcal{M}(x,y))$$

or

$$\mathcal{M}(y,a)\,\mathcal{M}(a,x) \quad (= \mathcal{M}(y,x))$$

which are clearly group inverses of each other. The first of these acts on N(x) to yield N(y), the second acts on N(y) to yield N(x).

The possibility of such intimate group-theoretic interconnections between one part and another of a natural light field such as that summarized above stems from the fact that a given natural light field in a well-defined optical medium with non-pathological values of its optical constants or its geometric form, has a strong inner structural bond, so to speak, such that if the light field over a small region is known, the form of the light field in the remaining regions of

the medium can be inferred. This property was commented upon once before in the discussions of Example 6. Because of the importance of this property, we shall pause to discuss it in some more detail by means of two main examples, the first of which is as follows.

As an everyday example of what is meant by "strong inner structure" of some conceptual object, consider a sphere. To make the illustration take on practical overtones consider one of those silvered spheres seen reposing on concrete pedestals in certain gardens. Suppose that one such specimen is dropped and shattered. The gardener takes a remnant shard to a dealer in silvered garden spheres. The dealer is now confronted with the task of inferring from the shard the diameter of the sphere from which it came (so that the new sphere will sit on the old concrete pedestal). Now, such dealers are aware at least on an empirical level, of the "strong inner structure" of spheres in the form of their *constant* positive curvature over their extents. Hence if the curvature of the shard is estimated (and there are little tripod-like devices which are designed just for such tasks) this estimate is then the numerical reciprocal of the radius of the entire sphere.

Perhaps the preceding illustration will now serve to direct with new insight the reader's attention to the matter of verifying the intuition of the "strong inner structure" of natural light fields. Mathematical readers will recall that an outstanding example of such a property is possessed by analytic functions on open connected sets of the complex plane.

Further examples of strong inner structures in everyday light fields can be verified without excessive theoretical preliminaries: take the case of a finite sphere of uniform surface radiance in a vacuum. By measuring the normal irradiance at one point at a known distance from the center of the sphere it is possible to infer the normal irradiance produced by the sphere's radiant output at every other point in the space around the sphere (cf. Example 4, Sec. 2.11). Further, if the radiant intensity of the sphere is known, then by measuring the normal irradiance over a known interval of distance, however small, the reconstruction of the whole irradiance field is possible. The light field in this case is represented by an analytic function of a simple kind. The approximate but practically effective exponential law of decay of downward irradiance with depth in natural waters is still another basis for a group-theoretic property of natural light fields. We shall return to these ideas in Chapters 7 and 8. For the present we give one final illustration of a group-theoretic property of light fields.

The second of the two main illustrations of the group-theoretic structures of natural light fields to be given in this example will now be considered. Recall that "$\Gamma_4(a,b)$" denotes the set of all operators of the form $\mathscr{M}(v,x;u,w)$, with u,v,w,x arbitrary levels in $X(a,b)$. Then, as in the case of $\Gamma_2(a,b)$, the set $\Gamma_4(a,b)$ is a partial group in the sense that the closure property (79), the associativity property (80), the identity property (81), and the inverse property (82) can be verified to hold for $\Gamma_4(a,b)$. In particular, the closure

property is of the form:

$$\boxed{\mathcal{M}(v,z;u,y) = \mathcal{M}(v,x;u,w)\mathcal{M}(x,z;w,y)} \tag{84}$$

using standard matrix multiplication.

Here, at last, the full role of all four parameters defining the members of $\Gamma_4(a,b)$ can be seen. Notice, e.g., how the transition from v to z on the left is made on the right side of the equation in two steps: v to x and x to z, similarly for the transition from u to y. From (84), four operator equations are obtainable, namely the fully symmetrized fourth-order versions of (52) and (53), and the proper generalizations of (63)-(66). These are:

$$\mathcal{T}(u,y;v,z) = \mathcal{T}(u,w;v,x)\mathcal{T}(w,y;x,z) + \mathcal{R}(u,x;v,w)\mathcal{R}(x,y;w,z) \tag{85}$$

$$\mathcal{T}(v,z;u,y) = \mathcal{R}(v,w;u,x)\mathcal{R}(w,z;x,y) + \mathcal{T}(v,x;u,w)\mathcal{T}(x,z;w,y) \tag{86}$$

$$\mathcal{R}(u,z;v,y) = \mathcal{T}(u,w;v,x)\mathcal{R}(w,z;x,y) + \mathcal{R}(u,x;v,w)\mathcal{T}(x,z;w,y) \tag{87}$$

$$\mathcal{R}(v,y;u,z) = \mathcal{R}(v,w;u,x)\mathcal{T}(w,y;x,z) + \mathcal{T}(v,x;u,w)\mathcal{R}(x,y;w,z) \tag{88}$$

This set of equations or (84) will be called the *fourth-order semigroup relations*.

The preceding two main illustrations of the group theoretic structure of natural light fields constituting the present example will serve to show the novel directions in which the interaction principle can lead us. In particular we are led toward useful modes of representing and computing light fields in radiative transfer problems. The operators introduced in these illustrations will be explored further in Chapters 7 and 8.

We close this discussion on the group-theoretic structures of various sets of interaction operators with some observations on general partial groups of the form $\Gamma_3(a,b)$; i.e., the collection $\Gamma_3(a,b)$ of operators $\mathcal{M}(x,y,z)$ where $x \leq y \leq z$. First of all, $\Gamma_3(a,b)$ does not satisfy the group properties of the type (79)-(82). Under the standard definition of matrix product the closest we can come to closure--the first important group property required of $\Gamma_3(a,b)$ is:

$$\mathcal{M}(a,z,b) = \mathcal{M}(a,y,b)\mathcal{M}(y,z,y) \tag{89}$$

which clearly requires $y = z$ for the second operator on the right to be defined in the usual invariant imbedding relation. This special relation follows immediately from (84) and (76). However, despite the failure of $\Gamma_3(a,b)$ to have group structure (a defect adequately remedied by $\Gamma_2(a,b)$ and $\Gamma_4(a,b)$ introduced above) using standard matrix multiplication, the

component $\mathscr{R}$ and $\mathscr{T}$ operators of $\mathscr{M}(x,y,z)$ have the important and useful semigroup properties (52) and (53). Still further group-theoretic possibilities for interaction operators of $\Gamma_3(a,b)$ will be studied in Secs. 7.4 and 8.7. In particular, certain "non standard" products will be defined in order to obtain various desired group structures.

Finally, we observe two useful connections between the members of $\Gamma_3(a,b)$ and $\Gamma_4(a,b)$ which follow from (76) and (84). The first connection links two members of $\Gamma_3(a,b)$ via a member of $\Gamma_4(a,b)$ thus:

$$\mathscr{M}(a,y,b) = \mathscr{M}(a,x;b,z)\,\mathscr{M}(x,y,z) \tag{90}$$

where $a \leq x \leq y \leq z \leq b$. The second connection links two members of $\Gamma_4(a,b)$ via a member of $\Gamma_3(a,b)$ thus:

$$\mathscr{M}(a,z;b,y) = \mathscr{M}(a,x,b)\,\mathscr{M}(x,z;x,y) \tag{91}$$

where levels x,y and z are arbitrary within $X(a,b)$.

Group Theory, Radiative Transfer and Quantum Theory

In conclusion, it can hardly be overemphasized that the group-theoretic formulation of radiative transfer problems in the manner of this example, and as summarized in (79)-(82), gives rise to perhaps the most fundamental physical formulation possible at present. This formulation, when suitably generalized (see (102) of Sec. 7.4), begins to indicate a completely unified approach to radiative transfer theory foundations which can be made to rest on quantum mechanics. A study of Landé's formulations of quantum mechanics will illustrate the general manner of approach envisioned. See Chapter VI of [151] and in particular equation (5a), which can be written in matricial form as $\Psi(A,C) = \psi(A,B)\psi(B,C)$, and which may be compared with (79). It now appears possible that problems II and IV of Sec. 141 in Ref. [251] may not only be successfully solved, but also in an elegant, unified mathematical manner. The program would in outline be as follows: starting with Landé's quantum mechanical formulation, one derives $\Gamma_2(a,b)$ (and, for generality, also its coherent-flux form), and obtains (79)-(82). Then using the techniques to be developed in Sec. 7.4, one can construct the operators in $\Gamma_4(a,b)$. These operators yield all the operators of radiative transfer theory presently known, including $\mathscr{A}(x;a,b)$ as defined in (6) of Sec. 3.8 or in Sec. 25 of [251]. Using the steps of Sec. 3.15 below or those of Sec. 126 in [251], one arrives at the equation of transfer. Once the equation of transfer is obtained, then the gateway to the classical theory has been entered.

The approach to the foundations of radiative transfer just outlined is but one of the many possible approaches that

may be developed. The preceding approach is specifically designed to play up the deep group-theoretic similarities of the quantum and phenomenological levels of radiative transfer theory. Unquestionably, the simplest connection between quantum mechanics and radiative transfer theory--the connection that would require a minimum of re-doing of existing constructions, is that which would derive the interaction principle (Sec. 3.2) from the tenets of quantum mechanics with a specific representation of the interaction operator in terms of the quantum properties of matter. Then all the constructions of [251] and the present work would stand ready-made for use without any further effort. In this way one can go on to solve important remaining problems of radiative transfer theory (Sec. 141 of [251]) with a minimum of duplication of effort. For further observations on the similarity of the structures of radiative transfer equations and quantum dynamical equations, see the closing remarks of Sec. 8.2.

3.8 Interaction Operators for General Spaces

The third step in the present sequence of constructions of the main concepts of radiative transfer theory will be taken in this section. We shall develop the concept of the interaction operator for a general three-dimensional optical medium. This development therefore augments the store of operators constructed in Sec. 3.3 for surfaces and in Sec. 3.6 for plane-parallel media. As in the latter case, the present constructions will utilize the space-level interpretation of the interaction principle.

The subsets of Euclidean three-space we shall consider in this section are those that are connected--i.e., in one piece--and we may envision them as members of an ensemble of interacting connected sets. The interaction equations will be stateable for the ensemble once the interaction operator for each connected component of the ensemble is known. The connected spaces we consider may be of finite or infinite extent and fall conveniently for radiometric purposes into two main classes: Those that have convex surfaces and those that have non convex (concave) surfaces.

Geometrical Conventions

Figure 3.22 depicts a general connected optical medium X and a point x on the boundary Y of X. Let "$\mathbf{k}(x)$" denote the unit outward normal to Y at x. Then "$\Xi_+(x)$" will denote, as usual, the set of all directions in Ξ such that $\xi\cdot\mathbf{k}(x)>0$, and "$\Xi_-(x)$" will denote the set of all directions ξ in Ξ such that $\xi\cdot\mathbf{k}(x)<0$. The directions in $\Xi_\pm(x)$ are called the *outward* (+) or *inward* (-) directions at x. Radiance distributions $N(x,\cdot)$ at points x of the boundary Y are split, as usual, into two parts: the *outward radiance distribution* $N_+(x,\cdot)$ and the *inward radiance distribution* $N_-(x,\cdot)$. If "a" denotes a part of Y, then "$N_+(a)$" and "$N_-(a)$" denote the outward and inward radiance distributions of Y restricted to part a. The part a can vary from a set {x} consisting of one point x of Y, up to

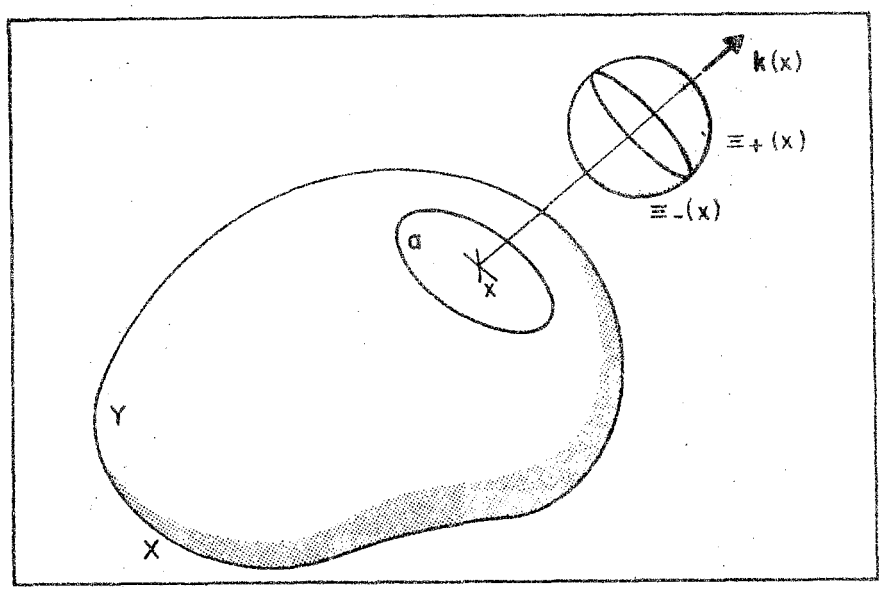


FIG. 3.22 Direction convention for interaction operators on general spaces.

Y itself.

The geometrical conventions for the empirical quantities D',D,S',S, established in Sec. 3.3, will also hold below.

Another geometric convention we shall require is that based on the process of *convexification* of a concave optical medium. This process will allow in many instances both convex and concave media to be treated alike during a given discussion. Let any of parts (a), (b), (c) of Fig. 3.23 represent an optical medium which has a concave boundary. This, it will be recalled, means that some points of the boundary can be joined by straight lines lying partially outside the surface Y. To be specific, we have pictured *solid* subsets X for the present discussion. It should be noted that all that we say below can be applied, *mutatis mutandis* to surfaces also. Now imagine a rubber sheet to be neatly applied all around X, enclosing X like a tight-fitting cocoon. On those parts of X where its surface is convex, the rubber sheet will cling and follow the contours of the original surface. On those parts of X where the surface is concave, the rubber sheet will soar as a plane surface across the concave hollow and will thereby establish a smooth convex surface enclosing X, of minimal possible area. Thus the step-like concavity of X in (b) of Fig. 3.23 will be ideally bridged by the rubber coating as sketched by the dashed lines in the figure, and the hollows and holes of (a) and (b) of the figure will be enclosed likewise. The net result will be a new region X' containing X with the rubber sheet as a *convex* boundary of the newly encased volume X'.

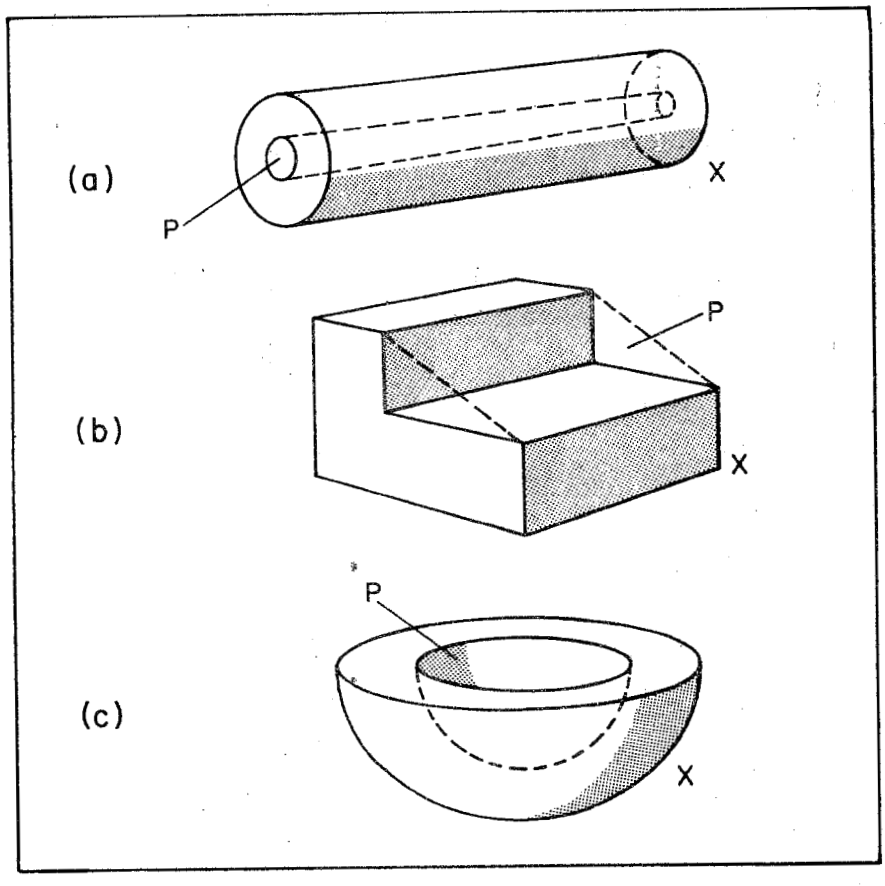


FIG. 3.23 Illustrating the convexification of concave media.

The new surface so formed is called the *convex hull* of X. In short, X' is the smallest convex solid containing X. So far we have engaged in pure geometry.

Next we introduce a radiometric element into the discussion. We consider all the regions which comprise the difference P between X and its convex hull X', including any holes inside X. For example, the triangular prism region P in Fig. 3.23 is one such region, and the hole in (a) and the hemisphere in (b) are further examples of the difference P between X' and X. It is found that certain theoretical considerations of X are facilitated by considering all such regions like P as filled either (a) by a hypothetical vacuum of unit transmittance and zero reflectance, or (b) by its antithesis: a hypothetical black material of zero transmittance and zero reflectance. In the case (a), we use X' and say that X has been *white convexified* and in case (b) we use X' and say that

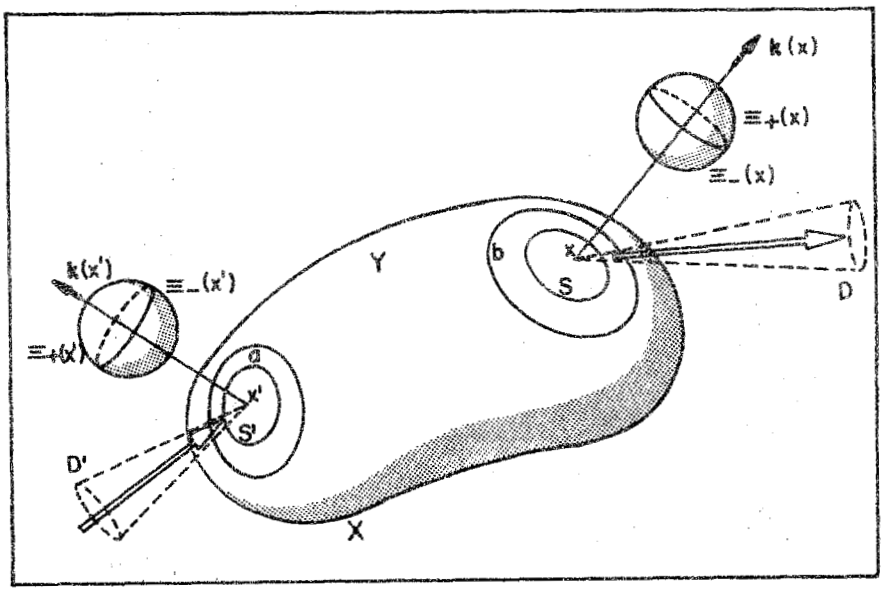


FIG. 3.24 Details for defining empirical scattering functions on arbitrary optical media.

X has been *black convexified*. It is obvious that if X is convex to begin with (and hence also with no holes), then either its black or white convexification results in X once again.

It is perhaps needless to add that a convexified X (either way) is still a *conceptual* object which can be considered irradiated or probed at will at any point of its surface or interior. However, the definitions of convexified media have an operationally meaningful cast which, if the necessity ever arose, could quite possibly be realized in many instances.

The Empirical Scattering Functions

The empirical scattering functions will now be established for a general optical medium X. The medium X may be convex or concave. If X has a concave boundary Y then we shall consider X to have been either white or black convexified. The present discussion is independent of the particular choice of these convexifications and hence we need not distinguish between them.

Consider two parts a and b of the boundary Y. Let S' be a small patch of part a around point x', and S be a small patch of b around point x, as in Fig. 3.24. Thus the present geometric situation is similar--as far as the present general geometry will allow--to Fig. 3.18. Let an amount N(S',D') of radiance be incident over S' and within the narrow conical solid angle D' which lies wholly inside $\Xi_-(x')$. This is the only source of irradiation either in or on X. (Again "S'",

as in Sec. 3.6, should be replaced in "N(S',D')" by the name of the projection of S' on a plane normal to the axis of D'. However, brief notation wins out over logical notation once again.) Let N(S',D';S,D) be the resultant radiance of S within the conical solid angle D which lies wholly inside $\Xi_+(x)$. Then let us write:

$$\text{"S(X;S',D;S,D)"} \quad \text{for} \quad \frac{N(S',D';S,D)}{N(S',D')A(S')\Omega(D')} \tag{1}$$

The non negative valued function $S(X;\cdot,\cdot,\cdot,\cdot)$ is the *standard* (empirical) *scattering function* for X.

Occasionally it is convenient to know if X has been white or black convexified, and when it is necessary to explicitly note this fact in the symbol for the standard empirical scattering function we shall write:

$$\text{"}S_w(X;S',D';S,D)\text{"} \quad \text{for} \quad S(X;S',D';S,D) \tag{2}$$

if X has been white convexified and:

$$\text{"}S_b(X;S',D';S,D)\text{"} \quad \text{for} \quad S(X;S',D';S,D) \tag{3}$$

if X has been black convexified.

At about this point in the corresponding developments of Secs. 3.3 and 3.6, it was customary to observe that the counterparts to $S(X;\cdot,\cdot,\cdot,\cdot)$ obeyed D and S additivity and continuity properties. The observance of this procedure is now well established and, therefore, in order not to repeat unnecessarily, these facts need only be alluded to here with an observation that these properties are stated in detail in Sec. 18 of Ref. [251]. Of course, while we are currently giving slight attention to these properties, this does not in any way mitigate their supreme importance in allowing the rigorous deduction from the interaction principle of the standard $\mathscr{S}$ and $\mathscr{K}$ operators below, and hence, ultimately, all of radiative transfer theory on discrete or continuous optical media. At any rate, the formal establishment of all these functions in 3.3, 3.6 and the present function, starting from the interaction principle, will be discussed in detail in Sec. 3.10. In particular, it will be shown in that section that each of the various S'-additivity and D'-additivity properties will take its formal place as an appropriate property of the interaction measure, and the various D' and S' continuity properties will be formulated as the so-called *AC property* of the interaction measure.

The Theoretical Scattering Functions

Let us write:

$$\text{"S(X;S',D';x,}\xi\text{)"} \quad \text{for} \quad \lim_{\substack{S\to\{x\} \\ D\to\{\xi\}}} S(X;S',D';S,D) \tag{4}$$

and

$$\text{"}S(X;x',\xi';x,\xi)\text{"} \quad \text{for} \quad \lim_{\substack{S'\to\{x'\}\\ D'\to\{\xi'\}}} S(X;S',D';x,\xi) \ . \tag{5}$$

These limits exist by virtue of the various D and S additivity and continuity properties of the empirical scattering function. If black or white convexification is to be explicitly noted, then "b" and "w" subscripts are inherited, appropriately, from (2),(3). We go on to write:

$$\text{"}\mathscr{S}(X;a,b)\text{"} \quad \text{for} \quad \int_a \int_{\Xi_-(x')} [\ ]S_b(X;x',\xi';x,\xi)\ d\Omega(\xi')dA(x') \tag{6}$$

where a and b are parts of the original surface Y of X and x is in b, and ξ is in $\Xi_+(x)$. Further, we write:

$$\text{"}\mathscr{C}(X;a,b)\text{"} \quad \text{for} \quad \int_a \int_{\Xi_-(x')} [\ ]S_w(X;x',\xi';x,\xi)\ d\Omega(\xi')dA(x') \tag{7}$$

where a and b are parts of the (original) surface Y of X and x is in b and ξ is in $\Xi_+(x)$. $\mathscr{S}(X;a,b)$ (or "$\mathscr{S}$" for short when X,a,b are understood) is the *standard $\mathscr{S}$-operator for* X *over* a *and* b. $\mathscr{C}(X;a,b)$ (or "$\mathscr{C}$" for short) is the *standard $\mathscr{C}$-operator for* X *over* a *and* b.

To see the relative roles played by $\mathscr{S}$ and $\mathscr{C}$ we observe that $\mathscr{S}$ is to a black convexified X as $\mathscr{C}$ is to a white convexified X. The theoretical connections between $\mathscr{S}$ and $\mathscr{C}$ for a given concave space X have been given in Sec. 25 of Ref. [Z51]. It suffices to say that this connection is intricate and its applications have not yet been completely explored. Of the two, the standard $\mathscr{S}$-operator is by far the more useful in the immediate generalizations of classical radiative transfer theory, especially in the theory of one-parameter carrier and general spaces (Examples 4,5 in Sec. 3.9). The operators $\mathscr{S}$ promise to help organize and systematize the theory on the more general spaces which have little or no symmetry or regular structure.

It will be instructive for the reader to give simple verbal proofs, based on the appropriate definitions, of the following statements:

(a) For every X, a, b, if X is convex and a and b are parts of the boundary of X, then $\mathscr{S}(X;a,b) = \mathscr{C}(X;a,b)$.

(b) For every X,a,b, if X is concave and a and b are parts of the boundary of X, then $|N_-(a)\ \mathscr{S}(X;a,b)| \le |N_-(a)\ \mathscr{C}(X;a,b)|$.

In statement (b) above, we have used the definition of radiometric norm (Example 5, Sec. 3.4) extended to curved surfaces

Y. Thus, we write in general:

$$\text{"}|N_{\pm}(Y)|\text{"} \quad \text{for} \quad \frac{1}{A(Y)} \int_Y \int_{\Xi_{\pm}(x)} N(x,\xi)\, d\Omega(\xi)\, dA(x). \qquad (8)$$

Furthermore, we have written in statement (b):

$$\text{"}N_-(a)\mathscr{S}(X;a,b)\text{"} \quad \text{for}$$

$$\int_a \int_{\Xi_-(x')} N_-(x',\xi')S_b(X;x',\xi';x,\xi)\, d\Omega(\xi')dA(x') \quad .$$

A similar definition holds for the term $N_-(a)\mathscr{L}(X;a,b)$.

Since we have defined the radiometric norm for radiance distributions over surfaces Y bounding general optical media X, it is natural to try to extend the definition of the norm of a reflectance operator, as given in Sec. 3.4, to a more general object such as the $\mathscr{S}$-operator for a medium X. The requisite sequence of definitions for the norm of $\mathscr{S}(X;a,b)$ is patterned closely after (44)-(49) of Sec. 3.4, and proceeds as follows. First we agree that if X has a boundary Y of finite area A(Y) then we normalize all radiometric norms of radiance distributions $N_{\pm}(a)$, defined over parts a of Y, with respect to A(Y) rather than with respect to A(a). Thus on a fixed finite boundary surface Y of an optical medium X we agree to write:

$$\text{"}|N_{\pm}(a)|_Y\text{"} \quad \text{for} \quad \frac{1}{A(Y)} \int_a \int_{\Xi_{\pm}(x)} N(x,\xi)\, d\Omega(\xi)\, dA(x) \quad . \qquad (9)$$

If A(Y) if infinite, then, as in Sec. 3.4, we employ a limit process to define the norm. In practice, when working with a fixed boundary Y, then "Y" may be dropped from the norm notation, for brevity.

Next we write:

$$\text{"}S_b(X;x',\xi';x)\text{"} \quad \text{for} \quad \int_{\Xi_+(x)} S_b(X;x',\xi';x,\xi)\, d\Omega(\xi) \qquad (10)$$

where x' and x are in Y and ξ' is in $\Xi_-(x')$. We have chosen to work with S_b simply to be specific. All that follows below holds also for S_w. Further, we agree to write:

$$\text{"}\beta(X,N;x',x)\text{"} \quad \text{for} \quad \frac{\displaystyle\int_{\Xi_-(x')} N(x',\xi')S_b(X;x',\xi';x)\, d\Omega(\xi')}{\displaystyle\int_{\Xi_-(x')} N(x',\xi')\, d\Omega(\xi')} \qquad (11)$$

Next, we write:

$$\text{"}\beta(X,N;x)\text{"} \quad \text{for} \quad \frac{\int_a \beta(X,N;x',x)\left[\int_{\Xi_-(x')} N(x',\xi')\, d\Omega(\xi')\right] dA(x')}{\int_a\int_{\Xi_-(x')} N(x',\xi')\, d\Omega(\xi')\, dA(x')} \tag{12}$$

And finally:

$$\text{"}\beta(X,N)\text{"} \quad \text{for} \quad \int_b \beta(X,N;x)\, dA(x) \quad . \tag{13}$$

The motivation for this sequence of definitions is made clear by computing the norm $|N_-(a)\,\mathcal{J}(X;a,b)|_Y$. Thus:

$$|N_-(a)\,\mathcal{J}(X;a,b)|_Y =$$

$$= \frac{1}{A(Y)}\int_b\int_{\Xi_+(x)} N_-(a)\,\mathcal{J}(X;a,b)\, d\Omega(\xi)\, dA(x)$$

$$= \frac{1}{A(Y)}\int_b\int_{\Xi_+(x)}\left[\int_a\int_{\Xi_-(x')} N(x',\xi')S_b(X;x',\xi';x,\xi)d\Omega(\xi')dA(x')\right]d\Omega(\xi)dA(x)$$

$$= \frac{1}{A(Y)}\int_b\int_a\left[\int_{\Xi_-(x')} N(x',\xi')S_b(X;x',x';x)\, d\Omega(\xi')\right] dA(x')dA(x)$$

$$= \frac{1}{A(Y)}\int_b\int_a\left[\beta(X,N;x',x)\int_{\Xi_-(x')} N(x',\xi')\, d\Omega(\xi')\right] dA(x')dA(x)$$

$$= \frac{1}{A(Y)}\int_b \beta(X,N;x)\left[\int_a\int_{\Xi_-(x')} N(x',\xi')\, d\Omega(\xi')dA(x')\right] dA(x)$$

$$= \beta(X,N)\ |N_-(a)|_Y \tag{14}$$

As in the case of the norm of the surface reflectance operators $r_\pm(a)$, $t_\pm(a)$ (Sec. 3.4) it can be shown with the help of the energy conservation principle that:

$$0 \le \beta(X,N) \le 1 \tag{15}$$

for every X,a,b on X, and every radiance function N. For a given X, we write

$$\text{"}\beta(X)\text{"} \quad \text{for} \quad \max_N \beta(X,N) \quad , \tag{16}$$

where the maximum operation is taken over the set of all radiance functions on Y. Then the conclusion in (14) implies:

$$|N_-(a)\,\mathscr{S}(X;a,b)|_Y \leq \beta(X)|N_-(a)|_Y \quad . \tag{17}$$

We say that $\mathscr{S}(X;a,b)$ is *norm contracting* if:

$$0 < \beta(X) < 1 \quad . \tag{18}$$

Variations of the Basic Theme

The operators $\mathscr{S}$ or $\mathscr{U}$ can be used as a basis for further definitions of operators which work with radiometric quantities other than radiance. Thus, following the patterns established in Secs. 3.3, 3.6, we could redesign $\mathscr{S}$ so as to map radiance into radiant emittance, or irradiance into radiance, etc. These brief comments will suffice to make the reader aware of the potential variations he himself may wring from $\mathscr{S}$ and $\mathscr{U}$ as the occasion may arise.

It should be noted in conclusion that the operators $\mathscr{S}(X;a,b)$ and $\mathscr{U}(X;a,b)$ serve the capacities of both reflectance and transmittance operators depending on the relative disposition of parts a and b over the boundary of X. Thus we agree to call $\mathscr{S}(X;a,b)$ or $\mathscr{U}(X;a,b)$ a *reflectance operator* whenever a = b, and call it a *transmittance operator* whenever a and b are disjoint, i.e., have no points in common. This convention attains its greatest conceptual utility when X is very irregular and no simple directional conventions are possible, such as are available in the case of plane-parallel media. Observe, that if X is a plane-parallel medium X(a,b), then our present convention essentially reduces to that established earlier for a plane-parallel medium X(a,b) with upper boundary a and lower boundary b. (See, e.g., (8)-(11) of Sec. 3.6).

3.9 Applications to General Spaces

The applications of the interaction principle will now be extended to general optical media. We will begin with some relatively simple but important extensions of the principles of invariance to curvilinear media such as spherical, cylindrical and toroidal media. Then the abstract versions of these media--one-parameter carrier spaces--are considered, and finally the illustrations culminate in the principles of invariance for completely arbitrary media which are not represented explicitly as one-parameter media. Throughout this section, the proceedings may best be viewed once again from the two vantage points defined and discussed in the introduction to Sec. 3.7. In regard to these vantage points, Sections 3.4-3.8 and the present section begin to illustrate the efficacy of the interaction principle, not only as a theoretical tool, but as one which shows promise in fostering novel methods of numerical computations in radiative transfer problems.

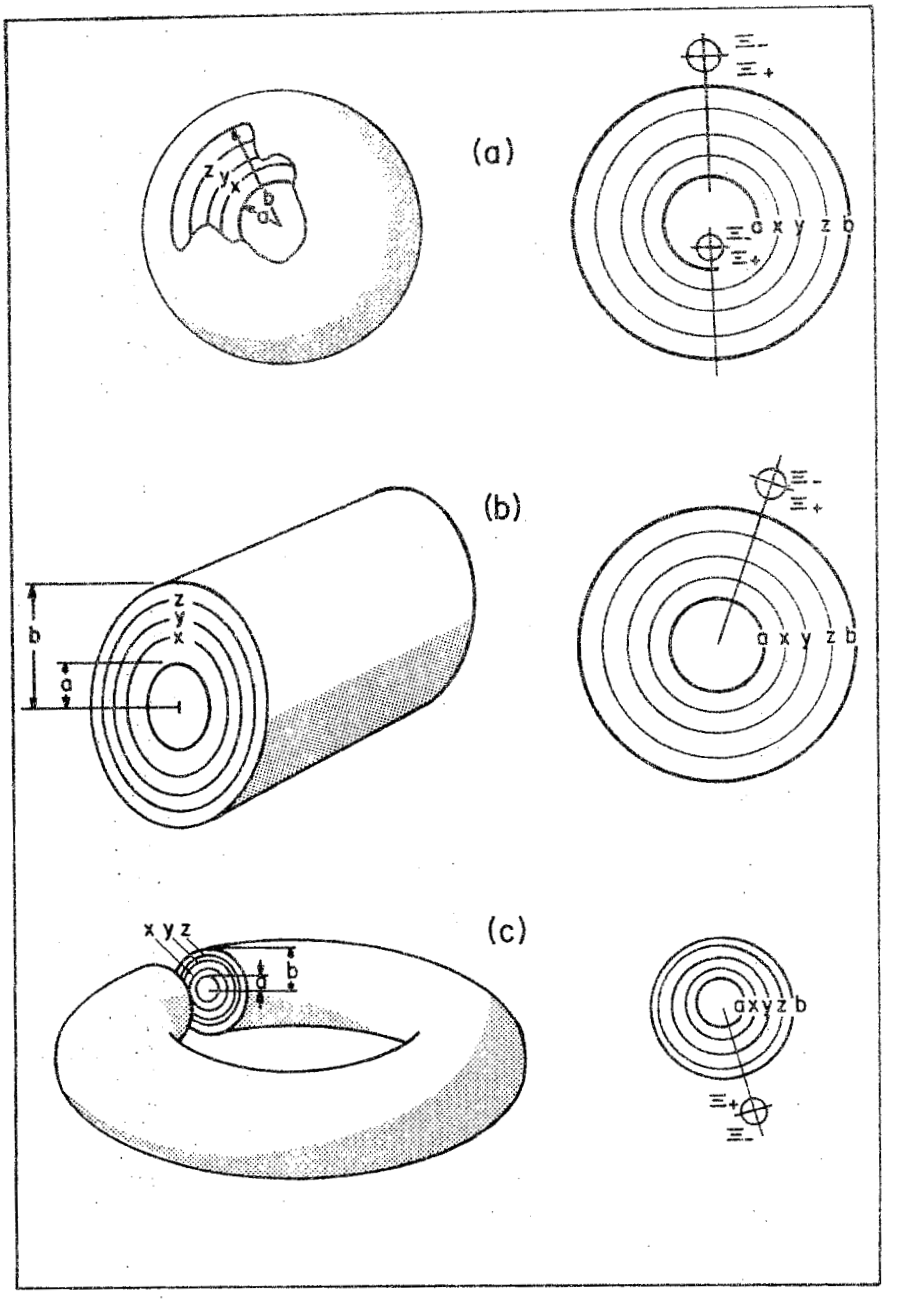


FIG. 3.25 Illustrating some applications of the interaction principle to various optical media.

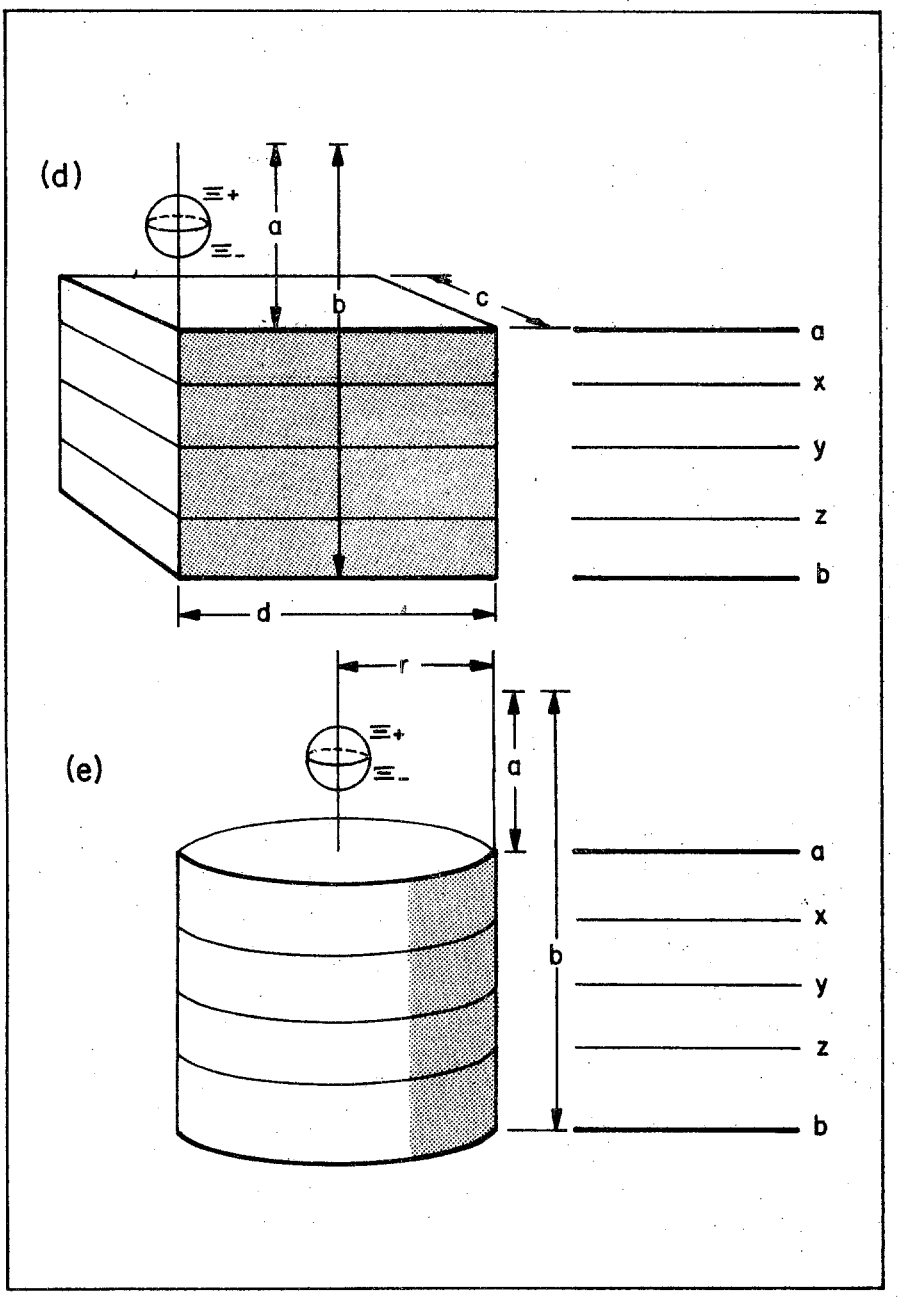


FIG. 3.25, concluded

Example 1: Principles of Invariance on Spherical Cylindrical, Toroidal Media

Our present goal is to use the interaction principle to formulate the principles of invariance on three common types of curvilinear media. Figure 3.25 depicts four instances of a curvilinear inhomogeneous optical medium X and one linear inhomogeneous optical medium. Part (a) depicts a spherical medium in the form of a spherical shell with inner radius a, and outer radius b. Adjacent to the schematic cutaway of the spherical shell is a diagram showing a partition of X into concentric spherical shells of radii x,y,z, with $a \leq x \leq y \leq z \leq b$. Similar descriptions can be made of the hollow cylindrical medium X in part (b) of Fig. 3.25, the hollow toroidal medium in part (c), the rectangular parallepiped medium in part (d), and the solid vertical cylindrical medium of part (e). In the case of the hollow cylindrical medium, its axial length may be finite or infinite. In the case of the parallelepiped, it may be of infinite extent in one or both lateral dimensions. In all five cases we may have $a = 0$. However, for the present illustration, we consider for generality $a \geq 0$.

We shall use as a prototype for the present formulations, the four principles of invariance derived in Example 3 of Sec. 3.7 for the case of plane-parallel media. As in that earlier example, we shall for brevity use the letters "a", "x", "y", etc., as names for both the parameter of the associated surface and the surface itself. Each medium in Fig. 3.25 will be designated by the name "X(a,b)", and subsets of X(a,b) as "X(x,z)", etc., just as in the plane-parallel case. Each medium is irradiated over surface a and b by incident external radiance distributions; $N_-^-(a)$ for a, $N_+^-(b)$ for b. No other sources are incident on or within X(a,b). The direction conventions are also analogous to the plane-parallel cases: we agree that at each point x on a parameter surface, the unit normal k(x) is directed toward the direction of *decreasing* parameter values.

Now, isolating X(a,y) and considering it black convexified, we enumerate the sets of incident radiance distributions:

A_1: all field radiance distributions like $N_-^-(a)$

A_2: all field radiance distributions like $N_+^-(y)$.

Enumerating the response radiance distributions, we have:

B_1: all surface radiance distributions like $N_+^+(a)$

B_2: all surface radiance distributions like $N_-^+(y)$.

The four interaction operators s_{ij} are:

s_{11} -- R(a,y)

s_{12} -- T(a,y)

s_{21} -- T(y,a)

s_{22} -- R(y,a)

These four operators are instances of the standard $\mathscr{S}$-operator $\mathscr{S}(X;a,b)$ in (6), where X is now X(a,y) and "b" is replaced by "a" where y is now a spherical surface in X(a,b). For the standard reflectance operator R(a,y) we have, explicitly:

$$\int_{\Xi_-(x')} \int_a [\ ] S_b(X;x',\xi';x,\xi)\ d\omega(\xi')dA(x')$$

where x is in spherical surface a, ξ is in $\Xi_+(x)$, and X is X(a,y). Similar constructions are made for the remaining three standard R and T operators. The R-T notation has been chosen so as to be uniform with the plane-parallel case of Sec. 3.7.

The interaction principle then states that:

$$N_+^+(a) = N_-^-(a)R(a,y) + N_+^-(y)T(y,a) \qquad (1)$$

$$N_-^+(y) = N_-^-(a)T(a,y) + N_+^-(y)R(y,a) \quad . \qquad (2)$$

By repeating this process now for X(y,b) we arrive at the analogous pair of statements:

$$N_+^+(y) = N_+^-(b)T(b,y) + N_-^-(y)R(y,b) \qquad (3)$$

$$N_-^+(b) = N_+^-(b)R(b,y) + N_-^-(y)T(y,b) \quad . \qquad (4)$$

The similarity of (1)-(4) with (15)-(18) of Sec. 3.7 is unmistakable: the interaction principle unifies all these instances. When we append the following two auxiliary equations:

$$N_+^+(y) = N_+^-(y) \qquad (5)$$

$$N_-^+(y) = N_-^-(y) \qquad (6)$$

the set (1)-(6) becomes autonomous, as usual. The remaining discussion of Example 2 of Sec. 3.7 now holds--virtually unchanged--including the definition (27) of iterated operators. Now, however, we use the standard $\mathscr{S}$-operator. It is not necessary to rewrite the principles of invariance I-IV of Example 3 of Sec. 3.7. They apply, as they stand to the present context. The only salient change is in the basis of the R and T operators: *we now use the standard $\mathscr{S}$-operator, as defined in (6) of Sec. 3.8, as a basis.* As in the plane-parallel case, the four principles of invariance are instrumental in allowing one to solve for $N_\pm(y)$ for every y, $a \leq y \leq b$, assuming

the standard R and T operators are known. These, in turn, are obtained from solutions of functional equations of the kind to be studied in Chapter 7.

Example 2: Invariant Imbedding Relation for One-Paramenter Media

The comprehensiveness of the principles of invariance, as extended from their classical plane-parallel settings by means of the interaction principle, begins to emerge as the five specific media in Example 1 are re-examined. In this example we systematically extend the results of Example 1 to their immediate logical limit. To do this, we ask: what is common to all the specific instances of Example 1? The answer is that these media are all constructed by assembling layer upon layer of surfaces of *geometrically similar shapes*. In part (a) of Fig. 3.25, we can imagine the hollow sphere to be built up from spherical surfaces of radii y, $a \leq y \leq b$, much in the way an onion is built up layer by layer. Parts (b) and (e) of Fig. 3.25 show that the cylindrical medium can be built up from cylindrical surfaces or circular plane surfaces. This two-way slice can be done for every instance shown in Fig. 3.25, and many others not shown. In each of the five instances displayed in Fig. 3.25, the medium X(a,b) may be imagined to consist of a set of geometrically similar surfaces X_x with $a \leq x \leq b$, i.e., with x a point in the interval [a,b] of real numbers. Thus we may set:

$$X(a,b) = \{X_x: \quad x \in [a,b]\} \tag{7}$$

i.e., X(a,b) is equal to the set of all geometrically similar surfaces X_x, each being indexed (identified) by a single parameter x drawn from an interval [a,b] of real numbers.

The examples of Fig. 3.25 only begin to illustrate first of all the great number of three-dimensional subsets of Euclidean space which are one-parameter spaces and available for study, and secondly the multiplicity of ways in which a given solid can be represented as a one-parameter space (viz. (b) and (e) of Fig. 3.25). Indeed, as can readily be verified any solid of Euclidean three-space may be represented as the union of a one-parameter family of two-dimensional surfaces, and in many distinct ways! Despite this great variety of shapes and sizes for each set X(a,b) and each source-free subset X(x,a) of X(a,b), we can isolate X(x,z), consider X(x,z) black convexified if it is concave, and enumerate the sets of incident radiance distributions on X(x,z):

A_1: all incident (surface) radiance distributions like $N_+(z)$

A_2: all incident (surface) radiance distributions like $N_-(x)$

where we are now following the pattern established in Example 4 of Sec. 3.7 and using surface radiances throughout (see, e.g., (21)-(24) of Sec. 3.7). The sets of response functions of interest are

B_1: all response (surface) radiance distributions like $N_+(y)$

B_2: all response (surface) radiance distributions like $N_-(y)$

In the present enumerations, $N_+(z)$ is the outward radiance distribution over the parameter surface X_z, $a \leq z \leq b$. The unit outward normal $\mathbf{k}(p)$ at point p on X_z is in the direction of decreasing parameter values.

The interaction principle then asserts the existence of four interaction operators s_{ij}:

$$s_{11} \text{ -- } \mathcal{T}(z,y,x)$$

$$s_{12} \text{ -- } \mathcal{R}(z,y,x)$$

$$s_{21} \text{ -- } \mathcal{R}(x,y,z)$$

$$s_{22} \text{ -- } \mathcal{T}(x,y,z) \quad .$$

These four operators are not instances of the operators defined in (6) of Sec. 3.8. Rather, they are exactly analogous to the complete reflectance and transmittance operators (40)-(43) of the plane-parallel case in Example 4 of Sec. 3.7. The interaction principle now yields the two statements:

$$\text{I.} \qquad N_+(y) = N_+(z)\,\mathcal{T}(z,y,x) + N_-(x)\,\mathcal{R}(x,y,z) \tag{8}$$

$$\text{II.} \qquad N_-(y) = N_+(z)\,\mathcal{R}(z,y,x) + N_-(x)\,\mathcal{T}(x,y,z) \tag{9}$$

which we can write as:

$$\left(N_+(y),N_-(y)\right) = \left(N_+(z),N_-(x)\right)\mathcal{H}(x,y,z)$$

where we have written:

$$\text{"}\mathcal{H}(x,y,z)\text{"} \quad \text{for} \quad \begin{bmatrix} \mathcal{T}(z,y,x) & \mathcal{R}(z,y,x) \\ \mathcal{R}(x,y,z) & \mathcal{T}(x,y,z) \end{bmatrix} .$$

The preceding equation is the *invariant imbedding relation* for one-parameter media. It is exactly analogous to (36) of Sec. 3.7. On the strength of this analogy, we summarize the preceding results as follows:

Let $X(= \{X_x: \ x \in [a,b]\})$ *be a one-parameter optical medium where* $[a,b]$ *is a closed interval in the extended real-number system. For every* $y \in [a,b]$, *there is a pair* $(N_+(y),N_-(y))$ *of (real or vector valued) response functions on* X_y. *Let "*$\mathcal{H}$*" denote the set of all ordered pairs* $(N_+(z),N_-(x))$ *of incident functions,* $[x,z] \subset [a,b]$ *with subsets* $\mathcal{H}_+$ *and* $\mathcal{H}_-$ *defined as* $\{N_+(z): \ z \in [a,b]\}$ *and* $\{N_-(x): \ x \in [a,b]\}$, *respectively. Then for every* x,y,z *with* $y \in [x,z] \subset [a,b]$ *there exists an interaction operator* $\mathcal{H}(x,y,z)$ *of* $\mathcal{H}$ *into* $\mathcal{H}$ *such that:*

$$\boxed{\left(N_+(y),N_-(y)\right) = \left(N_+(z),N_-(x)\right)\mathcal{H}(x,y,z)} \tag{10}$$

where we have written:

$$\text{"}\mathcal{M}(x,y,z)\text{"} \quad \text{for} \quad \begin{bmatrix} \mathcal{T}(z,y,x) & \mathcal{R}(z,y,x) \\ \mathcal{R}(x,y,z) & \mathcal{T}(x,y,z) \end{bmatrix} \tag{11}$$

in which $\mathcal{R}(z,y,x), \mathcal{R}(x,y,z)$ *are the complete reflectance operators with domains* $\mathcal{H}_+, \mathcal{H}_-$ *and ranges* $\mathcal{H}_-, \mathcal{H}_+$, *respectively;* $\mathcal{T}(z,y,x), \mathcal{T}(x,y,z)$ *are the complete transmittance operators with domains* $\mathcal{H}_+, \mathcal{H}_-$ *and ranges* $\mathcal{H}_+, \mathcal{H}_-$, *respectively. In addition,* $\mathcal{T}(x,z,z) = T(x,z)$ *is the standard transmittance operator for* $X(x,z)$ *and* $T(x,x,z) = I$, *the identity operator;* $\mathcal{R}(x,x,z) = R(x,z)$ *is the standard reflectance operator and* $\mathcal{R}(x,z,z) = 0$, *the zero operator.*

The preceding statement of the invariant imbedding relation is essentially that given in Ref. [233]. It is now a simple matter to deduce from (10) the semigroup properties.

$$\begin{aligned} \mathcal{T}(a,z,b) &= \mathcal{T}(a,y,b)\mathcal{T}(y,z,b) \\ \mathcal{R}(a,z,b) &= \mathcal{T}(a,y,b)\mathcal{R}(y,z,b) \end{aligned} \tag{12}$$

for complete transmittances (cf. (52), (53) in Example 5, Sec. 3.7). Furthermore, the principles of invariance for one-parameter media are readily forthcoming from (10)--or the equivalent set (8),(9). Indeed, setting $x = y$ in (8):

$$\text{I.} \qquad N_+(y) = N_+(z)T(z,y) + N_-(y)R(y,z) \quad .$$

Setting $z = y$ in (9):

$$\text{II.} \qquad N_-(y) = N_-(x)T(x,y) + N_+(y)R(y,x) \quad .$$

Principles III and IV now follow from I, II, as in Example 3 of Sec. 3.7. The present instances of the principles are identical in form to those in Sec. 3.7 and therefore need not be repeated in detail here. Furthermore, the representations of the present complete reflectance and transmittance operators in terms of the standard operators are identical in form to those given in (40)-(43) of Sec. 3.7 for the plane-parallel setting. Furthermore, the properties (44)-(47) also are easily shown to hold for the present complete reflectance and transmittance operators. The present forms of the standard R and T operators are important enough to repeat here. Thus for an arbitrary one-parameter optical medium X(a,b) we write:

$$\text{"R(a,b)"} \quad \text{for} \quad \int_a \int_{\Xi_-(x')} [\ ] S_b(X;x',\xi';x,\xi)\, d\Omega(\xi')dA(x')$$

if x is in a and ξ is in $\Xi_+(x)$.

$$\text{"T(a,b)"} \quad \text{for} \quad \int_a \int_{\Xi_-(x')} [\ ] S_b(X;x',\xi';x,\xi)\, d\Omega(\xi')dA(x')$$

if x is in b and ξ is in $\Xi_-(x)$.

"R(b,a)" for $\int_b \int_{\Xi_+(x')} [\ ]S_b(X;x',\xi';x,\xi)\ d\Omega(\xi')dA(x')$

if x is in b and ξ is in $\Xi_-(x)$.

"T(b,a)" for $\int_b \int_{\Xi_+(x')} [\ ]S_b(X;x',\xi';x,\xi)\ d\Omega(\xi')dA(x')$

if x is in a and ξ is in $\Xi_+(x)$.

Example 3: One-Parameter Media with Internal Sources

In this example we show how the interaction principle may be used in the task of formulating the equations governing the radiance distribution N(y) over a parameter surface X_y in a one-parameter optical medium X(a,b) which has internal sources generally distributed over an internal parameter surface X_s, $a \leq s \leq b$. To see at the outset the essential structure of the resultant equations, we assume that no other sources are incident on X(a,b).

Figure 3.26 depicts the one-parameter optical medium X(a,b) with the incident source (field) radiance distributions $N_+^o(s)$ and $N_-^o(s)$ over level s in X(a,b). We imagine $N_+^o(s)$ to irradiate X(a,s) and $N_-^o(s)$ to irradiate X(s,b). Thus, it is

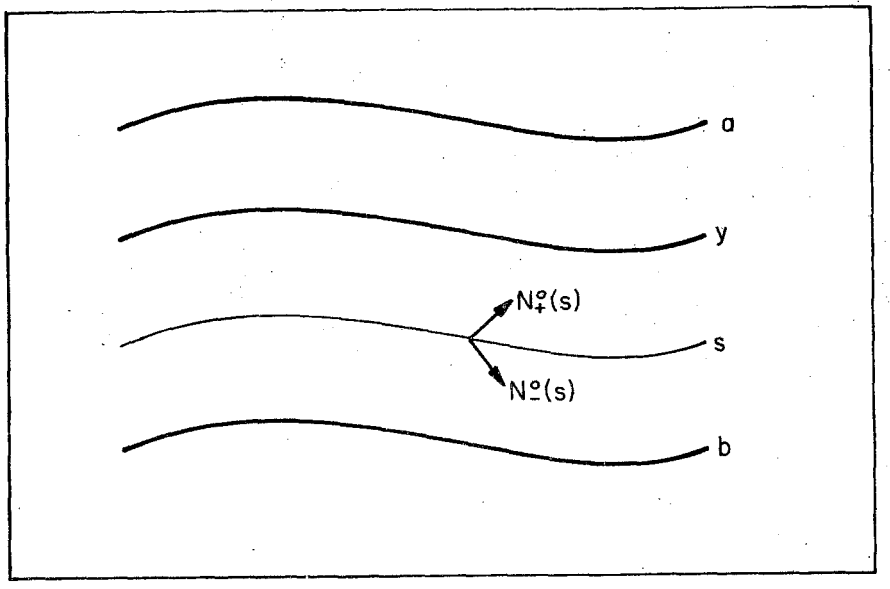


FIG. 3.26 Taking into account internal sources in general one-parameter media.

as if the incident source radiance distribution $N^o(s)$ $(= (N^o_+(s), N^o_-(s))$) were placed (like a thin transparent luminous vanilla filling) into $X(a,b)$ after the latter had been momentarily sliced open (like a layer cake) along X_s. It follows that the light field generated by this source may be viewed as being distinct from $N^o(s)$. We assume $N^o(s)$ to vary from point to point over X_s, and to be of arbitrary directional structure at each point of X_s. Thus in particular, $N^o(s)$ could consist of a narrow pencil of radiation at one point only, or it could be of uniform radiance over all directions at each point, etc. As usual $X(a,b)$ is generally inhomogeneous. The only requisite regularity in $X(a,b)$ is its geometric one-parameter structure (and even this can eventually be relaxed); optical properties and radiometric properties are left unconstrained--except for a modicum necessary to define integration and to have the usual additivity and continuity properties on which to build the operator algebra.

The given internal source over X_s suggests a partition of $X(a,b)$ into two parts $X(a,s)$ and $X(s,b)$. In order to invoke the interaction principle we could employ the usual notation "$N^+(y)$" for surface radiance of X_y, and "$N^-(y)$" for field radiance over X_y, $a \leq y \leq b$; however, now that some specific examples have shown how to systematically use surface radiance, we shall limit our use mainly to that kind of radiance. When "N" has no superscript, surface radiance is understood. The outward and inward directions over X_y for radiance distributions are as defined in Example 2.

Isolating $X(a,b)$ and enumerating the sets of incident radiance distributions on $X(a,b)$ we have:

A_1: all radiance distributions like $N^o_+(s)$

A_2: all radiance distributions like $N^o_-(s)$.

Enumerating the sets of response radiance distributions:

B_1: all radiance distributions like $N_+(y)$

B_2: all radiance distributions like $N_-(y)$.

Then $m = 2$, $n = 2$, and the interaction principle yields four interaction operators s_{ij} such that:

$$s_{11} \text{ -- } \Psi_{++}(s,y)$$

$$s_{12} \text{ -- } \Psi_{+-}(s,y)$$

$$s_{21} \text{ -- } \Psi_{-+}(s,y)$$

$$s_{22} \text{ -- } \Psi_{--}(s,y) \quad .$$

The fact that these four operators belong to the medium $X(a,b)$ is implicit in the notation. Occasionally it will be desirable to explicitly denote this fact (see, e.g., Sec. 7.13) and we shall then write "$\Psi_{++}(s,y:a,b)$" for $\Psi_{++}(s,y)$; "$\Psi_{+-}(s,y:a,b)$" for $\Psi_{+-}(s,y)$, etc. The interaction principle then states that, for every pair of levels y,s in $X(a,b)$:

$$N_+(y) = N_+^o(s)\Psi_{++}(s,y) + N_-^o(s)\Psi_{-+}(s,y) \qquad (13)$$

$$N_-(y) = N_+^o(s)\Psi_{+-}(s,y) + N_-^o(s)\Psi_{--}(s,y) \quad . \qquad (14)$$

In matrix form, (13) and (14) become:

$$\boxed{N(y) = N^o(s)\Psi(s,y)} \qquad (15)$$

where we have written:

$$\text{"}N(y)\text{"} \quad \text{for} \quad \left(N_+(y), N_-(y)\right)$$

$$\text{"}N^o(s)\text{"} \quad \text{for} \quad \left(N_+^o(s), N_-^o(s)\right)$$

$$\text{"}\Psi(s,y)\text{"} \quad \text{for} \quad \begin{bmatrix} \Psi_{++}(s,y) & \Psi_{+-}(s,y) \\ \Psi_{-+}(s,y) & \Psi_{--}(s,y) \end{bmatrix} .$$

We next show how the four operators $\Psi_{++}(s,y), \ldots, \Psi_{--}(s,y)$ can be represented in terms of the standard operators associated with the space $X(a,b)$ and its subsets $X(x,z)$. The derivation of the representation will proceed in two parts. The first part obtains a representation of $\Psi(s,s)$. The second part obtains the representation of $\Psi(s,y)$ with $s \neq y$.

We turn now to the case of $\Psi(s,s)$. Consider, for example, the subset $X(a,s)$. Isolating this subset and enumerating its incident functions and response functions under the present hypothesized conditions, we have $N_+^o(s)$ and the surface radiance $N_+(s)$ of $X(s,b)$ as incident functions which both act on the lower boundary of $X(a,s)$. These are the only incident functions on $X(a,s)$. Hence by principle of invariance II in Example 2 (with $z = y = s$, $x = a$) we have:

$$N_-(s) = \left(N_+^o(s) + N_+(s)\right)R(s,a) \quad . \qquad (16)$$

Similarly, for subset $X(s,b)$ and principle I of Example 2 (with $x = y = s$, $z = b$):

$$N_+(s) = \left(N_-^o(s) + N_-(s)\right)R(s,b) \quad . \qquad (17)$$

From (16) and (17):

$$N_+(s) = \left[N_-^o(s)R(s,b) + N_+^o(s)R(s,a)R(s,b)\right]\left[I - R(s,a)R(s,b)\right]^{-1} \qquad (18)$$

$$N_-(s) = \left[N_+^o(s)R(s,a) + N_-^o(s)R(s,b)R(s,a)\right]\left[I - R(s,b)R(s,a)\right]^{-1} \qquad (19)$$

Comparing (18) and (19) with (15) (in which s = y) and recalling that $N_{\pm}^{o}(s)$ are arbitrary, we deduce for the case* $a \leq s \leq b$:

$$\Psi_{++}(s,s) = R(s,a)R(s,b)\left[I-R(s,a)R(s,b)\right]^{-1} \tag{20}$$

$$\Psi_{+-}(s,s) = R(s,a)\left[I-R(s,b)R(s,a)\right]^{-1} \tag{21}$$

$$\Psi_{-+}(s,s) = R(s,b)\left[I-R(s,a)R(s,b)\right]^{-1} \tag{22}$$

$$\Psi_{--}(s,s) = R(s,b)R(s,a)\left[I-R(s,b)R(s,a)\right]^{-1} \tag{23}$$

We now go on to the second part of the representation derivation for $\Psi(s,y)$ with $s \neq y$. For definiteness we first assume $a \leq y < s \leq b$, as in Fig. 3.26. Then by the invariant imbedding relation (10) applied to X(a,s) we have:

$$\left(N_{+}(y),N_{-}(y)\right) = \left(N_{+}^{-}(s),N_{-}^{-}(a)\right)\mathcal{M}(a,y,s) \quad . \tag{24}$$

Here, by our present source conditions and choice of notation, we have:

$$N_{-}^{-}(a) = 0 \tag{25}$$

$$N_{+}^{-}(s) = N_{+}(s) + N_{+}^{o}(s) \quad . \tag{26}$$

From (24)-(26) we have:

$$N_{+}(y) = \left(N_{+}^{o}(s) + N_{+}(s)\right)\mathcal{T}(s,y,a) \tag{27}$$

$$N_{-}(y) = \left(N_{+}^{o}(s) + N_{+}(s)\right)\mathcal{R}(s,y,a) \quad . \tag{28}$$

Using (13) and (14) to give a representation of $N_{+}(s)$, (27) and (28) become:

$$\begin{aligned} N_{+}(y) &= \left[N_{+}^{o}(s) + \left(N_{+}^{o}(s)\Psi_{++}(s,s) + N_{-}^{o}(s)\Psi_{-+}(s,s)\right)\right]\mathcal{T}(s,y,a) \\ &= N_{+}^{o}(s)\left[\mathcal{T}(s,y,a) + \Psi_{++}(s,s)\,\mathcal{T}(s,y,a)\right] + \\ &\quad + N_{-}^{o}(s)\Psi_{-+}(s,s)\,\mathcal{T}(s,y,a) \end{aligned} \tag{29}$$

Further,

$$\begin{aligned} N_{-}(y) &= \left[N_{+}^{o}(s) + \left(N_{+}^{o}(s)\Psi_{++}(s,s) + N_{-}^{o}(s)\Psi_{-+}(s,s)\right)\right]\mathcal{R}(s,y,a) \\ &= N_{+}^{o}(s)\left[\mathcal{R}(s,y,a) + \Psi_{++}(s,s)\mathcal{R}(s,y,a)\right] + \\ &\quad + N_{-}^{o}(s)\Psi_{-+}(s,s)\mathcal{R}(s,y,a) \end{aligned} \tag{30}$$

---

*In the notation of Sec. 7.13, "$\Psi_{++}(s,s)$" becomes "$\Psi_{++}(s,s:a,b)$"; "$\Psi_{+-}(s,s)$" becomes "$\Psi_{+-}(s,s:a,b)$", and so on. The abbreviated notation is used whenever [a,b] is understood, as in the present discussion.

We return now to (13) and (14) which are the alternative representations of $N_{\pm}(y)$ in terms of the Ψ-operator. Since $N^{o}_{\pm}(s)$ are arbitrary, a comparison of (29) and (30) with the earlier representations implies that:

$$\text{—— a} \quad \Psi_{++}(s,y) = \left(I + \Psi_{++}(s,s)\right) \mathcal{T}(s,y,a) \tag{31}$$

$$\text{—— y} \quad \Psi_{+-}(s,y) = \left(I + \Psi_{++}(s,s)\right) \mathcal{R}(s,y,a) \tag{32}$$

$$\text{—— s} \quad \Psi_{-+}(s,y) = \Psi_{-+}(s,s)\, \mathcal{T}(s,y,a) \tag{33}$$

$$\text{—— b} \quad \Psi_{--}(s,y) = \Psi_{-+}(s,s) \mathcal{R}(s,y,a) \tag{34}$$

for the case $a \leq y < s \leq b$.* Here "I" denotes the identity operator for radiance distributions over X_s. With this set of equations the representation of the operator $\Psi(s,y)$, $s \neq y$, is essentially complete; for by interchanging "a" with "b", and "+" with "-" everywhere in (31)-(34), the case $a \leq s < y \leq b$ is obtained (see (39)-(42) of Sec. 7.13).

Two observations should be made on the nature of the operator $\Psi(s,y)$, $s \neq y$. First $\Psi(s,y)$ is not continuous at the diagonal points (s,s) of its domain. This can be seen from a study of its transmission component in (31). This discontinuity may be traced back to the equations (27) and (28), and the meanings ascribed to the radiances $N^{o}_{+}(s)$ and $N_{\pm}(y)$. A careful re-reading of the derivations will show that for every level y, $N_{+}(y)$ is the surface radiance of the boundary X_y of the subset $X(y,b)$ of $X(a,b)$. A similar reading holds for $N_{-}(y)$. Hence, as we approach the surface X_s from below we at first see $N_{+}(s)$, the surface radiance of X_s; then, abruptly, the source radiance $N^{o}_{+}(s)$ is added as we continue to move upward through the surface. In symbols, by (27), and for $y \leq s$:

$$\lim_{y \to s} \left(N_{+}(y) - N_{+}(s)\right) = N^{o}_{+}(s) \quad . \tag{35}$$

Since we have concentrated the origin of the source radiance $N^{o}_{\pm}(s)$ on a surface X_s, it is small wonder then that the outward (or inward) light field receives a jolt across X_s as we move upward (or downward) through level s. This jolt is duly recorded in $\Psi_{++}(s,y)$ in the manner shown in (31) (for $y < s$) to wit:

$$\lim_{y \to s} \left(\Psi_{++}(s,y) - \Psi_{++}(s,s)\right) = I \quad .$$

The second observation is that the use of the concept embodied in $\Psi_{++}(s,y)$ can be extended by postulating a *continuous* distribution of source radiance over a parameter interval $(s-\varepsilon, s+\varepsilon)$, $\varepsilon > 0$, or simply defining continuous functions

---

* In the notation of Sec. 7.13, "$\Psi_{++}(s,y)$" becomes "$\Psi_{++}(s,y:a,b)$", "$\Psi_{+-}(s,y)$" becomes "$\Psi_{+-}(s,y:a,b)$", and so on.

$N_{\pm}^{o}(s)$ for all s, $a \leq s \leq b$ and adjusting them to represent given physical situations as needed. The latter treatment is more general. The resulting operator equation will then be of the form:

$$N(y) = \int_a^b N^o(s)\Psi(s,y)\, ds \quad . \tag{36}$$

The interaction method then yields the integral operator:

$$\int_a^b [\ ]\Psi(s,y)\, ds \tag{37}$$

as a matter of course and the detailed decomposition of the new 2×2 matrix $\Psi(s,y)$ of operators proceeds analogously to that of the original $\Psi(s,y)$ above. The basis of this new integral operator will not be discussed here. The interested reader will find the general theorems leading to (37) in Sec. 3.16. Furthermore, readers interested in the discrete space version of (36) are referred to Chapters IX and X of Ref. [251]. A representation of (37) will be obtained in the irradiance context in Sec. 8.5.

We shall leave it as a simple exercise for the reader to show that, if there are a finite number of distinct external or internal sources $N_{\pm}^{o}(s_i)$, i = 1,...,n, of radiant flux of the discontinuous kind $N_{\pm}^{o}(s)$ considered above, then by virtue of the interaction principle, the resultant light field at level y within X(a,b) is given by

$$N(y) = \sum_{i=1}^{n} N^o(s_i)\Psi(s_i,y) \quad , \tag{38}$$

where we have written:

$$\text{"}N(y)\text{"} \quad \text{for} \quad (N_+(y), N_-(y))$$

$$\text{"}N^o(s_i)\text{"} \quad \text{for} \quad (N_+^o(s_i), N_-^o(s_i)) \quad ,$$

and where each $\Psi(s_i,y)$, i = 1,...,n is a 2×2 matrix of operators of the same structure as $\Psi(s,y)$ in (15), and where for every i, $1 \leq i \leq n$, the components $\Psi_{++}(s_i,y),\ldots,\Psi_{--}(s_i,y)$ are as defined in (20)-(23) and (31)-(34). If n = 1, then (38) reduces to (15).

The net result of this example is the demonstration that the problem of the internal sources of radiant flux in an arbitrary one-parameter medium can be reduced to a straightforward, albeit nontrivial, calculation using only the standard reflectance and transmittance for the medium (cf. Example 4 of Sec. 3.7). These standard reflectance and transmittance operators, in turn, are governed by certain differential equations which, when solved, yield the standard reflectance and transmittance functions for each given medium in terms of the inherent optical properties of the medium. These differential

equations will be developed in Chapter 7.

Example 4: Principles of Invariance for General Media

What is the geometric limit of validity of the principles of invariance? Can the principles of invariance be written down for radiance distributions in highly irregular media such as clouds, lakes, ponds, and wind-blown regions of the sea and other irregularly bounded natural hydrosols? The purpose of this example is to show that the answer to the latter question is in the affirmative. Once a few preliminary geometric conventions have been dispatched, an application of the interaction method yields the requisite principles of invariance.

Figure 3.27 depicts a general connected optical medium X. One may envision X as a cloud or a part of some natural hydrosol. As usual we assume no internal sources or reflecting boundaries. The boundary surface Y of X may be concave or convex.

In order to invoke the interaction principle we must have some idea of our goal. Let us re-examine a simple geometric setting in which the principles of invariance were derived. In such a simple case we know what the goal looks like. Fig. 3.20--the setting for Example 2 in Sec. 3.7--is a good starting point. We ask: what are the bare essentials of the

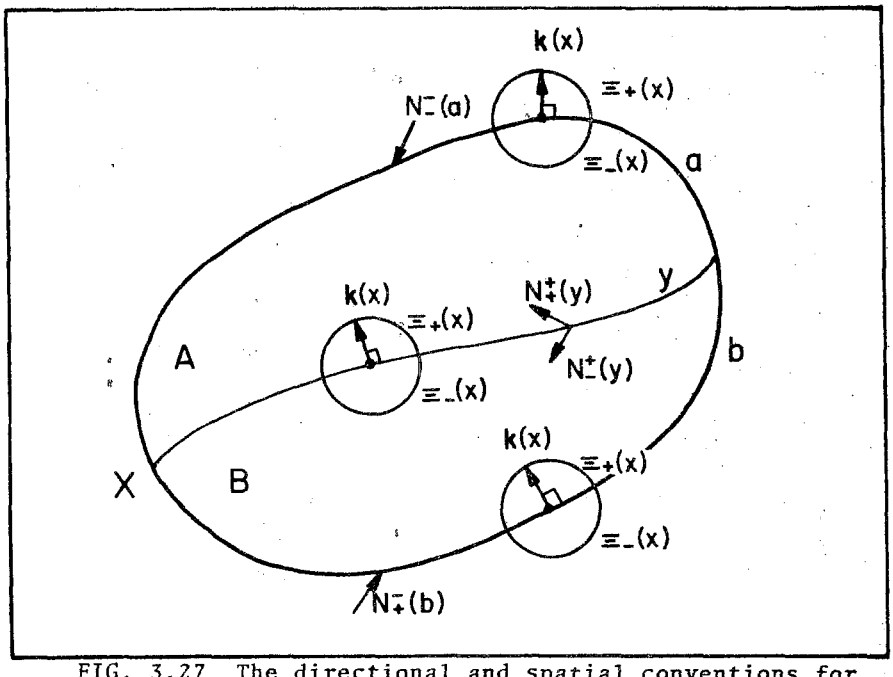


FIG. 3.27 The directional and spatial conventions for applying the principles of invariance to arbitrary optical media.

geometric setting there? First of all examine the medium: it is a plane-parallel medium; its boundary consists of two planes a and b. Thus if we are to emulate this for the case of X we should also divide the boundary of X into two parts a and b. Furthermore, the plane-parallel medium X(a,b) in Fig. 3.20 is partitioned into two parts: X(a,y) and X(y,b) by the internal plane y. The corresponding activity in X of Fig. 3.27 would be a partitioning of X into two parts: A and B by an internal hypothetical surface y. By combining these two requirements, we take the first step in the present derivation: *the arbitrary optical medium* X *is partitioned into two parts* A *and* B *by an internal surface* y*; this also partitions the surface of* X *into two parts* a *and* b*, the exterior boundaries of* A *and* B*, respectively.* Therefore the total boundary of A is the union of a and y; that of B is the union of b and y. There is one final geometric essential in Fig. 3.20 to be taken into account. That is the matter of direction. Our present goal is to emulate the classical plane-parallel case as far as possible. *Hence we establish the directionality conventions depicted in Fig. 3.27. The vector* k(x) *at each representative point of* a,y*, and* b *is an "outward" (or "upward" or "forward") direction* (the appropriate adjective is governed by the context within which one encounters X). As long as one is consistent in choice of directions, the following principles can be recast using any reference system an investigator cares to choose. The choice of directions depicted in Fig. 3.27 is internally consistent and coincides with the plane-parallel conventions in the limit as X approaches X(a,b), i.e., when a,b, and c are laterally extended and continuously deformed to become parallel planes.

With the foregoing partitioning and orientation of partition elements of the medium X established, we can specify the incident radiance distributions over X. Incident radiance distributions over part a of X will be denoted by "$N_-^-(a)$", outward radiance distributions over a will be denoted by "$N_+^+(a)$". On the internal surface y, "$N_+^+(y)$" will denote outward surface radiance distributions, "$N_+^-(y)$" will denote outward field radiance distributions, etc. Over b, incident radiance distributions will be denoted by "$N_+^-(b)$", etc. As usual, $N_\ddagger^\pm(y)$ is a function which assigns to each point x on surface y and direction ξ in $\Xi_+(\mathbf{k}(x))$ the surface radiance $N_\ddagger^\pm(x,\xi)$ of y directed into part A. At the same point and same direction $N_\ddagger^-(y)$ has the value $N_\ddagger^-(x,\xi)$, and by (32) of Sec. 2.5, $N_\ddagger^+(x,\xi) = N_\ddagger^-(x,\xi)$.

The geometrical prerequisites are sufficiently established to permit us to apply the interaction method to X. In particular part A is isolated and we enumerate the sets of incident radiometric quantities:

A_1: all incident radiance distributions like $N_-^-(a)$

A_2: all incident radiance distributions like $N_+^-(y)$.

The sets of response radiometric quantities are:

B_1: all response radiance distributions like $N_+^+(a)$

B_2: all response radiance distributions like $N_-^+(y)$.

In the present case m = 2, n = 2, and the interaction principle yields the four operators s_{ij}:

$$s_{11} \text{ -- } \mathscr{S}(A;a,a)$$

$$s_{12} \text{ -- } \mathscr{S}(A;a,y)$$

$$s_{21} \text{ -- } \mathscr{S}(A;y,a)$$

$$s_{22} \text{ -- } \mathscr{S}(A;y,y) \quad .$$

The four operators $\mathscr{S}(A;a,a),\ldots,\mathscr{S}(A;y,y)$ are instances of the $\mathscr{S}$-operator (6) of Sec. 3.8 (and hence the appropriate black convexification of A has been achieved). The interaction principle also yields the two interaction equations:

$$N_+^+(a) = N_-^-(a)\mathscr{S}(A;a,a) + N_+^-(y)\mathscr{S}(A;y,a) \tag{39}$$

$$N_-^+(y) = N_-^-(a)\mathscr{S}(A;a,y) + N_+^-(y)\mathscr{S}(A;y,y) \quad . \tag{40}$$

In a similar way the interaction method is applied to the subset B of X, with the following interaction equations as a result:

$$N_+^+(y) = N_+^-(b)\mathscr{S}(B;b,y) + N_-^-(y)\mathscr{S}(B;y,y) \tag{41}$$

$$N_-^+(b) = N_+^-(b)\mathscr{S}(B;b,b) + N_-^-(y)\mathscr{S}(B;y,b) \quad . \tag{42}$$

The auxiliary equations for the present formulation are:

$$N_+^+(y) = N_+^-(y) \tag{43}$$

$$N_-^+(y) = N_-^-(y) \quad . \tag{44}$$

The set of six equations (39)-(44) is autonomous. We can use (43) and (44) to reduce the set of six equations to four in terms of surface radiance only, and so the superscripts are no longer needed. Thus, with incident source radiances $N_-(a)$ and $N_+(b)$, we have:

$$N_+(a) = N_-(a)\mathscr{S}(A;a,a) + N_+(y)\mathscr{S}(A;y,a) \tag{45}$$

$$N_-(y) = N_-(a)\mathscr{S}(A;a,y) + N_+(y)\mathscr{S}(A;y,y) \tag{46}$$

$$N_+(y) = N_+(b)\mathscr{S}(B;b,y) + N_-(y)\mathscr{S}(B;y,y) \tag{47}$$

$$N_-(b) = N_+(b)\mathscr{S}(B;b,b) + N_-(y)\mathscr{S}(B;y,b) \quad . \tag{48}$$

Equations (45)-(48) are the requisite principles of invariance for X under the present partition into parts A and B. The middle two equations are autonomous. Their solutions are:

$$N_+(y) = \left[N_+(b)\mathcal{S}(B;b,y)+N_-(a)\mathcal{S}(A;a,y)\mathcal{S}(B;y,y)\right]\left[I-\mathcal{S}(A;y,y)\mathcal{S}(B;y,y)\right]^{-1} \tag{49}$$

$$N_-(y) = \left[N_-(a)\mathcal{S}(A;a,y)+N_+(b)\mathcal{S}(B;b,y)\mathcal{S}(A;y,y)\right]\left[I-\mathcal{S}(B;y,y)\mathcal{S}(A;y,y)\right]^{-1} \tag{50}$$

A comparison of these solutions with (25),(26) in Example 2 of Sec. 3.7 is instructive. Furthermore, the remaining part of the discussion of that example concerned with computation details pertains also, in essence, to the present situation.

Example 5: Invariant Imbedding Relation in General Media

By writing the solution $N_\pm(y)$ of (49),(50) in Example 4 in matrix form, the invariant imbedding relation for the general medium X of that example is obtained. Thus, we write:

$$\text{"}\mathcal{R}(a,y,b)\text{"} \quad \text{for} \quad \mathcal{S}(A;a,y)\mathcal{S}(B;y,y)\mathcal{D}(a,y,b)$$

$$\text{"}\mathcal{R}(b,y,a)\text{"} \quad \text{for} \quad \mathcal{S}(B;b,y)\mathcal{S}(A;y,y)\mathcal{D}(b,y,a)$$

$$\text{"}\mathcal{T}(a,y,b)\text{"} \quad \text{for} \quad \mathcal{S}(A;a,y)\mathcal{D}(b,y,a)$$

$$\text{"}\mathcal{T}(b,y,a)\text{"} \quad \text{for} \quad \mathcal{S}(B;b,y)\mathcal{D}(a,y,b)$$

$$\text{"}\mathcal{D}(a,y,b)\text{"} \quad \text{for} \quad \left[I - \mathcal{S}(A;y,y)\mathcal{S}(B;y,y)\right]^{-1}$$

$$\text{"}\mathcal{D}(b,y,a)\text{"} \quad \text{for} \quad \left[I - \mathcal{S}(B;y,y)\mathcal{S}(A;y,y)\right]^{-1} .$$

Then (49), (50) may be put into matricial form:

$$\left(N_+(y),N_-(y)\right) = \left(N_+(b),N_-(a)\right)\mathcal{M}(a,y,b) \tag{51}$$

where we have written

$$\text{"}\mathcal{M}(a,y,b)\text{"} \quad \text{for} \quad \begin{bmatrix} \mathcal{T}(b,y,a) & \mathcal{R}(b,y,a) \\ \mathcal{R}(a,y,b) & \mathcal{T}(a,y,b) \end{bmatrix} \tag{52}$$

We could have obtained (51) directly by means of the interaction method and deduced the form of $\mathcal{R}(a,y,b)$, $\mathcal{T}(a,y,b)$, etc., as in Example 4 of Sec. 3.7. We took the present route for the purpose of illustrating the manner in which (51) was first obtained (cf. Sec. 23, Ref. [251]). At any rate, the invariant imbedding relation is seen in its most general form in (51) and its basic role in radiative transfer theory is clear: *the invariant imbedding relation (51) relates a given incident boundary radiance distribution on X to a requisite internal radiance distribution in* X.

The specific evaluation of (52) is contingent on actual knowledge of the $\mathscr{S}$-operators for X. Functional relations for the $\mathscr{S}$-operator on a general medium may be found in I'-II' of Sec. 25 of Ref. [251]. The development of a *practical* general solution procedure in specific media of the functional relations I'-IV', especially relation I', constitutes one of the present outstanding (i.e., unsolved) applicational problems of modern radiative transfer theory (see Problem VIII, Sec. 141, Ref. [251]).

Example 6: Reflecting Boundaries and Interfaces

Hitherto we have examined all manners of radiometric interactions of surfaces with surfaces (Secs. 3.4, 3.5) and solids with solids (Secs. 3.7 and the present section). In this example we shall illustrate how the mixed radiometric interaction equations for surfaces and solids are derived using the interaction method. We shall consider a medium X with reflecting boundaries and an internal reflecting interface, but with no internal sources. Fig. 3.27 will serve to establish the geometric situation within X. We consider external sources incident on X over boundary surfaces a and b. In particular $N^{o}_{-}(a)$ and $N^{o}_{+}(b)$ are the two external sources.

When the radiometric problem at hand is analyzed into its essential components, it is seen that the medium X may be considered to consist of five interacting parts:

boundary	a
medium	A
interface	y
medium	B
boundary	b .

The interaction method then indicates the following desideratum: Each of these five subsets of X is required to be isolated and black-convexified, if concave; its sets of incident and response radiometric functions enumerated; the associated interaction equations written down; and the requisite auxiliary equations stated which will make the resulting system of interaction equations autonomous.

We begin by isolating the boundary part a. We first black-convexify surface a (see Sec. 3.8). This has the effect of reducing its interaction equations, (14)-(17) of Sec. 3.5, to a form identical to those of plane surfaces. In particular, the interaction equations for part a are:

$$N^{+}_{+}(a) = N^{o}_{-}(a)r_{-}(a) + N^{-}_{+}(a)t_{+}(a) \qquad (53)$$

$$N^{+}_{-}(a) = N^{-}_{-}(a)t_{-}(a) + N^{-}_{+}(a)r_{+}(a) \qquad (54)$$

which follow from (14), (15) of Sec. 3.5 in which the second two terms (the self-interaction terms) are zero by virtue of the black-convexification. The notation "$N^{o}_{-}(S)$" is now replaced by "$N^{o}_{-}(a)$"; "$N^{o}_{+}(S)$" by "$N^{-}_{+}(a)$"; etc.

Next, the medium part A is isolated and then black-convexified. In isolating A we are to imagine it separated from the boundary parts a and y as if we were carefully separating the meat of a two-slice piece of orange (A + B) from the peel (a + b) and interslice membrane (y) (see Fig. 3.27). Thus let us write, *ad hoc*, "$N_+^+(A)$" for the outward surface radiance of medium part A--i.e., without boundary part a. Similarly, "$N_+^-(A)$" will be the inward field radiance distribution over medium part A, just below the boundary part a, etc. We shall, however, retain the notation "$\mathscr{S}(A;a,a)$" for the interaction operator which describes the responses of the medium part A (i.e., the boundaryless A). The interaction method then yields the two equations (cf. Fig. 3.27):

$$N_+^+(A) = N_-^-(A)\mathscr{S}(A;a,a) + N_+^-(A)\mathscr{S}(A;y,a) \tag{55}$$

$$N_-^+(A) = N_-^-(A)\mathscr{S}(A;a,y) + N_+^-(A)\mathscr{S}(A;y,y). \tag{56}$$

Next, the interface part y is black-convexified if necessary (i.e., if non planar) and generally treated exactly in the manner of boundary part a. The resultant interaction equations are (cf. Fig. 3.27):

$$N_+^+(y) = N_-^-(y)r_-(y) + N_+^-(y)t_+(y) \tag{57}$$

$$N_-^+(y) = N_-^-(y)t_-(y) + N_+^-(y)r_+(y) . \tag{58}$$

The operators $r_\pm(y), t_\pm(y)$ are instances of those defined in (10),(11) of Sec. 3.3.

The medium part B is isolated and black-convexified exactly in the manner of medium part A. The resultant equations are (cf. Fig. 3.27):

$$N_+^+(B) = N_-^-(B)\mathscr{S}(B;y,y) + N_+^-(B)\mathscr{S}(B;b,y) \tag{59}$$

$$N_-^+(B) = N_-^-(B)\mathscr{S}(B;y,b) + N_+^-(B)\mathscr{S}(B;b,b) . \tag{60}$$

Finally the boundary part b is then isolated and black-convexified, if necessary, and generally treated in the manner of boundary part a. The resultant interaction equations are (cf. Fig. 3.27):

$$N_+^+(b) = N_-^-(b)r_-(b) + N_+^o(b)t_+(b) \tag{61}$$

$$N_-^+(b) = N_-^-(b)t_-(b) + N_+^o(b)r_+(b) . \tag{62}$$

The interaction method is brought to its final stage by appending the appropriate auxiliary equations. For the present problem, we have (cf. Fig. 3.27):

$$N_+^+(A) = N_+^-(a) \tag{63}$$

$$N_-^+(a) = N_-^-(A) , \tag{64}$$

which couple sets (53),(54) and (55),(56). Further:

$$N_+^+(y) = N_+^-(A) \tag{65}$$

$$N_-^+(A) = N_-^-(y) \quad , \tag{66}$$

which couple sets (55),(56) and (57),(58). Further:

$$N_+^+(B) = N_+^-(y) \tag{67}$$

$$N_-^+(y) = N_-^-(B) \quad , \tag{68}$$

which couple sets (57),(58) and (59),(60). Finally:

$$N_+^+(b) = N_+^-(B) \tag{69}$$

$$N_-^+(B) = N_-^-(b) \quad , \tag{70}$$

which couple sets (59),(60) and (61),(62). The set of 18 equations (53)-(70) is autonomous. Using the eight auxiliary equations (63)-(70) (all instances of (32) of Sec. 2.5), the set reduces to a less complex set of ten equations, which, written uniformly in response surface radiance form are:

$$N_+(a) = N_-^o(a)r_-(a) + N_+(A)t_+(a) \tag{71}$$

$$N_-(a) = N_-^o(a)t_-(a) + N_+(A)r_+(a) \tag{72}$$

$$N_+(A) = N_-(a)\mathcal{S}(A;a,a) + N_+(y)\mathcal{S}(A;y,a) \tag{73}$$

$$N_-(A) = N_-(a)\mathcal{S}(A;a,y) + N_+(y)\mathcal{S}(A;y,y) \tag{74}$$

$$N_+(y) = N_-(A)r_-(y) + N_+(B)t_+(y) \tag{75}$$

$$N_-(y) = N_-(A)t_-(y) + N_+(B)r_+(y) \tag{76}$$

$$N_+(B) = N_-(y)\mathcal{S}(B;y,y) + N_+(b)\mathcal{S}(B;b,y) \tag{77}$$

$$N_-(B) = N_-(y)\mathcal{S}(B;y,b) + N_+(b)\mathcal{S}(B;b,b) \tag{78}$$

$$N_+(b) = N_-(B)r_-(b) + N_+^o(b)t_+(b) \tag{79}$$

$$N_-(b) = N_-(B)t_-(b) + N_+^o(b)r_+(b) \quad . \tag{80}$$

Of these ten equations, the eight "interior" equations, i.e., (72)-(79) form an autonomous system. Hence the problem of a general optical medium with two reflecting boundaries and one reflecting interface, requires the solution of eight simultaneous integral equations. As in the case of earlier examples in this section, the solution of the present formulation is contingent on knowledge of the various operators $r_\pm(a)$, $r_\pm(b)$, $t_\pm(a)$, $t_\pm(b)$, $\mathcal{S}(A;a,a),\ldots,\mathcal{S}(B;b,b)$.

Example 7: The Unified Atmosphere-Hydrosphere Problem

The atmosphere of the earth and the surface of the earth (over both land and sea) form a system of radiometrically interacting optical media. In this example the interaction equations for the atmosphere and the hydrosphere are obtained as an autonomous system using the interaction method of formulating radiative transfer problems.

The formulations of Example 6 are readily adapted to the present task. It remains to specify the physical meanings of the five parts a,A,y,B,b of space X in that example. Consider Fig. 3.28 which is a schematic cross section of the earth. Let "A" denote the atmosphere, and "B" the hydrosphere, i.e., natural waters. Let "y" denote the air-water interface. Let "b" denote the bottom of the hydrosphere; at those places where B is solid earth, then b shall coincide with y. Finally, "a" denotes the transparent upper boundary of A.

With these choices, the system of equations (71)-(80) reduces to seven equations: (73)-(79) with $N_-(a) = N^o_+(a)$ and $N^o_+(b) = 0$. For the solution procedure of this problem using the techniques of discrete space theory, see Sec. 71 of Ref. [251]. (Whenever media of differing indices of refraction are considered, it is implicitly understood that radiance functions are divided by the n^2 of the medium in which they are defined, see (4) of Sec. 2.6.) Because of the lateral extensiveness of A and B, the problem of solving the set (75)-(79) can be considered within the domain of the plane-parallel case if attention is restricted to a region such as that enclosed in the dashed radial lines of Fig. 3.28.

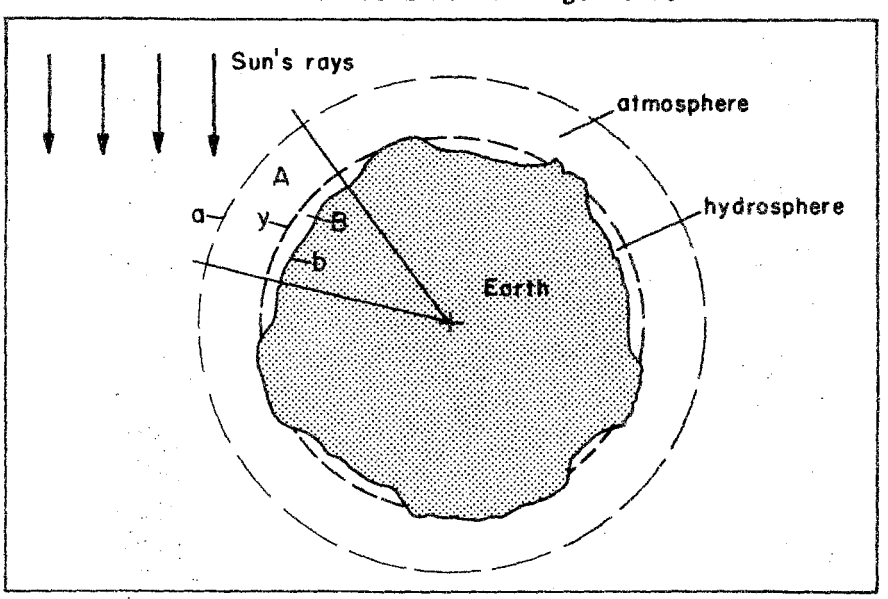


FIG. 3.28 A schematic diagram for the unified atmosphere-hydrosphere problem.

Example 8: Several Interacting Separate Media

The quantitative evaluation of the radiometric interaction of separate clouds in the atmosphere, or separate portions of lakes, oceans, or other natural hydrosols forms an interesting and difficult radiative transfer problem. The methods of discrete space theory have been used to develop a systematic means of computing the $\mathscr{S}$-operators of such irregular types of media (see, e.g., Chapter X, Ref. [251]). For actual computations, the latter formulation may use the estimates of the $\mathscr{S}$-operators as supplied by the methods of Chapter X of Ref. [251].

3.10 Derivation of the Beam Transmittance Function

In this and the following four sections the interaction principle is used to derive the basic inherent optical properties and the integral equation of transfer for radiance in general optical media. In the present section the beam tranmittance function is derived which in turn will yield the first of the inherent optical properties in Sec. 3.11, namely the volume attenuation function.

Let X be a general source-free optical medium. Consider a point x in X and a direction ξ at x. These together determine a natural path in X--a path for a light ray through x with direction ξ. By specifying a length r, a path segment $\mathscr{P}_r(x,\xi)$ is determined within X. For simplicity of exposition we shall introduce the beam transmittance concept in a medium X in which the index of refraction is constant. Then $\mathscr{P}_r(x,\xi)$ is a sensed, straight line segment in X with initial point x, terminal point $x+r\xi$ and length r. We shall write "z" for $x+\xi r$. The requisite steps for the definition in a completely general medium with variable index of refraction will be clear from the following derivation.

Our goal is to define the beam transmittance function--the well-known function of radiative transfer theory which assigns to an initial radiance $N_0(x,\xi)$ at point x of $\mathscr{P}_r(x,\xi)$ the residual radiance $N_r^0(z,\xi)$ at z after traversing $\mathscr{P}_r(x,\xi)$. This residual radiance $N_r^0(z,\xi)$ may be described as consisting of photons which have travelled the entire length of $\mathscr{P}_r(x,\xi)$ without having undergone scattering or absorption. Were we not in possession of the interaction principle, we would simply define the beam transmittance of $\mathscr{P}_r(x,\xi)$ as the ratio $N_r^0(x,\xi)/N_0(z,\xi)$. Such a definition is quite acceptable and implicitly assumes the terms "scattering" and "absorption" as understood, using their meanings as grounded in other departments of physics such as electromagnetic theory, and thus requiring no further elucidation. However, we choose not to use the words "scattering" and "absorption" at this stage of the exposition. Instead we ask: can the basic idea of "beam transmittance" be communicated without presupposing the physical notions of scattering and absorption? If this is possible then the theory of radiative transfer, at least in this area, is kept self-contained and freed from unnecessarily using

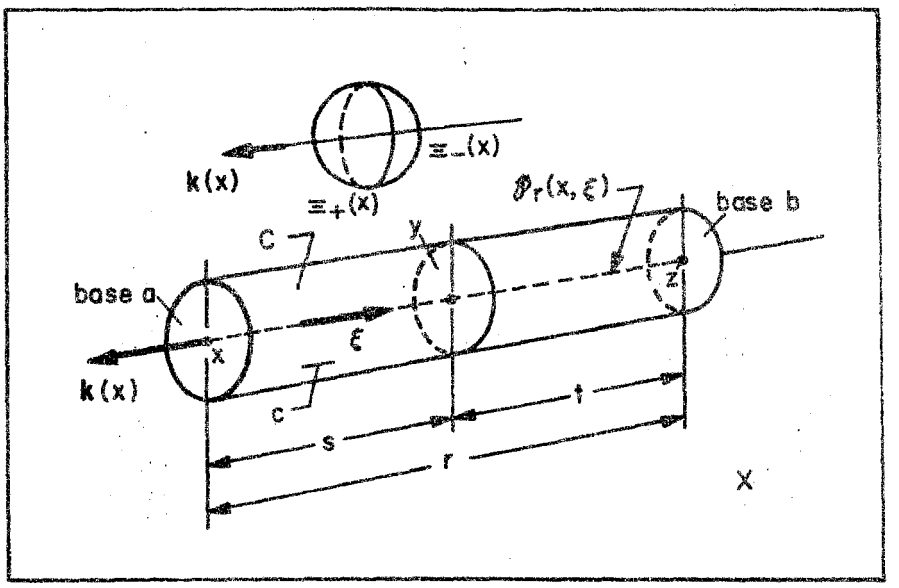


FIG. 3.29 The general setting for beam transmittance, path function, and path radiance derivations. The basic path $\mathscr{P}_r(x,\xi)$ is imbedded in a cylindrical volume located in a general optical medium X (a portion of the atmosphere or hydrosphere).

undefined terms. But to do this requires essentially a geometrically-based definition of absorption and scattering. Scattering and absorption are not ostensibly geometric concepts as they occur in, say, classical electromagnetic theory. Hence there appears to be an impasse between us and our present goal along the path we wish to travel as long as we try to retain the strict electromagnetic meanings of these terms. However, if we pause and examine closely the meanings of the terms "absorption" and "scattering" as they are customarily used in radiative transfer contexts, we find that there is no essential reference to physical processes beyond the simplest conservation-type of activity. Radiant flux is observed to enter and leave a medium and it is possible to take a census of the immigrating and emigrating photons. What happens inside the medium is not of immediate concern--*only the phenomenological aspects of the transfer of radiant flux through the medium is of concern.* Thus the terms "scattering" and "absorption" as they are used in radiative transfer theory proper are characterizable solely by means of a *geometric measure of radiant flux.* Such a view was taken in an earlier study (Secs. 11, 19, Ref. [251]) and it was found possible to give a complete geometric characterization of the concepts of residual radiance, scattering, absorption, attenuation, and hence beam transmittance as they are used in classical radiative transfer theory. We shall adopt the view of Ref. [251]

in the present discussion and with the concept of the $\mathscr{S}$-operator at our disposal, we shall use it to define the beam transmittance function geometrically.

Let C be a right cylinder in X containing $\mathscr{P}_r(x,\xi)$ as its axis. This cylinder is depicted schematically in Fig. 3.29. The plane base of C at x is denoted by "a", that at z by "b". Let $\mathbf{k}(x)$ $(= -\xi)$ be the unit outward normal to C over base a. Then C is an instance of a one-parameter optical medium with distance along the direction $-\mathbf{k}(x)$ as the parameter of the space (cf. Example 2, Sec. 3.9 and part (e) of Fig. 3.25). Consider C isolated. Since C is convex we shall not need to black-convexify it. Let $N_-(a)$ be an incident radiance distribution over a circular conical set D of directions with axis ξ and on base a. Let $N_-(a)$ be the only source of radiant flux in C. Since C is isolated, radiant flux leaving C is not considered to enter it again nor is the radiant flux in X considered to enter it. In short "isolated" means, as usual, that a subset is *conceptually* excised from its master set and placed in a dark vacuum for a controlled radiometric study.

We now direct attention to the radiance distribution

$$N_-(a)\,\mathscr{T}(a,y,b)$$

where y is an intermediate plane cross section of C, as in Fig. 3.29, and $\mathscr{T}(a,y,b)$ is the complete transmittance operator for C (cf. (11) of Sec. 3.9). This radiance distribution is thought of as that part of $N_-(a)$ transmitted from a to y. Hence:

$$\left(N_-(a)\,\mathscr{T}(a,y,b)\right)\,\mathscr{T}(y,b,b)$$

is that part of $N_-(a)$ transmitted from a to b and emerging over base b. By (12) of Sec. 3.9 (with z = b) we have:

$$N_-(a)\,\mathscr{T}(a,y,b)\,\mathscr{T}(y,b,b) = N_-(a)\,\mathscr{T}(a,b,b) \qquad (1)$$

or, alternatively:

$$N_-(a)\,\mathscr{T}(a,y,b)T(y,b) = N_-(a)T(a,b) \quad .$$

The crucial step in the definition of the beam transmittance can now be taken. We select the value of $N_-(a)\mathscr{T}(a,b,b)$ at z for the direction ξ and let $a \to \{x\}$, which is tantamount to letting $C \to \mathscr{P}_r(x,\xi)$. We then go on to write:

$$\text{"}N^o_r(z,\xi)\text{"} \quad \text{for} \quad \lim_{C \to \mathscr{P}_r(x,\xi)} [N_-(a)T(a,b)](z,\xi) \qquad (2)$$

where the notation indicates that $N_-(a)T(a,b)$ is evaluated at z in the direction ξ during the limit process. Let us also write:

$$\text{"}N_o(x,\xi)\text{"} \quad \text{for} \quad N_-(a)(x,\xi) \quad ,$$

i.e., $N_o(x,\xi)$ is the value of $N_-(a)$ at x and ξ. Then we shall write:

$$\text{"}T_r(x,\xi)\text{"} \quad \text{for} \quad \lim_{C \to \mathscr{P}_r(x,\xi)} [T(a,b)](z,\xi) \tag{3}$$

evaluated for a unit radiance function $N_o(a)$ and at z in the direction ξ. By (2) the definitional identity:

$$N_r^o(z,\xi) = N_o(x,\xi)T_r(x,\xi) \tag{4}$$

readily follows, where $z = x + r\xi$. For every x,ξ,r, the quantity $T_r(x,\xi)$ is a dimensionless, non negative real number associated with the path $\mathscr{P}_r(x,\xi)$, and is called the *beam transmittance* of $\mathscr{P}_r(x,\xi)$. The radiance $N_r^o(z,\xi)$ is the *directly transmitted* (or *residual*, or *reduced* or *unattenuated*) radiance.

The compatibility of the limit in (3) with classical theory can only be fully made after a sufficient amount of the theory has been developed. Those readers wishing an immediate indication of the compatibility may consult statements Ia and Ib of Sec. 23 in Ref. [251]. Statement Ia shows at any rate that the limit of $T(a,b)$ need not go to zero as $a \to \{x\}$ and $D \to \{\xi\}$ for short paths. Furthermore statement Ib states that $R(a,b)$ does go to zero under these conditions. The rigorous proof of the existence of the limit in (3) requires the specific postulation of very mild regularity properties of the underlying radiative process. We shall not digress here to establish such fine points. Interested readers are referred to Chapter III of Ref. [251] for a complete discussion of what regularity properties are needed in this matter.

Next, by means of (42) of Sec. 3.7 we see that the statement:

$$\lim_{C \to \mathscr{P}_r(x,\xi)} \mathscr{T}(a,y,b) = T(a,y) \tag{5a}$$

follows from:

$$\lim_{C \to \mathscr{P}_r(x,\xi)} R(y,b) = 0 \tag{5b}$$

for every y, $a \leq y \leq b$. Statement (5b), in turn, follows from the S-continuity of the function $S(X;\cdot,\cdot,\cdot,\cdot)$. The S-continuity of $S(X;\cdot,\cdot,\cdot,\cdot)$ follows from statement Ib of Sec. 23 of Ref. [251]. From (1), (3) and the preceding observation (5a) we have for every x,ξ,s and t such that $s + t = r$:

$$T_r(x,\xi) = T_s(x,\xi)T_t(x + s\xi,\xi) \tag{6}$$

or

$$T_{s+t} = T_s T_t$$

for short when x and ξ are understood. Furthermore, from the fact that

$$\mathcal{T}(a,a,a) = I \quad ,$$

we have for every x and ξ:

$$\boxed{T_o(x,\xi) = 1} \quad , \tag{7}$$

or simply:

$$T_o = 1 \quad ,$$

for short. In other words the beam transmittance of a path of length zero, is unity. Finally, from (15),(17) of Sec. 3.8 it follows that for every x,ξ,r:

$$\boxed{T_r(x,\xi) \leq 1} \tag{8}$$

or

$$T_r \leq 1 \quad ,$$

for short. Property (6) is called the *multiplicative (or semigroup) property*; property (7) is the *identity property*, and (8) is the *contraction property* of beam transmittance. These properties can also be based on properly presented physical arguments (see, e.g., Sec. 16 of Ref. [251]). However, it is of some interest to see that they can be made to follow formally from the interaction principle and (in the case of (8)) the law of conservation of radiant energy.

The intuitive meaning of $T_r(x,\xi)$ for a path $\mathcal{P}_r(x,\xi)$ in an optical medium can now be discerned. In everyday words, the radiance $N_-(a)\,\mathcal{T}(a,b,b)$ $(= N_-(a)T(a,b))$ transmitted through a finite cylinder C arrives at base b after having possibly been scattered several times. However, as C approaches $\mathcal{P}_r(x,\xi)$, so that base a approaches the set {x} consisting of the single point x, there is progressively less "elbow" room for the flux in transmission to scatter about within C on its way to b. In the limit as C goes to $\mathcal{P}_r(x,\xi)$, any radiant flux that travels from x to z along $\mathcal{P}_r(x,\xi)$ has the "straight and narrow path" to follow. If it happens that a photon in transit from point x to point z is scattered exactly forward or backward a certain number of times, this photon is considered to constitute part of the residual radiance at z, for it is impossible at present to operationally distinguish, in the time-averaged residual radiance, between a photon which has travelled along $\mathcal{P}_r(x,\xi)$ without scattering and one which has travelled along $\mathcal{P}_r(x,\xi)$ and which has been scattered. Hence it will not be inconsistent to speak of $N_r^o(z,\xi)$ as radiance consisting of those photons originally

constituting $N_0(x,\xi)$ which have not been absorbed or scattered from $\mathscr{P}_r(x,\xi)$ as they travelled from x to $x+r\xi = z$. This is the *phenomenological interpretation* of $N_r^0(x,\xi)$, and we shall adopt it in the present work.

The generalization of the foregoing results to the case of non-constant index of refraction is effected by repeating all steps with $(N_-(a)/n^2)$ instead of $N_-(a)$, and building a tube of natural paths around $\mathscr{P}_r(x,\xi)$ using cross section a as a base from which the paths begin with normal incidence. The motivation for using the quotient $N_-(a)/n^2$ rests in the n^2-law for radiance in (4) of Sec. 2.6.

This discussion is concluded with the observation that the beam transmittance $T_r(x,\xi)$ associated with the path $\mathscr{P}_r(x,\xi)$ in X is independent of the radiance distribution in X. This may be seen by returning to (3) and recalling that the standard transmittance operator R(a,b) is an integral operator whose kernel function S_b is derived from an interaction operator obtained via the interaction principle. A re-examination of the conclusion of the interaction principle will show that an interaction operator is independent of the members of its domain sets A_i and range sets B_j, in other words interaction operators do not depend on the radiance distributions (i.e., the light fields) in X.

Inherent and Apparent Optical Properties

The foregoing observation is of considerable importance in establishing the basic optical properties of a natural optical medium X. *An optical property* P *of an optical medium* X *(*P *in the form of a number, function, or operator) which is independent of the light fields (in the form of radiance distributions) in* X *will be called an* <u>*inherent optical property*</u> *of* X*; otherwise,* P *is an* <u>*apparent optical property*</u>. Hence, the beam transmittance function T_r which assigns to the path $\mathscr{P}_r(x,\xi)$ in X the beam transmittance $T_r(x,\xi)$, being independent of the radiance distribution in X, is an inherent optical property of X. We shall return to the systematic study of inherent and apparent optical properties in Chapter 9.

3.11 Derivation of the Volume Attenuation Function

The volume attenuation function is a measure of how much radiance a light beam loses per unit length of travel under the joint action of scattering and absorption processes. In this section we shall develop the concept of the volume attenuation function with the beam transmittance function as a starting point.

Let $\mathscr{P}_r(x,\xi)$ be a natural path in an optical medium X with associated beam transmittance $T_r(x,\xi)$. If a parcel of radiant flux of unit initial radiance traverses $\mathscr{P}_r(x,\xi)$, then on the one hand $T_r(x,\xi)$ is the amount of radiance transmitted over $\mathscr{P}_r(x,\xi)$, and on the other hand:

$$1 - T_r(x,\xi)$$

is the amount of radiance lost over $\mathscr{P}_r(x,\xi)$. Hence if we write:

$$\text{"}\alpha_r(x,\xi)\text{"} \quad \text{for} \quad \frac{1 - T_r(x,\xi)}{r}$$

then we can say that $\alpha_r(x,\xi)$ is the average amount of radiance lost per unit length for a beam of initial unit radiance traversing $\mathscr{P}_r(x,\xi)$.

We are almost at our goal. It remains to write:

$$\boxed{\text{"}\alpha(x,\xi)\text{"} \quad \text{for} \quad \lim_{r\to 0} \alpha_r(x,\xi)} \quad . \tag{1}$$

The function α which assigns to each x in X and ξ in Ξ the non negative value $\alpha(x,\xi)$ given in (1) is called the *volume attenuation function*. The dimensions of $\alpha(x,\xi)$ are L_r^{-1} (inverse radial length--see note (h) to Table III of Sec. 2.12). That $\alpha(x,\xi)$ is a non negative number follows from the contraction property of $T_r(x,\xi)$, ((8) of Sec. 3.10).

Once again we have assumed the existence of a limit without appropriate mathematical preamble. As in the case of (2), (3) and (5) of Sec. 3.10, we are concerned here only with the formal conceptual content of the interaction principle. There should not be any concern at present about the existence of the limit (1) above and the limits in (2), (3), and (5) of Sec. 3.10. These limits can always be made to exist in an acceptable and workable setting by postulating physically reasonable regularity properties of the underlying radiative process. However, what is of greatest importance here is the fact that there now exists a formal deductive chain of arguments connecting the volume attenuation function with the interaction principle. In this way we have shown that the volume attenuation function is an inherent optical property of an optical medium and a property whose conceptual roots are logically linked to the same principle which yields the reflectance and transmittance operators for surfaces and general subsets in the medium.

A useful connection between α and T_r is the exponential representation of T_r using α. This connection is derived by using the multiplicative property of T_r ((6) of Sec. 3.10) to write:

$$\frac{T_{r+s} - T_r}{s} = \frac{T_s - 1}{s} \cdot T_r = (-\alpha_s)T_r \quad .$$

The definition of α_s was used to obtain the second equality. Letting $s\to 0$, we obtain:

$$\boxed{\frac{dT_r}{dr} = -\alpha T_r} \tag{2}$$

For given x and ξ this is an elementary differential equation for T_r, with known function α, whose solution is:

$$T_r = \exp\left\{-\int_0^r \alpha\, dr'\right\} \quad ,$$

or in more explicit notation:

$$\boxed{T_r(x,\xi) = \exp\left\{-\int_0^r \alpha(x(r'),\xi(r'))\, dr'\right\}} \quad . \tag{3}$$

We have used the identity property for T_r ((7) of Sec. 3.10) to find the integration constant for the particular solution (3) of equation (2). Here x(r') and ξ(r') are the location and direction of a variable point within $\mathscr{P}_r(x,\xi)$, at distance r' from x along the path. If the index of refraction were constant, then x(r') = x + r'ξ; ξ(r') = ξ for every r', $0 \le r' \le r$; and (3) would become:

$$T_r(x,\xi) = \exp\left\{-\int_0^r \alpha(x + r'\xi,\xi)\, dr'\right\} \quad , \tag{4}$$

and (2) would take the form:

$$\frac{dT_r(x,\xi)}{dr} = -\alpha(x + r\xi,\xi)T_r(x,\xi) \quad . \tag{5}$$

For a discussion of (5) in the case of variable index of refraction, see Sec. 17 of Ref. [251]. In that section there is also an alternative derivation of the function α using empirical radiances and empirical attenuating volumes. An experimental procedure for determining α is given in Sec. 13.4 of this work.

3.12 Derivation of Path Radiance and Path Function

We continue the sequence of derivations, begun in Sec. 3.10, leading to the derivation of the integral equation of transfer for radiance along a path $\mathscr{P}_r(x,\xi)$ in an optical medium X. In this section we give a derivation of two important components of this equation: the path radiance, and the path function associated with $\mathscr{P}_r(x,\xi)$, and conclude with a derivation of an important connection between them.

The Path Radiance

Let $\mathscr{P}_r(x,\xi)$ be a path in an optical medium X such as that depicted in Fig. 3.29. Once again, for simplicity, we assume constant index of refraction over X and no internal

sources. Let C be a right cylinder about $\mathscr{P}_r(x,\xi)$. To introduce the concept of the path radiance of $\mathscr{P}_r(x,\xi)$ we conceptually isolate C, think of it as a one-parameter optical medium as in Sec. 3.10, and direct attention to the incident radiance distribution $N_-(c)$ on its flank c. The surface of C consists of three parts: two bases a,b, and flank c. The unit outward normal to c is assumed defined at each point of c. The incident radiance over flank c generates a light field within C by scattering processes within C with response radiance $N_-(b)$ over base b given by:

$$N_-(b) = N_-(c)\,\mathscr{S}(C;c,b)$$

where $\mathscr{S}(C;c,b)$ is the requisite interaction operator supplied by the interaction principle. The physical significance of this equation should be thoroughly understood before proceeding. We have enclosed $\mathscr{P}_r(x,\xi)$ (itself a conceptual object) in a hypothetical cylinder C. The radiance distribution at the base b of C is then the response of C to the incident natural light field all along its flank c. We direct attention now to point z in base b and direction $\xi(= -k(x))$ at z. The value of $N_-(b)$ at z and ξ is the radiance, generated within C, in response to the incident distribution $N_-(c)$. We now let C approach $\mathscr{P}_r(x,\xi)$, and keep an analytic eye on the limit of the value of $N_-(b)$ at x and ξ. We write:

$$\text{"}N_r^*(z,\xi)\text{"} \quad \text{for} \quad \lim_{C \to \mathscr{P}_r(x,\xi)} [N_-(c)\mathscr{S}(C;c,b)](z,\xi) \qquad (1)$$

evaluated at z and ξ. We call $N_r^*(z,\xi)$ the *path radiance* associated with $\mathscr{P}_r(x,\xi)$. The path radiance $N_r^*(z,\xi)$ is commonly known as the "space light" or "haze of day" or "diffuse radiance" along a path of sight. If one looks along a path of sight under water or in the atmosphere such that the path terminates in some dark region, the veiling light between the observer and the dark region is due to path radiance. What we have just done is to go from the observable case for $N_-(b)$ to the mathematical limit (1). The latter is easier to work with for the same general reason that lines and planes in geometry are easier to work with than narrow rods and thin flat sheets.

The Path Function

We consider now the path radiance of very short paths $\mathscr{P}_r(x,\xi)$. In this way we shall come to the concept of the path function. Part (a) of Fig. 3.30 depicts a path $\mathscr{P}_r(x,\xi)$ with a right cylindrical subset C of X about $\mathscr{P}_r(x,\xi)$ as axis. As in Sec. 3.10, we consider C a one-parameter optical medium with the usual geometrical conventions. We start once again with a conceptual isolation of C and consider the equation for $N_-(b)$. We are interested now not only in the limit of $N_-(b)$ as C approaches $\mathscr{P}_r(x,\xi)$, but also in the limit of $N_-(b)$ as r is allowed to go to 0 while x and ξ are held fixed. We can anticipate the limit by examining the specific integral structure of $\mathscr{S}(C;c,b)$. From (6) of Sec. 3.8:

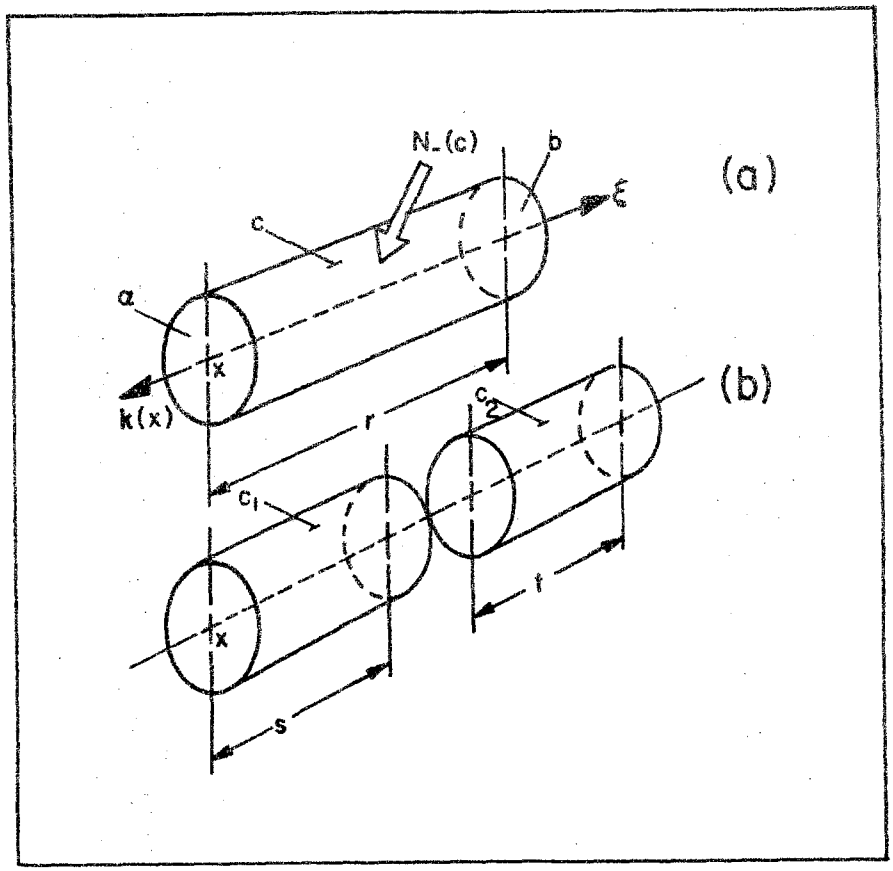


FIG. 3.30 Detail for the path function derivation.

$$N_{-}(b) = \int_{c}\int_{\Xi_{-}(x')} N(x',\xi')S_b(C;x',\xi';x,\xi)\, d\Omega(\xi')dA(x')$$

where $N(x',\xi')$ is the value of $N_{-}(c)$ at x' on c and ξ' in $\Xi_{-}(x')$. By the S'-continuity of the empirical funtion S_b given in (1) and (3) of Sec. 3.8, we conclude that $N_{-}(b)$ goes to zero as r goes to zero. We can say a little more than this: the value $N_r^*(z,\xi)$ eventually goes to zero linearly with r. This can be seen by conceptually slicing C into two parts C_1, C_2 shown in (b) of Fig. 3.30. This partitions the flank c into two parts c_1 and c_2 such that:

$$N_{-}(b) = N_{-}(c_1)\mathscr{S}(C_1;c_1,b) + N_{-}(c_2)\mathscr{S}(C_2;c_2,b)$$

where $c_1 \cup c_2 = c$, and $N_{-}(c_1)$ is equal to $N_{-}(c)$ over points of c_1 and zero over points of c_2. $N_{-}(c_2)$ is defined similarly. Each summand contributes a proportional share to $N_{-}(b)$. This

shows, with the foregoing continuity property, that $N_r^*(z,\xi)$ has additivity and continuity properties with respect to r; and so it becomes plausible that it has an r-derivative.

The foregoing observations lead us to write, for every path $\mathscr{P}_r(x,\xi)$ in X:

$$\text{"}N_*(x,\xi)\text{"} \quad \text{for} \quad \lim_{r\to 0} \frac{N_r^*(z,\xi)}{r} \quad . \tag{2}$$

From this definition and (1) we have the following useful definitional identity for every path $\mathscr{P}_r(x,\xi)$ in X:

$$N_*(x,\xi) = \lim_{\substack{C\to \mathscr{P}_r(x,\xi) \\ r\to 0}} \frac{[N_-(c)\,\mathscr{S}(C;c,b)](z,\xi)}{r} \tag{3}$$

where the limit is evaluated at x and ξ. The number $N_*(x,\xi)$ is value of the *path function* N_* at x,ξ, on $\mathscr{P}_r(x,\xi)$. The dimensions of $N_*(x,\xi)$ are those of radiance per unit radial length. That is, the dimension of $N_*(x,\xi)$ is $\mathbf{P}^{\pm}\mathbf{A}^{-1}\mathbf{\Omega}^{-1}\mathbf{L}_r^{-1} = \mathbf{P}^{\pm}\mathbf{V}^{-1}\mathbf{\Omega}^{-1}$ (see Table 3 of Sec. 2.12 and note (h) of that table). Hence $N_*(x,\xi)$ may be thought of in either of two ways: the *radiance per unit length* (in the direction ξ) generated by scattering of an incident radiance distribution at point z into the direction ξ; or the *radiant intensity per unit volume* generated by scattering of an incident radiance distribution at point z into the direction ξ.

To summarize: By imagining a path $\mathscr{P}_r(x,\xi)$ in an optical medium X as the axis of a narrow cylindrical subset C of X, as in Fig. 3.30, and by rigorously relating the incident and response radiances $N_-(c)$ and $N_-(b)$ over C by means of the interaction operator $\mathscr{S}(C;c,b)$, we can define precisely, and on a phenomenological level, the concepts of path radiance $N_r^*(z,\xi)$ and path function N_* associated with the path $\mathscr{P}_r(x,\xi)$.

The Connection Between Path Function and Path Radiance

We now reverse the step taken in (2) and obtain a formula for the path radiance $N_r^*(z,\xi)$ given knowledge of the path function values $N_*(y,\xi)$ at each point y of $\mathscr{P}_r(x,\xi)$ between x and $z = x + r\xi$. The result will be a generally useful relation which also forms an indispensible component of the integral equation of transfer for radiance.

The first step in the derivation is to set up once again the one-parameter optical medium C around an arbitrary path $\mathscr{P}_r(x,\xi)$ in X, as in Fig. 3.29. In this way we have a useful scaffolding around $\mathscr{P}_r(x,\xi)$ with many precise analytical ladders on which to clamber up and down its extent; ladders in the form of the complete reflectance and transmittance operators for the medium, and $\mathscr{S}$-operators for arbitrary portions of the medium. Now a novel aspect of the present light field within C--when C is considered as a one-parameter optical medium--is that radiant flux enters C from X through the flanking surface c of C. This is novel in the sense that normally the invariant imbedding relation formally provides for

external incident sources to enter C only on parameter levels a and b. However, by a simple device, to be introduced below, we can rigorously convert the incident radiance distributions on the flank c into internal sources at an arbitrary finite number of parameter levels in C. The problem can then be converted to one which, by Example 3 of Sec. 3.9, we know how to solve in principle. After that it will be a relatively simple matter to establish the requisite relation between path radiance and path function. The relation is established by means of six steps, the first of which begins below.

The conversion of the incident radiance distribution $N_-(c)$ on the flanking surface c of the medium C into internal source radiances is accomplished as follows. Part (a) of Fig. 3.31 depicts the one-parameter medium partitioned into a finite number n of cylindrical segments $C_1, C_2, \ldots, C_n$ by cross section surfaces $a_1, a_2, \ldots, a_{n-1}$, with $a_0 = a$ and $a_n = b$. This partition, in turn, partitions the flanking surface c of C into n pieces $c_1, c_2, \ldots, c_n$. Hence for each i, $i = 1, \ldots, n$, part C_i of C is bounded by parameter surfaces a_{i-1} on top and a_i on bottom (*upward* or *outward*, as usual, is the direction of $\mathbf{k}(x)$). The incident radiance distribution is partitioned correspondingly into n distributions $N_-(c_i)$ such that:

$$N_-(c_i) = N_-(c)$$

over c_i and

$$N_-(c_i) = 0$$

over c outside of c_i. In other words, we have decomposed $N_-(c)$ analytically by writing

$$N_-(c) = \sum_{i=1}^{n} \chi_i N_-(c_i) \tag{4}$$

where χ_i is the characteristic function for part c_i of c such that $\chi_i(y) = 1$ or 0 according as y is, or is not in c_i, respectively. Thus the radiometric effect of $N_-(c)$ on C is equivalent to the radiometric effect of this sum-decomposition of distributions on C. The proof of this is obvious: replace the left side of (4) by the right side of (4) in the interaction equation:

$$N_-(b) = N_-(c) \mathscr{S}(C;c,b) \quad . \tag{5}$$

Thus we have:

$$N_-(b) = \sum_{i=1}^{n} \chi_i N_-(c_i) \mathscr{S}(C;c,b) \tag{6}$$

$$= \sum_{i=1}^{n} N_-(c_i) \mathscr{S}(C;c_i,b) \quad . \tag{7}$$

We now begin the *second step*. The result (7) suggests

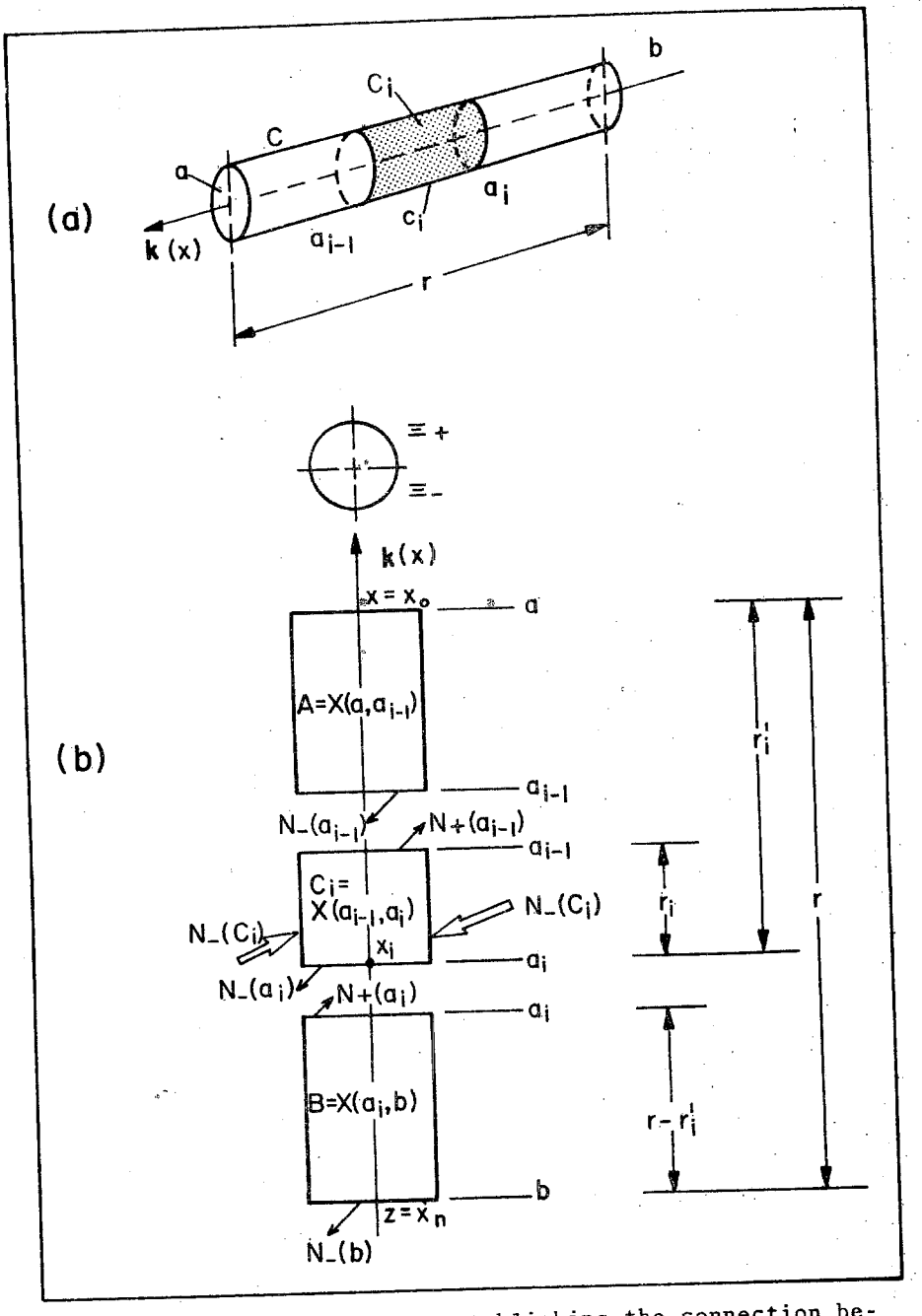


FIG. 3.31 Details for establishing the connection between path function and path radiance.

that the radiance distribution $N_-(b)$ can be thought of as generated by the response of C to the n incident radiance distributions $N_-(c_i)$. We wish to convert the external radiance distribution over c_i to an equivalent set of one or more source radiance distributions inside C. To see what is entailed in such a task, suppose for the moment that $N_-(c_j) = 0$ for every c_j except one, say $N_-(c_i)$. We now apply the interaction method to C_i by supposing C_i is conceptually isolated from C without disturbing the existing light field. Part (b) of Fig. 3.31 will be of help to visualize this conceptual excision of C_i from C. Parts A and B of C outside of C_i are shown in the diagram slightly pulled away from C_i. Next, the sets of incident radiometric functions on C_i are enumerated. These are:

A_1: all incident radiance distributions like $N_-(c_i)$

A_2: all incident radiance distributions like $N_+(a_i)$

A_3: all incident radiance distributions like $N_-(a_{i-1})$.

The sets of response radiance distributions of interest are:

B_1: all response radiance distributions like $N_+(a_{i-1})$

B_2: all response radiance distributions like $N_-(a_i)$.

The members of A_1 are field radiances. All the remaining radiances are surface radiances of their indicated surfaces. In the case of the members of A_2 and A_3, we have used the equality of field and surface radiance at a given point to replace the field radiances. In the present case of the principle, m = 3, n = 2, and the six operators s_{ij} given by the interaction principle are:

$$s_{11} \text{ -- } \mathscr{S}(C_i; c_i, a_{i-1})$$

$$s_{12} \text{ -- } \mathscr{S}(C_i; c_i, a_i)$$

$$s_{21} \text{ -- } T(a_i, a_{i-1})$$

$$s_{22} \text{ -- } R(a_i, a_{i-1})$$

$$s_{31} \text{ -- } R(a_{i-1}, a_i)$$

$$s_{32} \text{ -- } T(a_{i-1}, a_i)$$

These six operators have been given their standard notation as established earlier in the chapter. The first two are instances of (6) of Sec. 3.8. The remaining four are instances of the standard reflectance and transmittance operators for a general one-parameter medium, as given in Example 2 of Sec. 3.9.

The interaction principle then yields the equations for the response radiance distributions $N_-(a_i)$, $N_+(a_{i-1})$:

$$N_+(a_{i-1}) = N_-(c_i)\mathscr{S}(C_i;c_i,a_{i-1}) + \\ + N_+(a_i)T(a_i,a_{i-1}) + N_-(a_{i-1})R(a_{i-1},a_i) \qquad (8)$$

$$N_-(a_i) = N_-(c_i)\mathscr{S}(C_i;c_i,a_i) + \\ + N_+(a_i)R(a_i,a_{i-1}) + N_-(a_{i-1})T(a_{i-1},a_i) \; . \qquad (9)$$

The second step in the present derivation continues by comparing the preceding statements with those of the principles of invariance for the medium C under the same partitioning. The partitioning that is presently being used is such that:

$$C_i \text{ corresponds to } X(x,z)$$

in principles I and II of Example 2 in Sec. 3.9. Furthermore, equation (8) corresponds to I when we let y = x in I and also let:

$$y \text{ in principle I correspond to } a_{i-1} \text{ in (8)}$$

$$z \text{ in principle I correspond to } a_i \text{ in (8)} \; .$$

This detailed comparison is carried out in order to establish the fact that (8) may be interpreted as an instance of principle I applied to a source-free medium $X(a_{i-1},a_i)$ in C with a source of radiance of magnitude $N_-(c_i)\mathscr{S}(C_i;c_i,a_{i-1})$ at level a_{i-1}. In a similar way we see that (9) is interpretable as an instance of principle II applied to a source-free medium $X(a_{i-1},a_i)$ in C with a downward source (downward in C) of radiance of magnitude $N_-(c_i)\mathscr{S}(C_i;c_i,a_i)$ at level a_i. The conclusion we can now reach is the following: *with the current assumption about* $N_-(c_i)$ *as being the only source on* C *(to which we have momentarily agreed at the outset of this step of the discussion) we can then write the interaction equations for the light field in* C *with the interpretation that there are two sources of radiant flux in* C*: one upward flux confined to level* a_{i-1} *with magnitude* $N_-(c_i)\mathscr{S}(C_i;c_i,a_{i-1})$*, and one downward flux confined to level* a_i *with magnitude* $N_-(c_i)\mathscr{S}(C_i;c_i,a_i)$. This is the crucial observation needed in order to carry the derivation to completion, and since it appears somewhat complex, we state its message in still another way: Suppose a radiance meter were placed inside parts A or B of the cylinder C; for definiteness, say the meter was in B. What we have inferred from (9) is the prediction that the meter's reading would be unchanged were we to replace the external incident radiance distribution $N_-(c_i)$ by an internal source radiance distribution over parameter surface a_i equal to $N_-(c_i)\mathscr{S}(C_i;c_i,a_i)$ and an internal source radiance distribution over parameter surface a_{i-1} equal to $N_-(c_i)\mathscr{S}(C_i;c_i,a_{i-1})$.

The *third step* in the derivation consists in applying to the n-1 other partition pieces of C what we have just learned from consideration of the case of C_i. The net result is that the response radiance distribution $N_-(b)$ (the incip-

ient path radiance) may be considered to be generated within C by n-1 internal source radiances $N_{\pm}^{o}(a_i)$ at $i = 1,\ldots,n-1$, where we have written:

$$\text{"}N_{+}^{o}(a_{i-1})\text{"} \quad \text{for} \quad N_{-}(c_i)\mathcal{S}(C_i;c_i,a_{i-1})$$

$$\text{"}N_{-}^{o}(a_i)\text{"} \quad \text{for} \quad N_{-}(c_i)\mathcal{S}(C_i;c_i,a_i)$$

and one external radiance distribution $N_{-}^{o}(a_n)$, where we have written:

$$\text{"}N_{-}^{o}(a_n)\text{"} \quad \text{for} \quad N_{-}(c_n)\mathcal{S}(C_n;c_n,a_n) \ .$$

The *fourth step* in the derivation is the use of (38) of Sec. 3.9 to represent $N_{-}(b)$. For, by the conclusion of the third step, we have reduced the problem of representing $N_{-}(b)$ as generated by $N_{-}(c)$ to the problem of representing $N_{-}(b)$ as generated by the internal sources at the n-1 interval levels a_i of C. Toward this end, in (38) of Sec. 3.9, let $y = b$, $s_i = a_i$, and let the summation run from 1 to n-1. Then, for the contribution from the n-1 internal sources we have from (38) of Sec. 3.9:

$$[N_{+}(b),N_{-}(b)] = N(b) = \sum_{i=1}^{n-1} N^{o}(a_i)\Psi(a_i,b) \ .$$

From this, and the nth source $N_{-}^{o}(a_n)$ we have:

$$N_{-}(b) = \sum_{i=1}^{n-1} N_{+}^{o}(a_i)\Psi_{+-}(a_i,b) + \sum_{i=1}^{n-1} N_{-}^{o}(a_i)\Psi_{--}(a_i,b) + N_{-}^{o}(a_n) \quad (10)$$

wherein $a_i < b$ for $i = 1,\ldots,n-1$. The latter inequalities show that we may use the representations of the Ψ-operators in (31)-(34) of Sec. 3.9 once we have interchanged "+" with "-", and interchanged "a" with "b" everywhere in those equations. In particular, we have:

$$\Psi_{+-}(a_i,b) = \Psi_{+-}(a_i,a_i)\mathcal{T}(a_i,b,b) \quad (11)$$

$$\Psi_{--}(a_i,b) = [I + \Psi_{--}(a_i,a_i)]\mathcal{T}(a_i,b,b) \ . \quad (12)$$

With this observation, (10) becomes:

$$N_{-}(c)\mathcal{S}(C;c,b) = N_{-}(b) =$$

$$= \sum_{i=1}^{n-1} N_{-}(c_{i+1})\mathcal{S}(C_{i+1};c_{i+1},a_i)\Psi_{+-}(a_i,a_i)T(a_i,b) +$$

$$+ \sum_{i=1}^{n-1} N_{-}(c_i)\mathcal{S}(C_i;c_i,a_i)[I + \Psi_{--}(a_i,a_i)]T(a_i,b)$$

$$+ N_{-}(c_n)\mathcal{S}(C_n;c_n,a_n) \ . \quad (13)$$

The *fifth step* of the derivation consists in applying to (13) the limit process, $C \to \mathscr{P}_r(x,\xi)$. This application is facilitated by noting that:

$$\lim_{C \to \mathscr{P}_r(x,\xi)} \Psi_{+-}(a_i,a_i) = 0$$

$$\lim_{C \to \mathscr{P}_r(x,\xi)} \Psi_{--}(a_i,a_i) = 0$$

which follow from (21) and (23) of Sec. 3.9 and (5b) of Sec. 3.10. Applying the limit process $C \to \mathscr{P}_r(x,\xi)$ to (13) we have, with the aid of definition (3) of Sec. 3.10, and (1):

$$N^*_r(z,\xi) = \sum_{i=1}^{n-1} N^*_{r_i}(x_i,\xi) T_{r-r_i}(x_i,\xi) + N^*_{r_n}(x_n,\xi) \quad .$$

Hence:

$$\boxed{N^*_r(z,\xi) = \sum_{i=1}^{n} N^*_{r_i}(x_i,\xi) T_{r-r_i'}(x_i,\xi)} \qquad (14)$$

or, briefly:

$$N^*_r = \sum_{i=1}^{n} N^*_{r_i} T_{r-r_i'} \quad ,$$

where r_i, r_i' and x_i are defined as in (b) of Fig. 3.31. (Hence $x_0 = x$, $x_n = z$, $r_n' = r$, and $T_{r-r_n'}(x_i,\xi) = 1$.) Equation (14) is a useful exact formula for the path radiance of $\mathscr{P}_r(x,\xi)$ in terms of the path radiance over the component segments $\mathscr{P}_{r_i}(x_{i-1},\xi)$; $i = 1,\ldots,n$.

The *sixth and final step* of the derivation consists of dividing and multiplying the ith summand in (14) by r_i, and letting $n \to \infty$ in such a way that the maximum r_i in the set $\{r_1,\ldots,r_n\}$ goes to zero. Thus we first write (14) as:

$$N^*_r(z,\xi) = \sum_{i=1}^{n} \frac{N^*_{r_i}(x_i,\xi)}{r_i} T_{r-r_i'}(x_i,\xi) r_i$$

and apply the limit process $n \to \infty$, $\max\{r_i\} \to 0$, which formally yields, by means of (2), the following Riemann integral:

$$\boxed{N^*_r(z,\xi) = \int_0^r N_*(x',\xi) T_{r-r'}(x',\xi)\, dr'} \qquad (15)$$

which is the desired integral representation of the path radiance $N_r^*(z,\xi)$ for the path $\mathscr{P}_r(x,\xi)$ in terms of its path function N_* and beam transmittance function T_r. In the integral we have written:

$$\text{"z"} \quad \text{for} \quad x + r\xi$$

$$\text{"x'"} \quad \text{for} \quad x + r'\xi \quad ,$$

and have derived the formula under the conditions of constant index of refraction over $\mathscr{P}_r(x,\xi)$, and no internal sources over $\mathscr{P}_r(x,\xi)$. This completes the derivation.

The form of (15) is unchanged when the index of refraction is non constant, providing the appropriate form of the beam transmittance for such a case is used (see, e.g., Sec. 16 of Ref. [251]). The case of internal sources is covered by introducing the emission function N_η. (See, e.g., Sec. 19 of Ref. [251] and (1) of Sec. 5.8 below.)

It might be well to explicitly summarize what has transpired between (5) and (15). We started with (5) which is the precise formula for the observable path radiance $N_-(b)$ over path $\mathscr{P}_r(x,\xi)$ in an optical medium X. Here we have used one very convenient feature of the interaction principle, namely that it provides the exact mathematical rendition of all possible radiometric operations with radiance meters and other light measuring devices in natural optical media. Equation (5) summarizes how we can view the $N_-(b)$ as the response to incident flux $N_-(c)$, over the boundary c, of a subset C of X. This formula is then rearranged, using some of the laws of radiative transfer and the properties of the interaction operators we have deduced so far, into the classical representation (15) of the path function integral for path radiance. It should be emphasized that no intuitive observations were used as integral parts of the derivation between (5) and (15); however, they were occasionally invoked merely to make intuitively clear or to motivate the various steps in the derivation. The interaction principle has been stocked at the outset with all the intuition of the phenomenological point of view needed to establish radiative transfer theory. The principle thus provides the formal machinery for recovering all the known intuitions, and perhaps some new ones yet to be generated.

3.13 Derivation of Apparent-Radiance Equation

We turn next to the task of deducing from the interaction principle the well-known intuitive decomposition of the radiance of a distant object into two parts: the residual radiance transmitted from the object over the path of sight to the observer, and the path radiance generated over the extent of the path of sight.

The setting of Fig. 3.29 will serve for the first stage of the present discussion. We imagine the observer at point z of the path of sight $\mathscr{P}_r(x,\xi)$ in an optical medium X and that he is recording the radiance $N(z,\xi)$. For example the path of sight might begin with point x in a mountainside, or

on a lake bottom, or then again it could simply be a path segment beginning and ending in midair or midwater. Now for the purposes of the present derivation, we again imagine the cylindrical subset C of X about $\mathscr{P}_r(x,\xi)$ as axis. The set C is considered once again as a one-parameter optical medium with the usual direction conventions. The actual light field in X around $\mathscr{P}_r(x,\xi)$ is considered to be incident on the surface Y of C. As usual we imagine C isolated from X without disturbing the structure of the light field on and in C. This incident light field on Y can conveniently be thought of as incident on the bases a and b and the flank c of C. We are currently interested in the corresponding response radiance $N_-(b)$ emerging from the parameter surface with index b, as a result of the incident radiances over a, b, and c. In particular, we shall eventually concentrate on the value of $N_-(b)$ at point z in the direction ξ. For the present, the interaction principle yields three operators and the statement:

$$N_-(b) = N_-(a)\mathscr{S}(C;a,b) + N_+(b)\mathscr{S}(C;b,b) + N_-(c)\mathscr{S}(C;c,b). \quad (1)$$

Our goal is to obtain a limiting form of this statement as $C \to \mathscr{P}_r(x,\xi)$. The result will be the desired equation for apparent radiance. Toward this end, let us examine in turn each of the three terms in (1).

The limit of $N_-(a)\mathscr{S}(C;a,b)$ evaluated at z and ξ, as $C \to \mathscr{P}_r(x,\xi)$, is none other than the transmitted radiance $N_r^o(z,\xi)$ over the paths $\mathscr{P}_r(x,\xi)$. For we have:

$$\mathscr{S}(C;a,b) = T(a,b) \quad (2)$$

by definition (6) of 3.8 and the definition of T(a,b) given in Sec. 3.9. The observation now follows from (2) of Sec. 3.10.

The limit of $N_+(b)\mathscr{S}(C;b,b)$ evaluated at z and ξ, as $C \to \mathscr{P}_r(x,\xi)$, is zero. This may be seen by observing first of all that:

$$\mathscr{S}(C;b,b) = R(b,a) \quad (3)$$

by definition (6) of 3.8 and the definition of R(b,a) given in Example 2 of Sec. 3.9. The observation now follows from (5b) of Sec. 3.10.

The limit of $N_-(c)\mathscr{S}(C;c,b)$ evaluated at z and ξ as $C \to \mathscr{P}_r(x,\xi)$ is the path radiance $N_r(z,\xi)$, according to (1) of Sec. 3.12. Let us write,

$$\text{"}N_r(z,\xi)\text{"} \quad \text{for} \quad \lim_{C \to \mathscr{P}_r(x,\xi)} N_-(b)(z,\xi) \quad (4)$$

where $N_-(b)$ is as given in (1) and where "$N_-(b)(z,\xi)$" denotes as usual the value of $N_-(b)$ at z, and ξ. We call $N_r(z,\xi)$ the *apparent radiance* associated with $\mathscr{P}_r(x,\xi)$. Applying the limit operation $C \to \mathscr{P}_r(x,\xi)$ to each side of (1) and using the preceding observations we have:

$$N_r(z,\xi) = N_r^o(z,\xi) + N_r^*(z,\xi) \tag{5}$$

or in compact functional form:

$$N_r = N_r^o + N_r^* \quad .$$

Equation (5) is the desired equation for the apparent radiance associated with $\mathscr{P}_r(x,\xi)$.

Equation (5) is the exact statement of the primitive intuition we have about the apparent radiance of distant objects. Let us call the radiance $N_0(x,\xi)$ at the initial point of the path $\mathscr{P}_r(x,\xi)$ the *inherent radiance* of the field at x in the direction ξ. The point x may be on a tangible surface, or it may hang in empty space, mid air, or water. Furthermore, for a given fixed z and ξ the path $\mathscr{P}_r(x,\xi)$ may vary its length r without changing $N_0(z,\xi)$, but the associated $T_r(x,\xi)$ and $N_r^*(x,\xi)$ will vary accordingly. The apparent radiance of the field at x in the direction ξ as seen at z in the direction ξ is customarily thought of as consisting of two parts: the transmitted (or reduced or residual) inherent radiance $N_r^o(z,\xi) = N_0(x,\xi)T_r(x,\xi)$, and the path radiance (or "space light" or diffuse radiance) $N_r^*(z,\xi)$ scattered into the path between x and z. The preceding derivation has established a precise rendition of this intuitive judgment by means of a formal deduction from the interaction principle.

Equation (5) may be written in more detail using the results (4) of Sec. 3.10 and (15) of Sec. 3.12:

$$N_r(z,\xi) = N_0(x,\xi)T_r(x,\xi) + \int_0^r N_*(x',\xi)T_{r-r'}(x',\xi)\,dr' \tag{6}$$

To show the logical ancestry of the concepts in (6) in a striking way, the reader is invited to replace each of the five main terms in (6) by its definition as developed in this and in the preceding three sections. When these notational disguises are removed what should be left is an expression of the incident radiance distributions on a one-parameter cylindrical optical medium C being acted on by interaction operators and the result being embedded in a formidable battery of neatly interlocking limit processes. The interaction operators used for C are supplied by the interaction principle.

3.14 Derivation of the Volume Scattering Function

We come now to the derivation of the second inherent optical property of a general optical medium X. This property is summarized in the form of the volume scattering function which quantitatively describes at each point x of X the fraction of incident radiance $N(x,\xi')$ in the direction ξ' scattered in the generally different direction ξ, without change in wavelength. The volume scattering function together with the volume attenuation function derived in Sec. 3.11 are the two basic inherent optical properties for optical media used in radiative transfer theory.

Figure 3.32 depicts an arbitrary right-cylindrical subset C of the optical medium X. The axis of the cylinder is a straight line segment with initial point x, terminal point z, direction ξ, and length $r(\xi)$. Our present goal is to derive the volume scattering function from the appropriate interaction operator for C as supplied by the interaction principle.

Let a and b be the two bases of C and c be the flanking surface of C. Let $N_-(c)$ be a member of the set of incident radiance distributions on c. We assume C to be a one-parameter optical medium with the customary geometrical conventions. We are interested in the response radiances $N_-(b)$ emerging from the base b and such that $N_-(b)$ is of the same wavelength as $N_-(c)$. The interaction principle supplies an operator $\mathcal{S}(C;c,b)$ such that:

$$N_-(b) = N_-(c)\,\mathcal{S}(C;c,b) \quad . \tag{1}$$

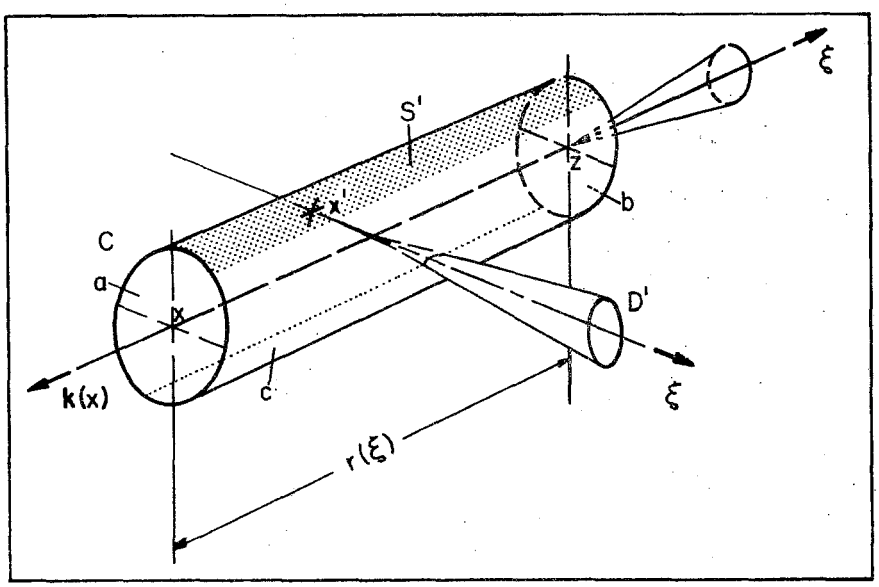


FIG. 3.32 Detail for the volume scattering function derivation.

The linear dimensions of C are now chosen so that the variation in size and shape of the radiance distributions from point to point over the volume of C is very small. This can always be done in natural light fields. The mathematical regularity property of $N_-(c)$ reflecting this natural phenomenon is that $N_-(c)$ be continuous and nearly constant with location over C. Furthermore we restrict the radiance distribution to be of uniform value over a fixed small solid angle D' containing a given direction ξ' distinct from ξ. Thus $N_-(c)$ is constant valued in D' and has zero value for directions outside D'. Again this is a feasible restriction, being operationally realizable in practice using appropriate optical shielding of C from appropriate parts of Ξ. The reason for these restrictions on $N_-(c)$ stems from the highly specialized geometrical structure of the volume scattering function which we are trying to characterize.

When these conditions are applied to (1) that equation takes on a considerably simpler form. Thus, evaluating $N_-(b)$ at z in the direction ξ we have from (1):

$$N(z,\xi) = \int_c \int_{\Xi_-(x')} N(x',\xi'')S_b(C;x',\xi'';x,\xi)\, d\Omega(\xi'')dA(x')$$

$$= N(S'',D') \int_{S'} \int_{D'} S_b(C;x',\xi'';x,\xi)\, d\Omega(\xi'')dA(x') \qquad (2)$$

Here S' is the part of the boundary of C which is bathed in the present radiance distribution. More precisely, S' is the set of all points x' of the flank c of C such that $\xi' \cdot \mathbf{k}(x') \leq 0$, where $\mathbf{k}(x')$ is the unit outward normal to c at x', and ξ' is the fixed direction of incidence of the radiance distribution $N_-(c)$ (see shaded region in Fig. 3.32). "N(S",D')" denotes the constant value of the radiance distribution $N_-(c)$ over S' and D'. S" is the projection of S' on a plane normal to ξ'. By (1), (3) and (4) of Sec. 3.8, $N(z,\xi)$ can be written as:

$$N(z,\xi) = N(S'',D')S(C;S',D';x,\xi)A(S')\Omega(D') \qquad (3)$$

and we are essentially back where we started from--except for one important difference: the radiance $N(z,\xi)$ is now clearly seen (by (1) of Sec. 3.12) to be the incipient path radiance for the path $\mathscr{P}_{r(\xi)}(x,\xi)$ under the present special lighting conditions. Hence by the agreement leading to (2) of Sec. 3.12 we are motivated to consider the ratio $N(z,\xi)/r(\xi)$, which by (3) is of the form:

$$\frac{N(z,\xi)}{r(\xi)} = N(S'',D')\left[S(C;S',D';x,\xi)\,\frac{A(S')}{r(\xi)}\right]\Omega(D') \qquad (4)$$

We have now reached the threshold of the definition of the volume scattering function. It remains only to write:

$$\text{"}\sigma(x;\xi';\xi)\text{"} \quad \text{for} \quad \lim_{\substack{C\to\mathscr{P}_{r(\xi)}(x,\xi)\\ r(\xi)\to 0\\ D'\to\{\xi'\}}} S(C;S',D';x,\xi)\,\frac{A(S')}{r(\xi)} \,. \tag{5}$$

The function σ which assigns to each point x and distinct pair of directions ξ',ξ the non negative number $\sigma(x;\xi';\xi)$ defined in (5) is called the *volume scattering function.* Since σ was derived from the operator $\mathscr{S}(C;c,b)$ with the kernel function $S(C;\cdot,\cdot,\cdot,\cdot)$ supplied by the interaction principle, it follows (as in the case of α in Sec. 3.11) that σ is an inherent optical property of X. The dimensions of σ are $\mathbf{L}_{\mathbf{r}}^{-1}\boldsymbol{\Omega}^{-1}$. This radial length and solid angle are associated with the direction ξ. The verbal interpretation of $\sigma(x;\xi';\xi)$ is obtained directly from (4) as follows: $\sigma(x;\xi';\xi)$ *is the scattered radiance generated, without change in wavelength, per unit length in the direction* ξ *by unit irradiance incident at* x *in the direction* ξ'.

An operational definition of σ, suitable for experimental use in natural optical media, is given in Sec. 13.6. An alternate approach to the volume scattering function is given in Example 1 of Sec. 3.17. The approach to (5) via empirical operations is given in Sec. 18 of Ref. [251].

Regularity Properties of σ

We have gone as far as we can in the present approach to the definition of σ: the value of σ at each x,ξ',ξ has been defined by means of (5). In order to go on and use σ in the mathematical theory of radiative transfer we must assume regularity properties of σ. In other words we must make explicit certain continuity properties of σ in order to use the calculus. Therefore σ will be assumed a piecewise continuous function of each of the arguments x,ξ', and ξ. Furthermore, we assume that for every point y in X there is a sphere $X_a(y)$ of center y and radius a such that for every connected subset C of $X_a(y)$:

$$S(C;x',\xi';x,\xi) = \sigma(x';\xi';\xi)\,\frac{r(\xi)}{A(S')} + o(r(\xi)) \tag{6}$$

for every pair of distinct directions ξ',ξ, and every pair of points x',x on the boundary of C. Here "$o(r(\xi))$" is the value of a function $o(\cdot)$ which has the property that $\lim_{x\to 0} o(x)/x=0$. $r(\xi)$ is the length of the shortest path through C with initial point x and in the direction $-\xi$. S' is that part of the boundary Y of C consisting of points x such that $\xi'\cdot\mathbf{k}(x)\leq 0$ where $\mathbf{k}(x)$ is the unit outward normal to Y at x. S' has a very simple geometric interpretation: shine a parallel beam of light on all of C along the direction ξ'. Then S' is that part of Y that is lit by the beam.

The Integral Representation of the Path Function

The final main step toward the derivation of the integral equation for radiance over a path $\mathscr{P}_r(x,\xi)$ can now be taken. This is the establishment of the integral representation of the path function N_* over $\mathscr{P}_r(x,\xi)$ using the volume scattering function σ along $\mathscr{P}_r(x,\xi)$.

Let $X_a(y)$ be a sphere of radius a about a point y on $\mathscr{P}_r(x,\xi)$. Let the boundary of $X_a(y)$ be b. Then:

$$N_-(b)(x,\xi) = \int_b \int_{\Xi_-(x')} N(x',\xi')S(X_a(y);x',\xi';x,\xi)\, d\Omega(\xi')dA(x') \tag{7}$$

Let the radius of $X_a(y)$ be such that (6) holds and $N_-(c)$ is independent of position over $X_a(y)$. Then (7) becomes:

$$N_-(b)(x,\xi) =$$

$$= \int_b \int_{\Xi_-(x')} N(y,\xi')\left[\sigma(y;\xi';\xi)\,\frac{r(\xi)}{A(S')} + o(r(\xi))\right] d\Omega(\xi')dA(x')$$

$$= \int_{\Xi} N(y,\xi')\left[\sigma(y;\xi';\xi)r(\xi) + A(S')o(r(\xi))\right] d\Omega(\xi') \quad .$$

Hence:

$$N_-(b)(x,\xi) =$$

$$= r(\xi) \int_{\Xi} N(y,\xi')\sigma(y;\xi';\xi)\, d\Omega(\xi') + o(r(\xi))\int_{\Xi} N(y,\xi')A(S')\, d\Omega(\xi') \; .$$

The radiance $N_-(b)(x,\xi)$ is thus seen to be nearly the path radiance for the path through $X_a(y)$ emerging at x in the direction y, and of length $r(\xi)$ inside the spherical region $X_a(y)$. Hence $N_-(b)(x,\xi)/r(\xi)$ is nearly the value of the path function at x in the direction ξ. By letting the radius a of $X_a(y)$ go to zero we have:

$$\lim_{a\to 0} \frac{N_-(b)(x,\xi)}{r(\xi)} = N_*(y,\xi) = \int_{\Xi} N(y,\xi')\sigma(y;\xi';\xi)\, d\Omega(\xi') \; .$$

Summarizing: For every point x in an optical medium, the relation between the path function $N_*(x,\cdot)$ at x, the radiance distribution $N(x,\cdot)$ at x and the volume scattering function $\sigma(x;\cdot,\cdot)$ is:

$$N_*(x,\xi) = \int_{\Xi} N(x,\xi')\sigma(x;\xi';\xi)\, d\Omega(\xi') \tag{8}$$

3.15 The Equation of Transfer for Radiance

All the pieces of the integral equation of transfer for radiance have now been carved out of the interaction principle. It remains only to assemble them into the desired statement.

Let $\mathscr{P}_r(x,\xi)$ be a path in an optical medium X. Let C be a right cylinder with $\mathscr{P}_r(x,\xi)$ as axis, and let C be a one-parameter optical medium with boundary Y composed of upper and lower parameter surfaces a and b, and flank c, as in Fig. 3.29. Let the incident radiance distribution over the outer surface Y of C be $N_-(Y)$. Then there is an interaction operator $\mathscr{J}(C;Y,b)$ such that:

$$N_-(b) = N_-(Y)\,\mathscr{J}(C;Y,b) \quad .$$

This is the interaction equation governing the response radiance of C over the base b. With this as a starting point, it was eventually reduced, as shown in Sec. 3.13, to the statement:

$$N_r(z,\xi) = N_r^o(z,\xi) + N_r^*(z,\xi)$$

where $N_r(z,\xi)$ is the apparent radiance of the field at z in the direction ξ. From (6) of Sec. 3.13 this can be written in the form:

$$N_r(z,\xi) = N_o(x,\xi)T_r(x,\xi) + \int_0^r N_*(x',\xi)T_{r-r'}(x',\xi)\, dr' \quad .$$

From (8) of Sec. 3.14 this becomes:

$$N_r(z,\xi) = N_o(x,\xi)T_r(x,\xi) + \int_0^r \left[\int_{\Xi} N(x',\xi')\sigma(x';\xi';\xi)\, d\Omega(\xi')\right] T_{r-r'}(x',\xi)\, dr' \tag{1}$$

which is the requisite *integral equation of transfer for radiance*. The subscripts "r" and "o" may be dropped wherever possible when it is convenient to divest (1) of all explicit ties with the path $\mathscr{P}_r(x,\xi)$. The result in such a case is:

$$N(z,\xi) = N(x,\xi)T_r(x,\xi) + \int_0^{r(z,\xi)} \int_\Xi N(x',\xi')\sigma(x';\xi';\xi)T_{r-r'}(x',\xi)\, d\Omega(\xi')dr' \,. \tag{2}$$

This is the mathematical form of the integral equation of transfer for the radiance function N defined on X×Ξ. Here $r(z,\xi)$ (abbreviated in (1) as "r") is the length of the shortest distance from z to the boundary of X along the direction $-\xi$ (see Fig. 3.33). Hence r depends on z and ξ and all are connected by the equation: $z = x + r\xi$. The distances r' and r are measured from the boundary point x to the point x' and the point z, respectively, along the direction ξ.

In mathematical discussions requiring the determination of N on X×Ξ, it is assumed that the volume attenuation function α and the volume scattering function σ are known, with appropriate regularity properties, and that $N(x,\xi)$ is given for each incident direction ξ at each point x on the boundary of X. As it stands, Equation (2) is an integral equation for N and applies to source-free optical media of arbitrary geometric shape with arbitrary α and σ and constant index of refraction n. With slight modifications (2) can be made to hold in media with internal sources by suitably including the emission function $N_\eta = N_e + N_s$ (where N_e represents true emission

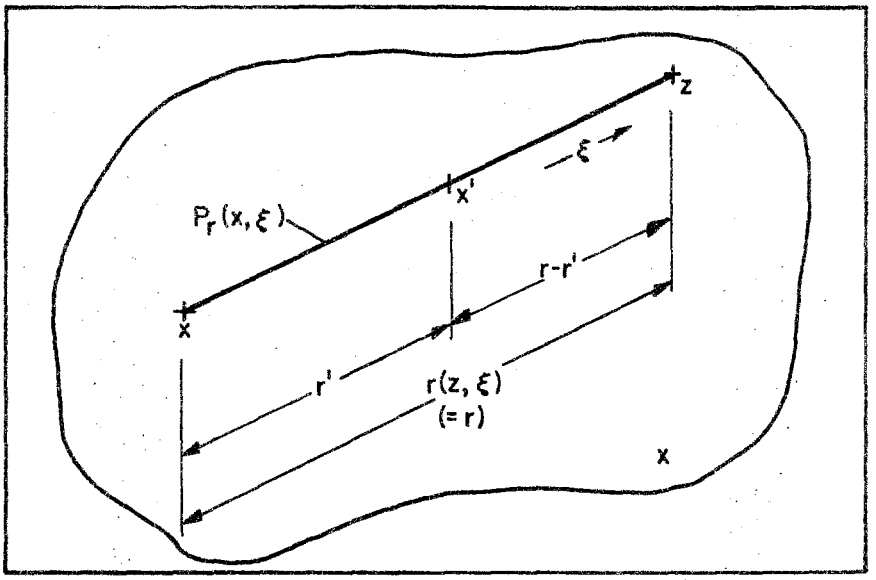


FIG. 3.33 The setting for the equation of transfer for radiance in a general medium X.

radiance and N_S transpectral scattered radiance), and variable index of refraction n. See, e.g., Sec. 21, Ref. [251], in which the boundary radiances and the optical properties can also be changing rapidly with time.

Steady State Equation of Transfer

We conclude the main discussion of this section by deriving from (1) the classical form of the integrodifferential equation of transfer. This equation is designed to describe the rate of change of $N_r(x,\xi)$ along $\mathscr{P}_r(x,\xi)$, with respect to r. Thus, holding x and ξ fixed, we shall let the path length r vary and then compute $dN_r(x,\xi)/dr$. By doing this, we in effect imbed $\mathscr{P}_r(x,\xi)$ in a family $\mathscr{P}$ of paths of the same x and ξ, and observe that the general functional form of (1) is invariant over the members of $\mathscr{P}$. In order to find $dN_r(x,\xi)/dr$, we can apply the operator d/dr to each of the two main terms of (1):

$$\frac{dN_o(x,\xi)T_r(x,\xi)}{dr} = N_o(x,\xi)\frac{dT_r(x,\xi)}{dr}$$

$$= N_o(x,\xi)\left[-\alpha(x,\xi)T_r(x,\xi)\right] \quad .$$

The second equality is based on (2) of Sec. 3.11. Next, we find (using (15) of Sec. 3.12 and (5) of Sec. 3.13):

$$\frac{dN_r^*(z,\xi)}{dr} = \frac{d}{dr}\int_0^r N_*(x',\xi)T_{r-r'}(x',\xi)\,dr'$$

$$= \int_0^r N_*(x',\xi)\frac{dT_{r-r'}(x',\xi)}{dr}\,dr' + N_*(z,\xi)$$

$$= -\alpha(x,\xi)\int_0^r N_*(x',\xi)T_{r-r'}(x',\xi)dr' + N_*(z,\xi)$$

$$= -\alpha(z,\xi)N_r^*(z,\xi) + N_*(z,\xi) \quad .$$

The second equality is obtained by means of the Leibniz rule of differentiating an integral with respect to a parameter. Collecting these results, we have:

$$\frac{dN_r(z,\xi)}{dr} = -\alpha(x,\xi)\left[N_r^o(z,\xi) + N_r^*(z,\xi)\right] + N_*(z,\xi) \quad .$$

Using (5) of Sec. 3.13, (8) of Sec. 3.14, and dropping "r" from "N_r" we finally obtain:

$$\boxed{\frac{dN(z,\xi)}{dr} = -\alpha(z,\xi)N(z,\xi) + \int_{\Xi} N(z,\xi')\sigma(z;\xi';\xi)\, d\Omega(\xi')} \quad (3)$$

which is the desired *integrodifferential form of the equation of transfer*. The settings in which (3) holds are also those for (1). The generalizations available for (3) are also those for (1).

It will be instructive for the reader to interpret each term of (3) by means of losses and gains of $N(x,\xi)$ due to attenuation and scattering. Thus the term $-\alpha(z,\xi)N(z,\xi)$ summarizes the rate of loss from $N(z,\xi)$ by means of attenuation, and the integral term $N_*(z,\xi)$ summarizes the rate of gain of $N_r(z,\xi)$ by means of scattering. A direct derivation of (3) by means of such loss-gain arguments is made, e.g., in Sec. 21 of Ref. [251]. We have chosen the present route to (3) to add to the accumulating evidence in this chapter that all of classical radiative transfer theory is derivable from the interaction principle. Having finally arrived at (3) in this manner we may essentially rest our case.

Time Dependent and Polarized Equations of Transfer

The chain of arguments in Sec. 3.10 to the present starting from the interaction principle and culminating in the equations of transfer (2) and (3), can now be repeated in all their essential steps but with more general formulations as the end result. For example, two immediate generalizations of (3) are obtained by considering time dependent radiances with time dependent optical properties, and by considering polarized radiance functions. The derivations of these generalizations of (3) will not be recorded here and are left as important (and nontrivial) exercises for interested students of the subject. (See Sec. 127 of [251].) The resultant time dependent equation of transfer is:

$$\boxed{\frac{1}{v}\frac{DN(z,\xi,t)}{Dt} = -\alpha(z,\xi,t)N(z,\xi,t) + \int_{\Xi} N(z,\xi',t)\sigma(x;\xi';\xi,t)\, d\Omega(\xi')} \quad (4)$$

where we have written:

$$\text{"}\frac{D}{Dt}\text{"} \quad \text{for} \quad \frac{\partial}{\partial t} + v\frac{d}{dr} \quad (5)$$

and where D/Dt, as in (4) is the usual (*mobile* or *substantial*) derivative operator along the path $\mathscr{P}_r(x,\xi)$. Furthermore, v is the speed of light at z, and we have written:

$$\text{"}\frac{d}{dr}\text{"} \quad \text{for} \quad \xi\cdot\nabla \quad , \quad (6)$$

where ∇ is the usual gradient operator of vector analysis.

The steady state equation of transfer for polarized radiance has the same Gestalt as (3); thus,

$$\boxed{\frac{dN(z,\xi)}{dr} = -\alpha(z,\xi)N(z,\xi) + \int_{\Xi} N(z,\xi')p(z;\xi';\xi)\, d\Omega(\xi')} \quad (7)$$

where $N(z,\xi)$ is as defined in Sec. 2.10 and $p(z;\xi';\xi)$ is the standard observable volume scattering matrix--the polarized counterpart to $\sigma(z;\xi';\xi)$. For a complete definition of $p(z;\xi';\xi)$ and related concepts, see Sec. 112 of Ref. [251]. Helpful background techniques for the derivation of (7) from the interaction principle are contained in Sec. 113, 114, and 126 of Ref. [251]. In particular, the work in Sec. 126 of Ref. [251] may serve as a prototype for the requisite derivation of (7). What is required for the general derivation is the consistent elevation of all concepts of the prototype derivation from the scalar to the vector level. The extension of (4) and (7) to the case of internal sources is accomplished by appending suitable source terms to each. For example, one may append $N_\eta(x,\xi,t)$ to (4), and $N\eta(x,\xi)$ to (7). A discussion of the relative consistency of (3) and (7) when the light field in a given medium is unpolarized is given in Sec. 13.11. In addition, the problem of the fidelity of (3) in the context of polarized light fields is raised and discussed in Sec. 13.11.

3.16 On the Integral Structure of the Interaction Operators

In this section we discharge a series of obligations which have been accumulating ever since Sec. 3.3. These concern the assertions that the interaction principle formally implies the existence of the various integral operators for reflectances and transmittances of surfaces, plane-parallel media, and the scattering properties for general optical media. Our purpose here is to cite and apply the appropriate mathematical theorems which, under suitable regularity conditions, yield the requisite integral operators arising under the use of the interaction principle along with their physically interesting kernels. By methodically applying these theorems to the various geometric settings encountered in Sec. 3.3, 3.6, 3.8 and 3.9, and many other settings, a veritable cornucopia of classical and novel integral operator formulas for radiative transfer phenomena is tapped and brought into formal existence by means of the interaction principle. The following discussion, while mainly mathematical in flavor, is written principally with the physicist in mind. Emphasis will constantly be on the physical or geometrical meanings of the terms discussed. As a result, mathematical rigor will not be of primary concern.

The Mathematical Prerequisites

In our everyday activities we frequently perform certain operations to determine various weights and measures of objects; measures of area, direction, length, volume, mass, and so on. When an attempt is made to formulate a mathematical representation of these operations and to summarize their salient characteristics, one arrives at the logical concepts of *linear functional* and *measure*.

Here is a simple example of a linear functional. Consider a straight flat sidewalk S one block long. A child's wagon loaded with objects is pulled over S from point a to b. As a result work is done in transporting these objects over S from a to b. The more objects in the wagon, the more force generally that must be exerted to push it from a to b. Suppose f(s) is the force parallel to the sidewalk that must be applied to a loaded wagon at point s on the sidewalk to keep it moving at constant speed from a to b. Then f is a force function defined on the sidewalk S. To f we can now assign the amount of work W(f) required to push its associated loaded wagon over S. For another wagon load, there is another function g and yet another amount W(g) of work done in transporting *that* wagonload over S. Hence W is a function which assigns to f or g a number, the work done in transporting over S the wagonload associated with f or g. W is called a *functional* because it acts on functions f, g and not numbers (as f and g do).

Now for typical sidewalks and most ordinary objects and wagons, the functional W is *linear* in the sense that:

$$\text{(i)} \quad W(f + g) = W(f) + W(g)$$

and

$$\text{(ii)} \quad W(cf) = cW(f)$$

where c is a real number. An instance of (i) arises, e.g., when we stack f's load of objects on top of g's load of objects and push the assembled load, whose force function then is f + g, along S. An instance of (ii) arises when several copies of an object with force function f are stacked on top of each other and transported along S. (In this case c would be an integer.)

Here is an example of a measure. Consider once again the sidewalk S described above. Every two points s, s' on the centerline of S determine a segment [s,s'] of S between them which has length say $l([s,s'])$. If [s,s'] and [t,t'] are two such separate segments along S, then $l([s,s'] \cup [t,t']) = l([s,s']) + l([t,t'])$, where "[s,s'] ∪ [t,t']" denotes the set of points on S consisting of the points of [s,s'] or [t,t']. Here l is an example of a measure--in this case a length measure: it assigns a number $l([s,s'])$ or "l(s,s')", for short, to the interval [s,s'] of points--the *length* of the interval [s,s'].

Here is an example of a linear functional closer to the present subject of radiative transfer theory. Let S be the same sidewalk as above, and let H be the function which

assigns to each point s of S, the irradiance on S at s. To this function H there is associated an amount P(H) of radiant flux falling on S. By simultaneous irradiation of S by skylight and street lamplight, etc., various irradiance functions H_1, H_2, etc. can be defined over S, and with resultant amounts $P(H_1)$, $P(H_2)$, etc. of radiant flux falling on S. P is therefore a linear functional in the sense that:

$$\text{(iii)} \quad P(H_1 + H_2) = P(H_1) + P(H_2)$$

$$\text{(iv)} \quad P(cH) = cP(H) \quad .$$

The pertinent measure in this example is area measure A of parts of S: and with the property that $A(S_1 \cup S_2) = A(S_1) + A(S_2)$ for two separate parts S_1 and S_2 of S.

Now the substance of the first of the two prerequisite theorems of present interest concerns the representation of linear functionals, such as W and P above, by means of an *integral with respect to a certain measure*. Thus, in the case of W, the theorem states that there is a measure μ over the sidewalk centerline such that for every force function f,

$$W(f) = \int_a^b f(s)\, d\mu(s) \quad . \qquad (1)$$

We know from elementary physical theory that μ is simply the length measure illustrated above; so

$$W(f) = \int_a^b f(s)\, dl(s) \quad . \qquad (2)$$

In the radiometric example, the theorem asserts that for every irradiance function H on S, there is a measure μ over the sidewalk such that:

$$P(H) = \int_S H(s)\, d\mu(s) \quad . \qquad (3)$$

From (6) of Sec. 2.4, we know that μ must be area measure A; so

$$P(H) = \int_s H(s)\, dA(s) \quad . \qquad (4)$$

The general version of the special cases (2), (4) may be stated after the following preliminaries are covered: Let S be a closed bounded subset in Euclidean n-space X_n, over which a family F(S) of real-valued continuous functions is defined. This setting is of sufficient generality to serve all our present purposes. Observe that F(S) is a vector space, (e.g., the sum of two functions in F(S) is again in F(S); the product of a member of F(S) by a real number is again in F(S)). By examples (i)-(iv) above we know what a linear functional L

is. A *positive* linear functional L on F(S) is a linear functional on F(S) such that $L(f) \geq 0$ if f is a non negative valued member of F(S) (both examples W and P above are examples of positive linear functionals). Then the requisite theorem (which is a general form of the Riesz representation theorem) goes as follows:

> *Theorem A. If* L *is a positive linear functional on* F(S), *then there exists a (Borel) measure* μ *on* S *such that for every* f *in* F(S)

$$L(f) = \int_S f(x)\, d\mu(x) \quad .$$

A complete general development of this theorem may be found, e.g., in Sec. 56 of Ref. [103].

The second theorem we shall need concerns measures which are absolutely continuous with respect to other measures. A measure μ is absolutely continuous with respect to a measure ν on a space X if $\mu(E) = 0$ whenever $\nu(E) = 0$ for every measurable subset E of X. This ostensibly forbidding-sounding description hides a very simple idea which may be illustrated as follows. To each subset X of ordinary Euclidean three-space assign the radiant energy content U(X) of that subset, as, e.g., we did in (14) of Sec. 2.7. Now it is obvious from the relation (14) of Sec. 2.7 that $U(X) = 0$ whenever $V(X) = 0$. That is, the radiant energy content of a set X of zero volume is zero. Using the present terminology we say that the radiant energy measure U is *absolutely continuous* with respect to the volume measure V. Other common examples may be found: mass measure, heat measure, etc., are absolutely continuous with respect to volume measure. Now the next theorem we have in mind says that for the case of U and V, for example, there is a function f on X such that:

$$U(X) = \int_X f(x)\, dV(x) \quad . \tag{5}$$

In other words, the theorem guarantees the existence of an energy density function f which when integrated over X gives the radiant energy content U(X) of X. In the other two cases cited we have the existence guaranteed of the mass density functions and heat density functions. Another way of writing f above is as:

$$\frac{dU}{dV}$$

pointing up the nature of f as a volume derivative of energy. We could then write the preceding integral as:

$$U(X) = \int_X \frac{dU}{dV}\, dV(x) \quad .$$

The function f above is a special instance of the general concept of a *Radon-Nikodym derivative* of one measure μ with respect to another ν. This derivative of μ exists whenever μ is absolutely continuous with respect to ν. The general statement is as follows:

> *Theorem B. Let* S *be a subset of Euclidean n-space* X_n *and let* ν *be a finite valued measure on* S. *Let* μ *be a finite valued measure on* S *which is absolutely continuous with respect to* ν. *Then there exists a finite valued function* f *on* S *such that*

$$\mu(E) = \int_E f(x)\, d\nu(x)$$

for every subset E *of* S.

The wording of this theorem, whose full version may be found in Sec. 31 of Ref. [103], has been deliberately simplified--references to fixed measure spaces and fixed families of measurable sets and functions have been suppressed and are to be implicitly understood. We are concerned here with only the essential conceptual content of Theorems A and B, what mathematical things they yield up for use, and their pertinence to the physical radiative transfer context. In the context of Theorem B we shall write

$$\text{"}\frac{d\mu}{d\nu}\text{"} \quad \text{for} \quad f \ . \tag{6}$$

The final theorem we shall need has been anticipated by the integral representation of U(X) above. Its statement goes as follows:

> *Theorem C. Let* S *be a subset of Euclidean n-space* X_n. *If* μ *and* ν *are finite valued measures on* S *and* μ *is absolutely continuous with respect to* ν *and* g *is a function on* S *such that* $\int_E g\, d\nu$ *is defined for every subset* E *of* S, *then:*

$$\int_E g\, d\mu = \int_E \left[g\, \frac{d\mu}{d\nu} \right] d\nu$$

for every subset E *of* S.

Again the wording of this theorem has been mercifully simplified so that one is encouraged to follow its physical applications below. Its unexpurgated and generalized version may be found in Sec. 32 of Ref. [103].

Interaction Operators for Surfaces

The preceding mathematical theorems will now be applied to the case of reflectance and transmittance operators for surfaces. Let us return to Fig. 3.3 of Sec. 3.3. We are interested in particular in the interaction properties of the surface Y depicted in part (a) of that figure. If a is any subset of Y, (e.g., a could be S of the figure) then the interaction method yields an operator $r_-(a)$ such that:

$$N_+(a) = N_-(a)r_-(a)$$

for the incident downward radiance distribution $N_-(a)$ on a and the reflected upward radiance distribution $N_+(a)$. Recall that $N_-(a)$ is a function which assigns to each point x on a and direction ξ in $\Xi_-(x)$ the incident (field) radiance $N_-(a)(x,\xi)$, called "$N_-(x,\xi)$" for short. The set $\Xi_-(x)$ is depicted in (b) of Fig. 3.3.

According to the interaction principle, $r_-(a)$ is a positive linear functional which, for a fixed choice x in a, and ξ in $\Xi_+(x)$, assigns to each $N_-(a)$ in the set of incident radiance functions on a the reflected radiance $N_+(a)(x,\xi)$ or "$N_+(x,\xi)$" for short. More specifically, the set S in Theorem A is the set $\Xi_-(x)$. F(S) is now the set of all incident radiance distributions $N_-(x,\cdot)$ at x on a. Hence, by Theorem A, there is a measure μ, depending on the current fixed choice of x and ξ, such that:

$$N_+(x,\xi) = \int_{\Xi_-(x)} N_-(x,\xi')\, d\mu(x;\xi';\xi) \quad . \qquad (7)$$

Here we have written the "μ" in Theorem A with sufficient notational paraphernalia so as to completely identify and keep track of it. The variable ξ' is like the x in the theorem. The variables x,ξ remind us that we have momentarily limited the values of $N_+(a)$ to x and ξ in a and $\Xi_+(x)$, respectively.

Now the measure $\mu(x;\cdot;\xi)$ just obtained from Theorem A is defined on $\Xi_-(x)$. That is, it assigns to subsets D' of $\Xi_-(x)$ a number whose geometric and physical significance becomes clear by letting $N_-(x,\cdot)$ be uniform valued with value 1 over subsets of $\Xi_-(x)$. For example, if D' is a subset of $\Xi_-(x)$ over which $N_-(x,\cdot)$ has value 1 and has value 0 outside D', then from (7):

$$N_+(x,\xi) = \int_{\Xi_-(x')} N_-(x,\xi')\, d\mu(x;\xi';\xi)$$

$$= \int_{D'} d\mu(x;\xi';\xi)$$

$$= \mu(x;D';\xi) \quad .$$

The construction of $\mu(x;D';\xi)$ in the present case is such (according to the proof of Theorem A) that for every D', if $\Omega(D') = 0$, then $\mu(x;D';\xi) = 0$ under all natural physical conditions. This means that a unit radiance distribution incident on surface a through solid angles of zero measure will induce zero radiance $N_+(x,\xi)$. Hence $\mu(x;\cdot;\xi)$ is to be absolutely continuous with respect to the solid angle measure Ω.

We are now ready to use Theorem B. The subset S is the same set just used in Theorem A. The measure ν is now solid angle measure Ω and μ is $\mu(x;\cdot;\xi)$. Hence Theorem B says that there is a finite valued function f--in this case call it "$r_-(x;\cdot;\xi)$", such that:

$$\mu(x;D';\xi) = \int_{D'} r_-(x;\xi';\xi)\, d\Omega(\xi')$$

for every subset D' of $\Xi_-(x)$. In other words:

$$r_-(x;\xi';\xi) = \frac{d\mu(x;\xi';\xi)}{d\Omega} \quad .$$

Theorem C completes the derivation when we observe that g is now to be $N_-(x,\cdot)$, μ is now $\mu(x;\cdot;\xi)$, and ν is Ω. We therefore have from (7):

$$N_+(x,\xi) = \int_{\Xi_-(x)} N_-(x,\xi')\, d\mu(x;\xi';\xi)$$

$$= \int_{\Xi_-(x)} N_-(x,\xi') r_-(x;\xi';\xi)\, d\Omega(\xi') \quad . \qquad (8)$$

Since x and ξ were arbitrary, (8) holds for every x in a and ξ in $\Xi_+(x)$, and the deduction of the form:

$$\int_{\Xi_-(x)} [\ ] r_-(x;\xi';\xi)\, d\Omega(\xi')$$

from the interaction principle is complete. The integral representations of the remaining three operators $r_+(a)$, $t_\pm(a)$ introduced in Sec. 3.3 are obtained similarly.

Interaction Operators for General Media

We go on now to consider a general optical medium X, bypassing the operators for plane-parallel media as being merely a special case of the present setting. Our goal is to derive the integral form of the linear operator $\mathscr{S}(X;a,b)$ in (6) of Sec. 3.8.

The interaction method yields a linear operator $\mathscr{S}(X;a,b)$ such that

$$N_+(b) = N_-(b)\,\mathcal{J}(X;a,b) \tag{9}$$

with the geometric conventions as defined in Sec. 3.8. The radiance distribution $N_-(b)$ is one of a family $F(S)$ of incident radiance distributions on S, where S is the set $a \times \Xi_-$ consisting of all pairs of points (x,ξ) with x in a and ξ in $\Xi_-(x)$. $F(S)$ is an instance of an incident set A_1 in the interaction principle and $N_+(b)$ is a member of the set B_1 of response functions. Hence $m = n = 1$ for the interaction principle yielding (9).

The interaction principle implies that $\mathcal{J}(X;a,b)$ induces a positive linear functional over $F(S)$ in the following way. We choose a *fixed* point x on b and *fixed* direction ξ in $\Xi_+(x)$ and consider the value $N_+(x,\xi)$ of $N_+(b)$ at (x,ξ). Theorem A then allows us to write:

$$\begin{aligned} N_+(x,\xi) &= \left(N_-(a)\,\mathcal{J}(X;a,b)\right)(x,\xi) \\ &= \int_S N_-(x',\xi')\, d\mu(X;x',\xi';x,\xi) \end{aligned} \tag{10}$$

where now $\mu(X;\cdot;x,\xi)$ is the measure on $a \times \Xi_-$, denoted by "S" in (10), whose existence is asserted in Theorem A. Again we have lavishly embellished μ of Theorem A with identifying variables: X,x,ξ. Further (x',ξ') in (10) acts like x in the theorem.

A glance at (10) shows that $\mu(X;\cdot;x,\xi)$ is a measure on S $(= a \times \Xi_-)$ and so the integral is a double integral over S. The geometric measure ν over S is the product of the area measure A over a and solid angle measure Ω over Ξ. If we let $N_-(a)$ be of uniform unit value over a subset S' of S and zero outside S', then (10) implies:

$$\begin{aligned} N_+(x,\xi) &= \int_{S'} d\mu(X;x',\xi';x,\xi) \\ &= \mu(X;S';x,\xi) \quad . \end{aligned} \tag{11}$$

As in the case of the reflectance operator for surfaces, we require, for obvious physical reasons, the radiance $N_+(x,\xi)$ to be zero when the measure ν of a subset S' of S is zero. That is, we require $\mu(X;S';x,\xi) = 0$ whenever $\nu(S') = 0$ and we shall assume that this is true.

Now we have:

$$\begin{aligned} \nu(S) &= \int_S d\nu = \int_S d(\Omega \times A) \\ &= \int_a \int_{\Xi_-(x')} d\Omega(\xi')\, dA(x') \quad . \end{aligned} \tag{12}$$

Recall that a general subset S' of S is a collection of ordered pairs (x',ξ') such that x' is in a subset a' of a and ξ' is in $\Xi_-(x')$; hence (12) is a special case of:

$$\nu(S') = \int_{a'} \int_{\Xi_-(x')} d\Omega(\xi')dA(x') \quad . \tag{13}$$

We now return to Theorem B which asserts that there is in this case a finite valued function f on $a \times \Xi_-$--call it $S(X;\cdot,\cdot;x,\xi)$--such that:

$$\mu(X;S';x,\xi) = \int_{a'} \int_{\Xi_-(x')} S(X;x',\xi';x,\xi)\, d\nu(x',\xi')$$

for every subset S' of S. In other words:

$$S(X;x',\xi';x,\xi) = \frac{d\mu(X;x',\xi';x,\xi)}{d\nu} = \frac{d\mu(X;x',\xi';x,\xi)}{d(\Omega \times A)} \quad .$$

Theorem C allows us to complete the derivation when we observe that g is to be $N_-(a)$, μ is now $\mu(X;\cdot;x,\xi)$ and ν is $\Omega \times A$, the Cartesian product of the solid angle and area measures. We therefore have from (10):

$$N_+(x,\xi) = \int_S N_-(x',\xi')\, d\mu(X;x',\xi';x,\xi)$$

$$= \int_a \int_{\Xi_-(x')} N_-(x',\xi')S(X;x',\xi';x,\xi)\, d\Omega(\xi')dA(x') \quad . \tag{14}$$

Since x and ξ were arbitrary, (14) holds for every x and ξ in b and $\Xi_+(x)$, respectively, and the deduction of $\mathscr{S}(X;a,b)$ in its integral operator form:

$$\int_a \int_{\Xi_-(x')} [\ ]S(X;x',\xi';x,\xi)\, d\Omega(S')dA(x') \quad ,$$

from the interaction principle, is complete.

Interaction Measures and Kernels

The features common to the two discussions just completed will now be summarized so as to extract the salient steps that must be generally taken in deducing from the interaction principle the requisite integral operator describing a given radiative transfer interaction.

Suppose a particular discussion using the interaction

method has reached the stage where the interaction principle yields for a subset S of an optical medium X the operator equation:

$$b = as \tag{15}$$

where a and b are elements of the incident and response classes A and B of radiometric functions, respectively (cf. Sec. 3.2). The functions a and b are quite general and may be either number-valued or vector-valued, or matrix-valued, etc., with domains of space, directional, frequency, time variables, singly or in combination. Let C be the domain of a. For the present discussion we shall view the class A explicitly as a set A(C) of continuous non negative valued functions on C. Similarly B is viewed explicitly as a set B(D) of non negative valued functions on some set D. (For example, in the case of the surface reflectance operators the medium X was three-space, S was a surface a, C was $\Xi_-(x)$, and D was $\Xi_+(x)$ for a fixed point x on surface a. A(C) was the set of all radiance distributions of the form $N_-(x,\cdot)$, and B(D) was the set of all radiance distributions of the form $N_+(x,\cdot)$.) Finally, the subset C is generally assumed to have some pertinent measure ν. (For example in the preceding discussion of the surface reflectance operators ν was Ω the solid angle measure on $\Xi_-(x)$.)

With these preliminaries established, the general method proceeds by selecting an arbitrary fixed point y in D. With this fixed y, we see that, the interaction operator s of (15) becomes a positive linear functional s(y). That is, if:

$$b_1(y) = a_1 s(y) \geq 0$$

and:

$$b_2(y) = a_2 s(y) \geq 0$$

then:

$$\alpha b_1(y) + \beta b_2(y) = (\alpha a_1 + \beta a_2) s(y) \geq 0 \quad ,$$

where α and β are non negative numbers, and where the $b_i(y)$ are images of the a_i under s(y), i = 1,2.

By Theorem A, there is a measure $\mu(S,\cdot,y)$ depending on the subset S of X and the point y in D such that:

$$s = \int_C [\ ]\, d\mu(S,\cdot,y) \tag{16}$$

so that for every y in D, (15) may be represented as:

$$b(y) = \int_C a(x)\, d\mu(S,x,y) \quad . \tag{17}$$

Let us say that $\mu(S,\cdot,y)$ has the *AC property (with respect to* ν*)* whenever ν and $\mu(S,\cdot,y)$ are such that if: $\nu(E) = 0$ then $\mu(S,E,y) = 0$ for every subset E of C. This property, it should be noted, is not asserted to hold universally. We view it as a *regularity property* whose validity must either be postulated (as an axiom, say) or demonstrated

in each situation. Thus the properties of each operator must be suitably stated so that the AC property holds (see remarks on the *Stages of the Interaction Method* in Sec. 3.18). The AC property of $\mu(S,\cdot,y)$ is the abstract version of all the S' and D' continuity statements made in Secs. 3.3, 3.6, and 3.8. The additive property of the measure $\mu(S,\cdot,y)$ is the abstract version of the S' and D' additivity properties stated in these sections. The initials "AC" stand for "absolute continuity".

The next step in the general method is to postulate (or verify) the AC property of $\mu(S,\cdot,y)$ so that we may go on formally to Theorem B which asserts that there exists a function $K(S,\cdot,y)$ on C such that:

$$\mu(S,E,y) = \int_E K(S,x,y)\, d\nu(x) \quad , \tag{18}$$

for every subset E of C. For example, $S(X;x',\xi';x,\xi)$ defined in the preceding example on general media is the special case of $K(S,x,y)$ for a general optical medium X. In the case of $S(X;x',\xi';x,\xi)$, "x" (in (18)) plays the role of "(x',ξ')", and "y" plays the role of "(x,ξ)", and of course "S" (in (18)) plays the role of X. Hence we have:

$$K(S,x,y) = \frac{d\mu(S,x,y)}{d\nu(x)} \quad . \tag{19}$$

We call $K(S,\cdot,y)$ the *interaction kernel* for the subset S of X and $\mu(S,\cdot,y)$ the *interaction measure* for S.

An application of Theorem C then completes the general method by allowing us to write:

$$b(y) = \int_C a(x)\, \frac{d\mu(S,x,y)}{d\nu(x)}\, d\nu(x) \quad .$$

That is, for every y in D, (15) may now be written:

$$b(y) = \int_C a(x)K(S,x,y)\, d\nu(x) \tag{20}$$

so that:

$$s = \int_C [\ ]K(S,x,y)\, d\nu(x) \quad . \tag{21}$$

Equation (21) is the requisite integral representation of the interaction operator s, associated with the subset S of the optical medium X.

3.17 Further Examples of the Interaction Method

We conclude the illustrations of the interaction method in this chapter with a brief listing of some further important radiative transfer phenomena which can be methodically subsumed under the interaction principle. We begin with two concepts which we have already studied: the path function and the path radiance (cf. Sec. 3.12). Now we approach these familiar concepts in perhaps the most interesting way of all.

Example 1: The Path Function Operator

The equation connecting a radiance distribution $N(x,\cdot)$ at a point in an optical medium X and the associated path function distribution $N_*(x,\cdot)$ at the same point in X was finally attained in Sec. 3.14 after a relatively laborious struggle which first had to bring into the light of day the concept of volume scattering function. We now connect $N_*(x,\cdot)$ and $N(x,\cdot)$ in an alternate and less arduous way. However, what we gain in elegance and mathematical insight by taking the present approach, we lose in physical meaning. The earlier route taken, however long and detailed, has the virtue that it suggests operational means of measuring σ *in situ*, i.e., within an optical medium. The present approach has the virtue of showing the logical structure of the relation between $N_*(x,\cdot)$, $N(x,\cdot)$, and $\sigma(x;\cdot;\cdot)$, and does so with unprecedented clarity.

Let X be an optical medium and let the present subset S of X be a singleton $\{x\}$, i.e., a one-point subset of X. Hence we will be using the special point-level interpretation of the interaction principle (re: Sec. 3.2). Let the set A_1 of incident radiometric functions on $\{x\}$ be radiance distributions like $N(x,\cdot)$. Let the set B_1 of response functions be the path functions like $N_*(x,\cdot)$ and defined using (3) of Sec. 3.12. Then $m = n = 1$ in the interaction principle of Sec. 3.2, and there exists an interaction operator $\mathbf{R}$ such that

$$N_*(x,\cdot) = N(x,\cdot)\mathbf{R} \quad . \tag{1}$$

In the terminology developed in the closing paragraph of Sec. 3.16, in particular with reference to (15) of Sec. 3.16, b is now $N_*(x,\cdot)$, a is now $N(x,\cdot)$, and s is now $\mathbf{R}$. The sets C and D are each now the unit sphere Ξ, and ν is solid angle measure Ω on Ξ. $\mathbf{R}$ gives rise for each fixed ξ in Ξ to a positive linear functional,so that for a particular fixed ξ in D (= Ξ) we obtain by Theorem A of Sec. 3.16, an interaction measure $\mu(x;\cdot;\xi)$ such that

$$N_*(x,\xi) = \int_{\Xi} N(x,\xi')\, d\mu(x;\xi';\xi) \quad . \tag{2}$$

$\mu(x;\cdot;\xi)$ clearly has the AC property with respect to Ω. Hence by Theorem B of Sec. 3.16 there is an interaction kernel $\sigma(x;\cdot;\xi)$ such that for every subset D' of Ξ:

$$\mu(x;D';\xi) = \int_{D'} \sigma(x;\xi';\xi)\, d\Omega(\xi') \quad . \tag{3}$$

Equation (2) corresponds to (17) of Sec. 3.16; Equation (3) corresponds to (18) of Sec. 3.16, in which y is now ξ. In the present instance the interaction kernel K for {x} is the volume scattering function σ. The present specific instance of (20) of Sec. 3.16 is obtained by means of Theorem C of Sec. 3.16:

$$N_*(x,\xi) = \int_{\Xi} N(x,\xi)\sigma(x;\xi';\xi)\, d\Omega(\xi') \tag{4}$$

and which is to be compared with (8) of Sec. 3.14. Thus we have:

$$\boxed{R = \int_{\Xi} [\ ]\sigma(x;\xi';\xi)\, d\Omega(\xi')} \quad . \tag{5}$$

We call **R** the *path function operator.*

Example 2: The Path Radiance Operator

The equation which represents the path radiance $N_r^*(z,\xi)$ over a path $\mathscr{P}_r(x,\xi)$ in an optical medium in terms of the path function $N_*(\cdot,\xi)$ defined over $\mathscr{P}_r(x,\xi)$ was obtained in (15) of Sec. 3.12 after some rather delicate analysis but in which each step was completely meaningful physically. We now establish (15) of Sec. 3.12 using the interaction principle in a radically different way; one that exhibits the logical interrelation of these concepts with a minimum of direct appeal to physical meaning.

We begin by choosing a path $\mathscr{P}_r(x,\xi)$ in an optical medium X (see Fig. 3.33). This path is a one-dimensional subset of X, and so we will be using the line-level interpretation of the interaction principle (Sec. 3.2). We let A_1 be the set of all incident radiometric functions on $\mathscr{P}_r(x,\xi)$, in this case all path functions like $N_*(\cdot,\xi)$. We let B_1 be the set of all path radiances like $N_r^*(z,\xi)$, where $z = x + \xi r$ (see Fig. 3.33). Then the interaction principle yields an interaction operator T such that:

$$N_r^*(z,\xi) = N_*(\cdot,\xi)T \quad . \tag{6}$$

In the terminology of Sec. 3.16, in particular (15) of Sec. 3.16, b is now $N_r^*(z,\xi)$, a is now $N_*(\cdot,\xi)$, and s is now **T**. The set C is $\mathscr{P}_r(x,\xi)$, and the set B(D) is the set of all path radiance values (non negative real numbers) $N_r^*(z,\xi)$ on the set $D = \{(z,\xi)\}$. The measure ν is now the length measure 1 along $\mathscr{P}_r(x,\xi)$.

T is a positive linear functional, so that we obtain directly from Theorem A of Sec. 3.16 an interaction measure $\mu(\mathscr{P}_r(x,\xi),\cdot,z)$ such that:

$$N_r^*(z,\xi) = \int_{\mathscr{P}_r(x,\xi)} N_*(x',\xi)\, d\mu\{\mathscr{P}_r(x,\xi),x',z\} \quad . \tag{7}$$

$\mu(\mathscr{P}_r(x,\xi),\cdot,z)$ clearly (i.e., on physical grounds) has the AC property with respect to $l(\cdot,z)$ the measure which assigns to x' in $\mathscr{P}_r(x,\xi)$ the distance $l(x',z) = r'$ between x and x'. Hence by Theorem B of Sec. 3.16 there is an interaction kernel $K(\mathscr{P}_r(x,\xi),x,z)$ such that:

$$\mu\{\mathscr{P}_r(x,\xi),E,z\} = \int_E K\{\mathscr{P}_r(x,\xi),x',z\}\, dl(x',z) \tag{8}$$

for every subset E of $\mathscr{P}_r(x,\xi)$. In the present instance, the interaction kernel is none other than the beam transmittance function such that:

$$T_{r-r'}(x',\xi) = K\{\mathscr{P}_r(x,\xi),x',z\} \tag{9}$$

so that in particular:

$$K\{\mathscr{P}_r(x,\xi),z,z\} = 1 \quad . \tag{10}$$

Being able to write down (9) at this time is of course the result of hindsight. Were we Martians developing radiative transfer theory for the first time, having been given only the interaction principle and none of the developments of Sec. 3.10, we would retain and work only with "$K(\mathscr{P}_r(x,\xi),x',z)$" and perhaps eventually deduce in our own way the multiplicative, identity and contraction properties of the beam transmittance function (Sec. 3.10), the differential governing it, and finally the volume attenuation function.

Finally, by means of Theorem C of Sec. 3.16 we come to:

$$\begin{aligned} N_r^*(z,\xi) &= \int_{\mathscr{P}_r(x,\xi)} N_*(x',\xi)K\{\mathscr{P}_r(x,\xi),x',z\}\, dl(x',z) \\ &= \int_0^r N_*(x',\xi)T_{r-r'}(x',\xi)\, dr' \quad , \end{aligned} \tag{11}$$

by virtue of (9). Hence

$$\boxed{T = \int_0^r [\ ]T_{r-r'}(x',\xi)\, dr'} \tag{12}$$

We call T the *path radiance operator* associated with the path $\mathscr{P}_r(x,\xi)$.

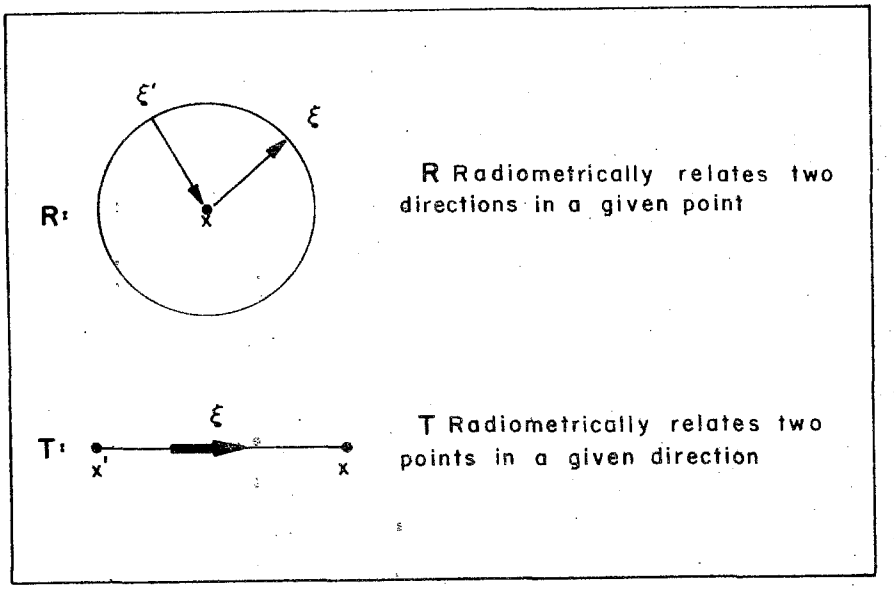


FIG. 3.34

Before closing this example we wish to point out an interesting geometrical duality between the path radiance operator T and the path function operator R. This duality is best described in ideographic form in Fig. 3.34. In other words, if we interchange the words "direction" and "point" in the description of R, we obtain that of T, and conversely.

Example 3: The Volume Transpectral Scattering Operator

We now formulate the definition of an important extension of the volume scattering function--the volume *transpectral* scattering function. As its name implies this new scattering function relates incident radiance of frequency ν' at a point x to resultant scattered radiance at x of frequency ν. In short we shall now consider scattering of flux not only from one direction to another, but also from one frequency to another.

The use of the interaction method has been illustrated often enough by now so that it will suffice in this and the remaining examples to be somewhat less detailed in the explanations.

Stage one: Construct a function $N_S(x,\cdot,\nu)$ (the transpectral path function) at point x in medium X such that its value at ξ is the radiance of frequency ν generated by inelastic (transpectral) scattering at x, of an incident radiance distribution $N(x,\cdot,\nu')$ of frequency ν'. This stage corresponds to the definition of the path function

using (3) of Sec. 3.12.

Stage two: Proceeds exactly as in Example 1 above. Instead of (1), we now have from the interaction principle:

$$N_s(x,\cdot,\nu) = N(x,\cdot,\nu')\mathbf{R} \quad .$$

The resultant integral representation is:

$$\mathbf{R} = \int_{\Xi} [\ ]\sigma(x;\xi';\xi;\nu',\nu)\, d\Omega(\xi') \quad . \tag{13}$$

The interaction kernel in (13) is called the *volume transpectral scattering function*. This function is a proper generalization of the monochromatic volume scattering function as can be seen by setting $\nu' = \nu$.

Alternate Stage two: Proceeds analogously to Example 1 but now the incident radiance distributions have *two* free variables ξ' and ν', so that the principle yields the operator equation:

$$N_s(x,\cdot,\nu) = N(x,\cdot,\cdot)\hat{\mathbf{R}} \quad .$$

The resultant integral representation is:

$$\hat{\mathbf{R}} = \int_{\Xi}\int_{\Lambda} [\ ]\hat{\sigma}(x;\xi';\xi;\nu',\nu)\, d1(\nu')d\Omega(\xi') \tag{14}$$

where Λ is the spectrum. The operator $\hat{\mathbf{R}}$ in (14) is called the *standard transpectral scattering operator.*

The *undecomposed transpectral scattering operator* combines $\hat{\mathbf{R}}$ of (14) and $\mathbf{R}$ of (5):

$$\int_{\Xi}\int_{\Lambda} [\sigma(x;\xi';\xi;\nu)\delta(\xi-\xi') + \hat{\sigma}(x;\xi';\xi;\nu',\nu)]\, d1(\nu')d\Omega(\xi') \tag{15}$$

where δ is the Dirac delta function. The dimensional distinction between $\hat{\sigma}$ and the two σ's should be noted. We shall also call $\hat{\sigma}$ the *volume transpectral scattering function.* Operator (13) is useful when only a finite number of discrete frequency transitions are considered. Operator (14) is a natural choice when continuous frequency transitions are considered.

Miscellaneous Examples

We leave the applications of the interaction principle open-ended at this stage and merely list some further possibilities for consideration by interested students of the subject:

(i) Interaction Operators for Internal Sources (cf. (37) of Sec. 3.9).

(ii) Path Function Operator for Polarized Radiance (and hence the genesis of the volume scattering matrix --see Sec. 112 of Ref. [251]).

(iii) The Path Radiance Operator for Polarized Radiance (and hence the genesis of the beam transmittance matrix--see Sec. 112 of Ref. [251]).

(iv) Time Dependent Operators--the time dependent versions of all the kinds of operators considered so far. (See (4) of Sec. 3.15 and Sec. 127 of Ref. [251]).

(v) The Photometric Operators $Y(\mathcal{Q},M)$, $Z(\mathcal{Q},M)$. (See (13) of Sec. 2.12 and (1) of Sec. 2.13.)

(vi) The Operator $C(x)$. (See Sec. 2.11.)

(vii) The Operators of the Mueller Phenomenological Algebra (Refs. [192], [193], [194], and Sec. 137 of Ref. [251]).

3.18 Summary of the Interaction Method

The interaction method is a method of formulating radiative transfer problems by means of the interaction principle. After some preliminary examples, the steps of the method were listed following Example 2 of Sec. 3.4. The method was then extensively applied throughout the remaining part of the chapter. In this section we summarize the method as developed throughout this chapter and include the steps of Sec. 3.17 leading to the integral representation of the interaction operators used in the method. The section concludes with some observations on the relative roles played by the interaction principle in this work and in Ref. [251].

Summary of the Interaction Method

There are three main stages of the Interaction Method. Let X be an optical medium and S be a subset of X. Then:

Stage I

(i) *Isolate the subset* S *of the optical medium. If* S *is concave decide how* S *is to be convexified (Sec. 3.8).*

(ii) *Enumerate the incident radiometric quantities* a_i *on* S. *This determines* A_j, $i = 1,\ldots,m$. *(Sec. 3.2)*

(iii) *Enumerate the requisite response radiometric quantities* b_j *on* S. *This determines* B_j, $j = 1,\ldots,n$. *(Sec. 3.2)*

(iv) *Enumerate the* mn *operators* s_{ij}, $i = 1,\ldots,m$; $j = 1,\ldots,n$, *supplied by the interaction principle (Sec. 3.2).*

(v) *Write the interaction equations:*

$$b_j = \sum_{i=1}^{m} a_i s_{ij}$$

for $j = 1,\ldots,n$.

(vi) *Append to (v) any auxiliary equations connecting various chosen* a_i *and* b_j *so that it is possible to algebraically solve the system of* n *functional equations in (v) for the requisite response functions* b_j. *Invariably, these auxiliary equations may be based on one or the other of the following radiometric laws:*

(a) *The radiance invariance law over a path in a vacuum (Sec. 2.6).*

(b) *The equality of field and surface radiance distributions at a point in a general optical medium (Sec. 2.5).*

Stage II *If the structure of* A,B, *and* S *indicate the possibility of an integral representation of the interaction operators* s_{ij}, *then use the technique of the interaction measures and interaction kernels of Sec. 3.16 to obtain:*

$$s_{ij} = \int_C [\ ] K_{ij}(x,y)\, dv(x)$$

for $i = 1,\ldots,m$: $j = 1,\ldots,n$.

Stage III *Determine by means of suitable functional relations the explicit structure of the* mn *interaction kernels* K_{ij}, *if they exist, and use the results in (vi) of Stage I to obtain a solution of the interaction problem.*

Remarks on the Stages of the Interaction Method

Stage I was fully illustrated in the present chapter; however, some further aspects of the details of Stage II and Stage III beyond those covered in Sec. 3.17, remain to be observed. As regards Stage III, the interaction kernels arising in homogeneous plane-parallel media and their governing functional relations have been exhaustively studied by Chandrasekhar (Ref. [43]). Further functional relations were given in Refs. [13], [14], by Bellman and Kalaba for inhomogeneous media. The functional relations for the complete set of four interactions kernels in non homogeneous one-parameter media, (i.e., the four reflectance and transmittance functions R and

T) were introduced and derived in Refs. [233] and [234]. (See also Sec. 7.1.) The functional relations governing the interaction kernel for the general operators $\mathscr{A}(X;a,b)$ were derived in Ref. [251]. The general procedures for the solution of the functional relations governing the operators R(a,b), T(a,b), R(b,a), T(b,a) for general one-parameter optical media for $\mathscr{A}(X;a,b)$ are given in Chapter 7 of this work.

Now that the conceptual structure of radiative transfer theory has been elucidated by the interaction principle, and its mathematical foundations established (ref. [251]) it remains to solve the important mathematical problems of modern radiative transfer theory centering around the functional relations governing the interaction kernels (see problem VIII, Sec. 141 of Ref. [251]).

One final remark on Stage II must be made. This concerns the AC property of an interaction measure. If the AC property is valid for a given interaction measure, then the interaction kernel of that measure is, according to Theorem B of Sec. 3.16, the Radon-Nikodym derivative of that measure. In this regard the development of interaction kernels will be occasionally simplified if the transmittance-type operators are decomposed into their residual and diffuse parts, i.e., into parts which, respectively, describe radiant flux which has not been scattered (i.e., beam transmitted) and which has been scattered. It turns out that transmittance operators for *diffuse* flux always have the AC property. (Reflectance-type operators generally have the AC property outright since they describe only diffuse flux.) The basis for these remarks rests in Ia, Ib of Sec. 23, Ref. [251], which, in the present work, may be taken as basic postulated regularity properties of interaction kernels. A model for this procedure of decomposing operators will be found in Sec. 7.1. The decomposition of the light field, which is a natural prerequisite to the decomposition of interaction operators, can easily be done in general since the concept of scattered and non scattered radiant flux is now rigorously definable by means of the path function and path radiance operators of Examples 1 and 2 of Sec. 3.17. This decomposition will be studied as a matter of course in Chapter 5. The net result of Stage II of the interaction method will be that the diffuse component of a transmittance operator (rather than the undecomposed operator) will be passed on to Stage III for the determination of its kernel. The prototype of this procedure may be found in Ref. [43], and in Refs. [234], [235].

The Interaction Method and Quantum Theory

We append here some final observations on the general methodology of the interaction method, an observation which will point up some points of similarity between the interaction method and two basic methods of solving dynamical problems in classical and modern physics. The observation is designed to be of especial interest to physicists, rather than radiative transferists per se. Nevertheless, since radiative transfer is ultimately derivable from quantum mechanics, the latter workers may peruse the following with some profit.

The first point of similarity was noted in the discussion following Example 2 of Sec. 3.4 where a comparison was made between the Newtonian laws of motion and the interaction principle, and note was made of the applicability of the method to linear hydrodynamics and general wave guide phenomena. We need not repeat it here. The second point of comparison appears to be even deeper than the first when we note the similarity between the interaction method and the formulation of the quantum mechanics of many-state atomic systems. To facilitate the comparison, the reader may consult, e.g., [92]. Here are the parallel correspondents: to an atomic or molecular system we pair an optical medium (step (i)). To the various base states of the atomic system we pair the sets of incident and response radiometric quantities (steps (ii), (iii)). To the Hamiltonian matrix of the atomic system we pair the set (s_{ij}) of interaction operators (step (iv)). To the transition probability equation (the linear superposition of amplitude functions) we pair the interaction equation (step (v)). To the finding of either the Hamiltonian matrix (using conservation laws and auxiliary physical arguments) or S-matrix, we pair the finding of the interaction operators (step (vi) and Stages II and III). The mystery of this remarkable similarity between the quantum mechanical and radiative transfer formalisms is only apparent and is resolved by noting that *each discipline is founded (for its own particular experimental reasons) on a set of linear superposition principles. Hence both methodologies come under the single unifying framework of vector space theory.* The salient difference between the two formalisms is that the possibility of interference of amplitudes exists in quantum theory, whereas interference of radiant fluxes is ruled out by fiat from radiative transfer theory (cf. Sec. 2.2). In the preceding point by point parallelism of the mathematics of quantum theory and radiative transfer theory lie the keys to the solutions of the basic problems II, and IV in Sec. 142 of Ref. [251]. It may be noted in passing that the applications of the linear interaction principle to quantum mechanics, linear hydrodynamics, acoustics, and electromagnetic theory, e.g., introduce complex-valued interfering amplitudes, and on this level the theoretical and numerical methodologies of all these fields are strikingly alike.

The Interaction Principle as a Means and as an End

Throughout this chapter there have been several occasions to refer to the developments of radiative transfer theory in Ref. [251] and in particular to the interaction principle in that work. A few words may be in order to help place in perspective the relative roles of the interaction principle in these two works.

The interaction principle in Ref. [251] was the end of a long series of generalizations and abstractions starting mainly with the work of Schuster, on through classical principles of invariance of Ambarzumian and Chandrasekhar, and up through the principle of invariant imbedding of Bellman and Kalaba as applied to transport phenomena, and finally on to the invariant imbedding relation, and the interaction principle

itself. It was shown, in particular, how all these principles could be deduced from the classical equation of transfer, and how the equation of transfer could itself be viewed as a local form of the principles of invariance. *Hence, in a word, the interaction principle was viewed in Ref. [251] as an end of a set of long conceptual and deductive trails, the main trail starting from Schuster's initial insight in 1905.* Thus in [251] the roots of interaction principle were established in the classical origins of the subject along with electromagnetic and axiomatic bases of the principle. *With this in mind we have taken the alternate view in the present chapter that the interaction principle is a basic means of formulating radiative transfer theory, a single working principle from which the salient algebraic structures of the theory may be deduced.* The thirty-eight enumerated examples throughout this chapter, starting in Sec. 3.4 and ending in Sec. 3.17, have shown that the interaction principle can indeed be used as a starting point for the construction of the principles of invariance on all types of three-dimensional media, the various classical interreflection problems of surfaces, the beam transmittance function for paths, the classical attenuation and scattering functions of the media used in the equation of transfer, and the equation of transfer itself.

Conclusion

In sum, then, the work of the monograph [251] constituted a necessary prerequisite for the establishment of the interaction principle. The present work no longer views the interaction principle as an end of research but rather as a means of application and a source of new research in radiative transfer theory and general linear transport theories (even beyond radiative transfer, as in hydrodynamics, acoustics, e. m. wave propagation, etc.). The first application of the interaction principle was to the development of the discrete-space theory of radiative transfer in Ref. [251]. These applications are continued in this chapter, and the following chapters of the present work.

3.19 Bibliographic Notes for Chapter 3

The interaction principle as given in Sec. 3.2 was first formulated in Ref. [251], the end result of an extended series of generalizations. A historical sketch of the evolution of the main lines of radiative transfer *theory* (not its manifold applications) which are pertinent to the interaction principle is given cumulatively in the bibliographic notes for the chapters in Ref. [251]. The formulation of the interaction method, as summarized in Sec. 3.18, is new.

BIBLIOGRAPHY FOR VOLUME II

1. Ambarzumian, V.A., "Diffuse reflection of light by a foggy medium," Compt. rend. (Doklady) Acad. Sci. U.R.S.S. 38, 229 (1943).

2. Ambarzumian, V.A., "On the problem of the diffuse reflection of light," J. Phys. Acad. Sci. U.S.S.R. 8, 65 (1944).

3. *American Institute of Physics Handbook* (McGraw-Hill, New York, 1957).

4. American Standards Association, "Nomenclature for radiometry and photometry (Z58.1.1-1953)," J. Opt. Soc. Am. 43, 809 (1953).

8. Barber, E., *Radiometry and Photometry of the Moon and Planets*, (Literature Search No. 345, Nat. Aero. and Space Adm., Contract NASw-6, Jet Propolusion Laboratory, California Institute of Technology, Pasadena, September 1961).

13. Bellman, R., and Kalaba, R., "On the principle of invariant imbedding and propagation through inhomogeneous media," Proc. Natl. Acad. Sci. 42, 629 (1956).

14. Bellman, R., and Kalaba, R., "On the principle of invariant imbedding and diffuse reflection from cylindrical regions," Proc. Natl. Acad. Sci. 43, 514 (1957).

26. Boileau, A.R., *Atmospheric Optical Measurements During High Altitude Balloon Flight*, parts I, II, III, IV (Scripps Inst. of Ocean. Refs. 59-32, 61-1, 61-2, 61-3, University of California, San Diego, December 1959, July 1961, July 1961, December 1961).

36. Burton, H.E., "The optics of Euclid," J. Opt. Soc. Am. 35, 357 (1945).

42. Chandrasekhar, S., "On the radiative equilibrium of a stellar atmosphere, II," Astrophys. J. 100, 76 (1944).

43. Chandrasekhar, S., *Radiative Transfer* (Oxford, 1950).

49. Committee on Colorimetry (Optical Society of America), "The psychophysics of color," J. Opt. Soc. Am. 34, 245 (1944).

50. Committee on Colorimetry (Optical Society of America), *The Science of Color* (Thomas Y. Crowell Co., New York, 1953).

51. Condas, G.A., "Maximum spectral luminous efficiency," J. Opt. Soc. Am. 54, 1168 (1964).

90. Fano, V., "Gamma ray attenuation, part 1, basic processes," Nucleonics 11, 8 (August 1953).

92. Feynman, R.P., Leighton, R.B., and Sands, M., *The Feynman Lectures on Physics* (Addison-Wesley Pub. Co., Reading, Mass., 1965), vol. III, Quantum Mechanics.

98. Gershun, A., "The Light field," (trans. by P. Moon and G. Timoshenko), J. Math. and Phys. 18, 51 (1939).

103. Halmos, P.R., *Measure Theory* (D. Van Nostrand, New York, 1950).

128. Johnson, F.S., "The solar constant," J. Met. 11, 431 (1954).

130. Jones, R.C., "Terminology in photometry and radiometry," J. Opt. Soc. Am. 53, 1314 (1963).

150. Landé, A., *From Dualism to Unity in Quantum Physics* (Cambridge Univ. Press, 1960).

151. Landé, A., *New Foundations of Quantum Mechanics* (Cambridge Univ. Press, 1965).

153. Le Grand, Y., *Light, Colour and Vision* (John Wiley and Sons, New York, 1957).

177. Middleton, W.E.K., *Vision Through the Atmosphere* (Univ. of Toronto Press, 1952).

182. Minnaert, M., *The Nature of Light and Color in the Open Air* (Dover Pub., Inc., New York, 1954), trans. by H.M. Kremer-Priest, revised by K.E. Brian Jay.

184. Moon, P., "A system of photometric concepts," J. Opt. Soc. Am. 32, 348 (1942).

185. Moon, P., *The Scientific Basis of Illuminating Engineering* (Dover Pub., Inc., New York, 1961), rev. ed..

186. Moon, P., and Spencer, D.E., "Illumination from a non-uniform sky," Illum. Eng. 37, 707 (1942).

187. Moon, P., and Spencer, D.W., "Theory of the photic field," Jour. Franklin Inst. 255, 33 (1953).

192. Parke, N.G., *Statistical Optics: I Radiation* (M.I.T. Research Laboratory of Electronics Report 70, Cambridge, Mass., 1948).

193. Parke, N.G., *Statistical Optics: II Mueller Phenomenological Algebra* (M.I.T. Research Laboratory of Electronics Report 119, Cambridge, Mass. 1949).

194. Parke, N.G., "Optical algebra," J. Math. and Phys. 28, 131 (1949).

206. Pivovonsky, M., and Nagel, M.R., *Tables of Blackbody Radiation Functions* (Macmillan Co., New York, 1961).

208. Pontrjagin, L., *Topological Groups* (Princeton Univ. Press, Princeton, N.J., 1946).

210. Preisendorfer, R.W., *Lectures on Photometry, Hydrologic Optics, Atmospheric Optics* (Lecture Notes, vol. I, Visibility Laboratory, Scripps Inst. of Ocean., University of California, San Diego, Fall 1953).

211. Preisendorfer, R.W., *A Preliminary Investigation of the Transient Radiant Flux Problem* (Lecture Notes, vol. II, Visibility Laboratory, Scripps Inst. of Ocean., University of California, San Diego, February 1954).

216. Preisendorfer, R.W., "A mathematical foundation for radiative transfer theory, "J. Math. and Mech. 6, 685(1957).

233. Preisendorfer, R.W., "Invariant imbedding relation for the principles of invariance," Proc. Nat. Acad. Sci. 44, 320 (1958).

234. Preisendorfer, R.W., "Functional relations for the R and T operators on plane-parallel media," Proc. Nat. Acad. Sci. 44, 323 (1958).

235. Preisendorfer, R.W., "Time-dependent principles of invariance," Proc. Nat. Acad. Sci. 44, 328 (1958).

251. Preisendorfer, R.W., *Radiative Transfer on Discrete Spaces* (Pergamon Press, New York, 1965).

264. Rickart, C.E., *General Theory of Banach Algebras* (D. Van Nostrand, New York, 1960).

296. Thekaekara, M.P., "The solar constant and spectral distribution of solar radiant flux," Solar Energy 9, 7 (1965).

298. Tyler, J.E., "Radiance distribution as a function of depth in an underwater environment," Bull. Scripps Inst. Ocean. 7, 363 (1960).

311. Walsh, J.W.T., *Photometry* (Dover Pub., Inc., New York, 1965).

☆ U.S. GOVERNMENT PRINTING OFFICE: 1976—677-882/ 37 REGION NO. 8